The EMBO Journal vol.13 no.14 pp.3245-3260, 1994

Isolation of high affinity human antibodies directly from large synthetic repertoires

Andrew D.Griffiths[1], Samuel C.Williams[2], Oliver Hartley[1], Ian M.Tomlinson[1], Peter Waterhouse[1,6], William L.Crosby[1,5], Roland E.Kontermann[1], Peter T.Jones[1], Nigel M.Low[2], T.John Allison[2], Terence D.Prospero[1], Hennie R.Hoogenboom[1,4], Ahuva Nissim[1], Jonathan P.L.Cox[1,7], Jacqueline L.Harrison[1], Manuela Zaccolo[3], Ermanno Gherardi[3] and Greg Winter[1,2]

[1]MRC Centre for Protein Engineering, [2]MRC Laboratory of Molecular Biology and [3]ICRF Cell Interactions Laboratory, MRC Centre, Hills Road, Cambridge CB2 2QH and [4]Cambridge Antibody Technology Ltd, The Science Park, Melbourn, Cambridgeshire SG8 6EJ, UK

Present address: [5]National Research Council of Canada, Plant Molecular Genetics, 110 Gymnasium Road, Saskatoon, Saskatchewan, Canada S7N-0W9, [6]CSIRO Division of Plant Industry, GPO Box 1600, Canberra ACT 2601, Australia and [7]University of Bath, School of Chemistry, Claverton Down, Bath BA2 7AY, UK

Communicated by G.Winter

Antibody fragments of moderate affinity (~μM) can be isolated from repertoires of ~10^8 immunoglobulin genes by phage display and rounds of selection with antigen, and the affinities improved by further rounds of mutation and selection. Here, as an alternative strategy, we attempted to isolate high affinity human antibodies directly from large repertoires. We first created highly diverse repertoires of heavy and light chains entirely *in vitro* from a bank of human V gene segments and then, by recombination of the repertoires in bacteria, generated a large (close to 6.5×10^{10}) synthetic repertoire of Fab fragments displayed on filamentous phage. From this repertoire we isolated Fab fragments which bound to a range of different antigens and haptens, and with affinities comparable with those of antibodies from a secondary immune response in mice (up to 4 nM). Although the VH-26 (DP-47) segment was the most commonly used segment in both artificial and natural repertoires, there were also major differences in the pattern of segment usage. Such comparisons may help dissect the contributions of biological mechanisms and structural features governing V gene usage *in vivo*.

Key words: high affinity/human antibodies/phage

Introduction

The display on the surface of filamentous bacteriophage (Smith, 1985) of antibody fragments (McCafferty *et al.*, 1990; Barbas *et al.*, 1991; Breitling *et al.*, 1991; Garrard *et al.*, 1991; Hoogenboom *et al.*, 1991) by fusion to a minor coat protein of phage (pIII), and selection of the phage with antigen, has provided a powerful means of making antibodies of predefined binding specificity from V gene repertoires (for review see Winter *et al.*, 1994). Starting from repertoires derived from the V genes of an immune response, antibody fragments have been readily isolated with high affinity (Clackson *et al.*, 1991), and even with neutralizing activities against virus (Barbas *et al.*, 1992a,b).

The display and selection of antibody fragments on the surface of phage mimics immune selection (for review see Marks *et al.*, 1992a), and antibodies have also been isolated without immunization, from repertoires of V genes rearranged *in vivo* (Marks *et al.*, 1991) or *in vitro* (Hoogenboom and Winter, 1992). The same phage repertoire may be used to generate many different binding specificities, including those that are difficult to raise by immunization, for example against self-antigens (Griffiths *et al.*, 1993; Marks *et al.*, 1993) or proteins of the lumen of the endoplasmic reticulum (Nissim *et al.*, 1994). Antibodies from such 'single pot' libraries appear to be highly specific and have been used for 'on-line detection' of antigen in biosensors (M.Malmqvist, unpublished), and for Western blotting and epitope mapping (Nissim *et al.*, 1994). But the affinities were not high (~μM), as expected for primary repertoires of 10^7-10^8 clones.

It should also be possible to isolate high affinity antibodies (<10 nM) directly from primary repertoires of sufficiently large diversity and size (Perelson and Oster, 1979), but the size and diversity of repertoires required is not known. The size and 'shape' of natural repertoires is limited not only by the number of B-lymphocytes (~5×10^8 in mouse and ~10^{12} in human), and the number of lymphocytes in each clone, but also by the processes of clonal deletion and anergy that lead to self tolerance (for review see Nossal, 1993).

Here we constructed a highly diverse combinatorial repertoire entirely *in vitro*, using V gene segments as building blocks. We used the vast majority of heavy and light chain segments used *in vivo* and encoded part, or all, of each CDR3 loop by random sequence. To make the repertoire as large as possible we infected bacteria harbouring a 'donor' heavy chain repertoire (on a plasmid) with an 'acceptor' light chain repertoire (on phage). The two chains were combined on the same (phage) replicon within the bacterium by Cre catalysed recombination at *loxP* sites. This process, termed 'combinatorial infection' (Waterhouse *et al.*, 1993), generates a large number of heavy and light chain combinations, potentially as large as the number of bacteria that have been infected. The repertoire was characterized by the properties of the selected Fab fragments.

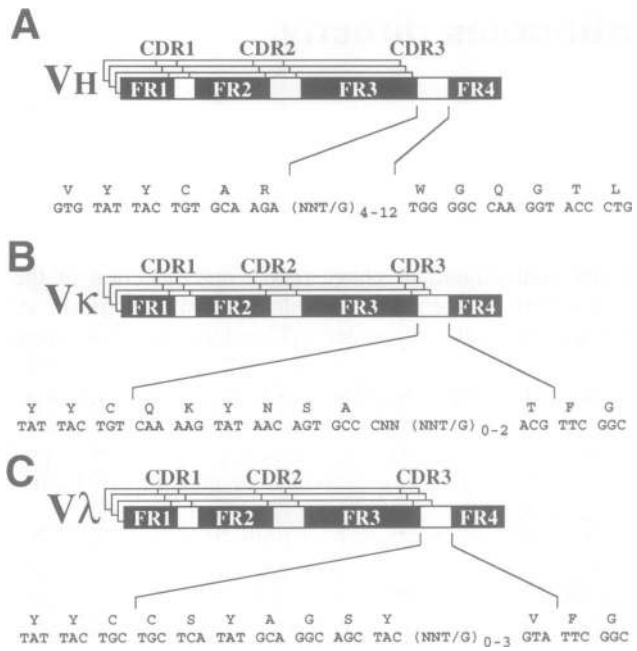

Fig. 1. Construction of synthetic heavy and light chain repertoires. (**A**) The repertoire of heavy chains ($>10^8$ different clones) was built from 49 cloned V_H segments (Tomlinson *et al.*, 1992; Nissim *et al.*, 1994), with CDR3 loops of four to 12 residues of random sequence. (**B**) The repertoire of κ light chains (9×10^4 clones) was built from 26 cloned V_κ segments (Cox *et al.*, 1994) with CDR3 loops of eight to 10 residues that included one, two or three residues of random sequence in all cases. DPK-4 is shown by way of an example. (**C**) The repertoire of λ light chains (7.4×10^5 clones) was built from 21 cloned V_λ segments (Williams and Winter, 1993), with CDR3 loops of eight to 13 residues that included zero, one, two, three, four or five residues of random sequence. DPL-12 is shown by way of example. CDR, complementarity determining region; FR, framework region. Neither the single segment of the $V_\kappa7$ subgroup, nor the few segments from the V_λ families 4, 5 or 6, were included in the repertoire. Segments rarely used *in vivo*, for example light chain segments (DPK-2, -7, -10, -12, -17, -19, -20, -23 and -25) from the distal V_κ locus (Cox *et al.*, 1994), were represented equally in the synthetic repertoire. Some of the V_H gene segments (DP-1, -12, -30, -39, -40, -44, -45 and -69) included in the repertoire are now known to be located on chromosomes 15 or 16, and therefore outside the functional locus (Tomlinson *et al.*, 1994).

Results

Repertoire size and diversity

Heavy and light chain repertoires were built from the majority of human V gene segments as described in Figure 1. The light chain repertoire was cloned into 'acceptor' fd phage (tetracycline resistant), with a 'dummy' heavy chain (Figure 2A), and the heavy chain repertoire into 'donor' plasmid (ampicillin resistant) (Figure 2B). The repertoires, $>10^8$ heavy chains and $>8\times10^5$ light chains, were combined by infection of *Escherichia coli* harbouring the donor heavy chains with fd phage carrying the light chains. If every heavy chain were paired with every light chain, this would generate a repertoire of $>8\times10^{13}$ antibodies. The culture was then co-infected with bacteriophage P1 (chloramphenicol resistant), which provides the Cre recombinase (Waterhouse *et al.*, 1993), leading to 6.5×10^{10} colonies resistant to ampicillin, tetracycline and chloramphenicol (for summary, see Table I). After growth of the cultures, the fd phage were used to infect *E.coli*. Twenty-eight per cent of the acceptor phage were shown to have acquired a heavy chain from the donor vector (see

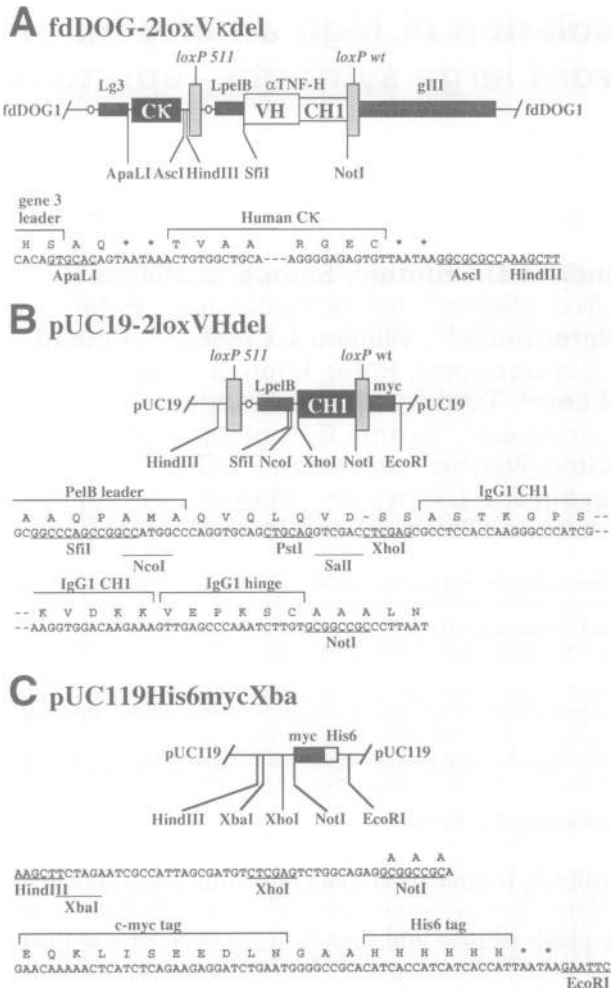

Fig. 2. Vectors. (**A**) The fd phage 'acceptor' vector fdDOG-2loxVκdel. Light chain genes ($V_\kappa-C_\kappa$ and $V_\lambda-C_\lambda$) are cloned into this vector as *Apa*LI–*Asc*I fragments. (**B**) The plasmid 'donor' vector pUC19-2loxVHdel. Heavy chain variable region genes (V_H genes) are cloned into this vector as *Nco*I–*Xho*I fragments. (**C**) The phagemid expression vector pUC119His6mycXba. Heavy and light chain genes encoding Fabs are cloned into this vector as *Xba*I–*Not*I fragments. Other features are marked as follows: sequences encoding Lg3, gene III leader sequence; LpelB, pelB leader sequence; Cκ, human kappa light chain constant region; VH, heavy chain variable region; CH1, first heavy chain constant domain (human Cμ1 in fdDOG-2loxVκdel and human Cγ1 in pUC19-2loxVHdel); αTNF-H, the V_H gene of the mouse anti-TNFα antibody mAb32 (Rathjen *et al.*, 1992) linked to a human Cμ1 constant domain gene; gIII, fd phage gene III; *loxP* wt, wild-type *loxP* site (Hoess *et al.*, 1982); *loxP 511*, a mutant *loxP* site with a single point mutation (Hoess *et al.*, 1986); myc, peptide from c-*myc* recognized by the monoclonal antibody 9E10 (Munro and Pelham, 1986); His6, six histidines.

Materials and methods). As there are multiple copies of plasmid and phage replicons in each bacterial cell when Cre recombinase is delivered by phage P1 infection and at least 60 phage are produced per bacterium after overnight growth, we believe that each bacterium should yield at least one phage containing the heavy chain from the donor vector and therefore we estimate that the repertoire contained close to 6.5×10^{10} different phage antibodies, with up to 60 copies of each.

Specificity of selected antibodies

The repertoire of phage was selected with a range of antigens, leading to isolation of binding specificities, as

Table I. Determination of repertoire size

Sample Point	Total no. of colony forming units ($\times 10^{10}$)							Significance	
	No antibiotic	amp^R	tet^R	chlor^R	amp^R + tet^R	tet^R + chlor^R	amp^R + tet^R + chlor^R		
1		0.17						size of pUC19-2loxVHlib inoculum from frozen stock:	1.7×10^9
2	3.4	2.3						size of pUC19-2loxVHlib inoculum from overnight culture:	2.3×10^{10}
3	65 (λ)		10 (λ)		19 (λ)			no. of *E.coli* containing pUC19-2loxVHlib infected with fdDOG-2loxVλlib:	1.9×10^{11} (λ)
	73 (κ)		11 (κ)		8.0 (κ)			no. of *E.coli* containing pUC19-2loxVHlib infected with fdDOG-2loxVκlib:	8.0×10^{10} (κ)
4	9.5 (λ)	22 (λ)	4.5 (λ)	6.0 (λ)	3.0 (λ)	2.0 (λ)	3.0 (λ)	no. of *E.coli* containing pUC and fdDOG replicons co-infected with phage P1:	3.0×10^{10} (λ)
	29 (κ)	28 (κ)	7.5 (κ)	28 (κ)	3.0 (κ)	4.5 (κ)	3.5 (κ)		3.5×10^{10} (κ)
5								titre of fdDOG phage (t.u.) immediately after P1 infection and centrifugation:	9.9×10^9 (λ)
									1.2×10^{10} (κ)
6	31 (λ)						16 (λ)	no. of viable *E.coli* containing pUC, fdDOG and P1 replicons after 24 h:	1.6×10^{11} (λ)
	35 (κ)						9.5 (κ)		9.5×10^{10} (κ)
7								titre of fdDOG phage (t.u.) after 24 h:	1.1×10^{13} (λ)
									3.0×10^{13} (κ)

Sample point, see Materials and methods; amp^R, ampicillin resistant; tet^R, tetracycline resistant; chlor^R, chloramphenicol resistant.

summarized in Table II. The repertoire was selected on all antigens and haptens by panning on antigen coated immunotubes; for the haptens NIP and fluorescein, the phage were also captured with biotinylated NIP–BSA and FITC–BSA and streptavidin-coated paramagnetic beads. To follow the selection process, *E.coli* were infected with the eluted phage after each round, and the phage screened for binding to antigen by ELISA, either as a 'polyclonal' population or as phage clones. DNA encoding the Fab fragments was amplified using the polymerase chain reaction from the population of phage after two to four rounds of selection and recloned into plasmid (Figure 2C) for expression of soluble Fab fragments.

We focused on the characterization of Fab fragments with specificities against the hapten–BSA conjugates of NIP (3-iodo-4-hydroxy-5-nitrophenyl-acetate) and FITC (fluorescein 5-isothiocyanate), as it would allow comparisons with the natural immune response to both haptens in mice, and facilitate the measurement of binding affinities to the free haptens (by fluorescence quench titrations). The fragments against hapten–BSA were first screened for binding to the free hapten by competition with fluorescein or NIP–caproic acid (NIP–CAP), or for binding to BSA. Most of the Fab fragments bound to hapten–BSA, but some also bound to BSA, or to free hapten.

We also characterized Fab fragments against four antigens of the kringle-serine protease family—hepatocyte growth factor/scatter factor (HGF/SF), plasmin, urokinase-type plasminogen activator (u-PA), and tissue-type plasminogen activator (t-PA)—as members of a family of related proteins. Binding of 'polyclonal' phage was detected after three rounds of selection and proved to be specific, despite homologies between the members of this family (Figure 3). We also characterized Fab fragments against the mouse monoclonal antibody NQ11/7.22 (Griffiths *et al.*, 1984), as the binding could be mapped to the variable regions, to the Fc region and to the remaining portions of the monoclonal antibody (Table III).

Table II. Binding specificities isolated from the library

Haptens
3-iodo-4-hydroxy-5-nitrophenyl-acetate (NIP)
Fluorescein
2-phenyloxazol-5-one (phOx)
N-(Carboxymethyl)-4-[(*p*-nitrobenzyloxyphosphoryl)-butyramide]
N-[2-hydroxy-3-(4-nitrophenyl)]-propyl-L-prolyl-glycine
Foreign antigens
Serum albumin (BSA) (bovine)
Tubulin (bovine)
Calmodulin (bovine)
Hepatocyte growth factor/scatter factor (HGF/SF) (murine)
Monoclonal antibody NQ11/7.22 (murine)
FixL (*Bradyrhizobium japonicum*)
Acetolactate synthase (ALS) (*Brassica napus*)
Lol pII (*Lolium perenne*)
Gene product of CDC4 (*Schizosaccharomyces pombe*)
Gene product of CDC8 (*Schizosaccharomyces pombe*)
Maltose binding protein (*E.coli*)
gp120 (HIV-1)
gp11 (T4 phage)
gp9 (T4 phage)
Human antigens
Tumour necrosis factor α (TNFα)
Thyroglobulin
High affinity IgE receptor (FcεRI)
Tissue-type plasminogen activator (t-PA)
Urokinase-type plasminogen activator (u-PA)
Plasmin
Carcinoembryonic antigen (CEA)
c-erb B2
Tau40
Elongation factor 1α (EF-1α)
Calreticulin
Calnexin
Ferritin light chain
Factor VIII
U1 snRNA
U1A protein
U1C protein

Segment usage of selected heavy and light chains

The phage selected from the large (6.5×10^{10} clones) repertoire by binding to antigen were characterized by DNA sequencing. Many of the clones were different: 137 unique antigen-binding Fab fragments (with differing light or heavy chain protein sequences) were identified from a total of 215 clones sequenced (Table III). A range of V gene segments was seen: 17 of the 49 V_H segments, 10 of the 26 V_κ segments and nine of the 21 V_λ segments

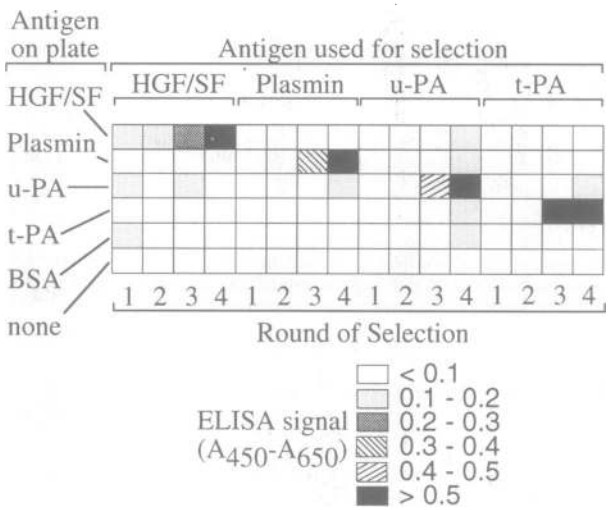

Fig. 3. Specificity of polyclonal phage against kringle serine protease family. 'Polyclonal' phage after one, two, three and four rounds of selection on either HGF/SF, plasmin, u-PA or t-PA, were assayed by ELISA for binding to the other members of the family, and to BSA.

(see Figure 4). Each of the major heavy and light chain families were represented (Chuchana *et al.*, 1990; Kabat *et al.*, 1991; Williams and Winter, 1993), but not all the minor families. Thus V_H segments were seen from families 1, 3, 4 and 5, but not 2 and 6; V_κ segments from subgroups 1, 2, 3 and 4, but not from 6; and V_λ segments from families 1, 2, 3, 7 and 8, but not 9. The heavy chain V_H segment DP-45 (included in the repertoire, but located on chromosome 16 outside the major locus on chromosome 14) was found in two Fab fragments (Table III and Figure 4A) binding to NIP–BSA.

Some V gene segments (V_H segments DP-7, DP-38, DP-47 and DP-67; V_κ segment DPK-15; and V_λ segment DPL-3) were seen frequently in the synthetic chains: of these only the V_H segment DP-47 is common in natural antibodies. Conversely, some segments (like V_H segments DP-63 and DP-71; V_κ segments DPK-1 and DPK-21; and V_λ segments DPL-5 and DPL-23) that are common in natural antibodies were not seen in the synthetic chains (Figure 4). Thus, except for DP-47, the pattern of usage of the segments from the synthetic repertoire, summed

Table III. CDR3 sequences and germline V gene segments from antigen-binding clones

Antigen[a]	Clone	Heavy chains[b]			Light chains[c]			Selection method[d]	Round[e]	No. of copies[f]
		Family	Segment	CDR3[g]	Family	Segment	CDR3[g]			
NIP–BSA	G09	VH3	DP-38	AGTL	Vλ1	DPL-3	AAWDDSLV	M	4	1
NIP–BSA	E01	VH3	DP-38	AGTL	Vκ2	DPK-12	MQSIQLPT	M	3/4	2
NIP–BSA	G10	VH3	DP-38	AGTL	Vκ2	DPK-12	MQSIQLPAT	M	4	1
NIP–BSA	G04	VH3	DP-38	AGTL	Vλ1	DPL-3	AAWDDGLSLV	M	4	1
NIP–BSA	H08	VH3	DP-38	AGTL	Vλ1	DPL-3	AAWDDSLSGV	M	4	1
NIP–BSA	G07	VH3	DP-38	AGTL	Vλ3	DPL-16	NSRDSSGSVRV	M	4	1
NIP–BSA	C09	VH3	DP-38	GGKD	Vλ7	DPL-18	LLYYGGAYV	Im	4	1
NIP–BSA	F03	VH1	DP-10	GGRL	Vλ3	DPL-16	NSRDSSGVSRV	M	3	1
NIP–BSA	E07	VH3	DP-38	GGTQ	Vλ1	DPL-3	AAWDDSLV	M	3	1
NIP–BSA	H05	VH3	DP-38	GGTQ	Vλ1	DPL-3	AAWDDSLPYV	M	4	1
NIP–BSA	H03	VH3	DP-38	HGQH	Vλ1	DPL-3	AAWDDSLCPEFV	M	4	1
NIP–BSA	H01	VH3	DP-38	KGSE	Vλ1	DPL-3	AAWDDSLAWFV	M	4	1
NIP–BSA	A12	VH3	DP-47	KGWS	Vλ1	DPL-4	LAWDTSPRWV	Im	3	1
NIP–BSA	A10	VH3	DP-47	KGWS	Vλ1	DPL-2	TAWDDSLAVV	Im	3	1
NIP–BSA	D08	VH3	DP-47	KGWS	Vλ3	DPL-16	NSRDSSGNHRV	Im	4	1
NIP–BSA	G02	VH3	DP-49	LGKA	Vκ3	DPK-22	QQYGSSQRT	M	4	1
NIP–BSA	E06	VH3	DP-38	NGYF	Vλ1	DPL-3	AAWDDSLRLV	M	3	1
NIP–BSA	D03	VH3	DP-49	PRGY	Vλ1	DPL-3	AAWDDSLRLV	Im	4	1
NIP–BSA	B02	VH3	DP-46	MYMRS	Vκ2	DPK-18	MQGTHWRPT	Im	3	1
NIP–BSA	E02	VH3	DP-46	MYRSV	Vκ2	DPK-18	MQGKHWPLT	M	3	1
NIP–BSA	A06	VH3	DP-42	NGGHV	Vλ1	DPL-3	AAWDDSLGF	Im	3	1
NIP–BSA	D05	VH3	DP-47	PAGSR	Vκ2	DPK-18	MQGTHRRAT	Im	4	1
NIP–BSA	A04	VH3	DP-38	PATRS	Vκ2	DPK-15	MQALQTPLT	Im	3	1
NIP–BSA	F06	VH3	DP-47	PFATF	Vκ2	DPK-18	MRGTHRRAT	M	3	1
NIP–BSA	E08	VH3	DP-51	PFLAH	Vκ2	DPK-18	MQGTHWHPT	M	3	1
NIP–BSA	C05	VH3	DP-32	PLGAH	Vκ2	DPK-15	MQALQSPT	Im	4	1
NIP–BSA	E12	VH3	DP-47	PMRGV	Vκ2	DPK-18	MQGTHRRAT	M	3	1
NIP–BSA	E05	VH3	DP-38	PNGDQ	Vλ1	DPL-3	AAWDDSLAFV	M	3	1
NIP–BSA	E04	VH3	DP-38	POTRR	Vκ2	DPK-15	MQALQTPT	M	3	1
NIP–BSA	A08	VH3	DP-47	PRLPR	Vκ1	DPK-9	QQSYSTRT	Im	3	1
NIP–BSA	E10	VH5	DP-73	PSGNV	Vκ2	DPK-19	MQGTHWPFT	M	3	1
NIP–BSA	A05	VH1	DP-25	QGLRN	Vκ2	DPK-15	MQALQTPLT	Im	3	1
NIP–BSA	D06	VH3	DP-47	RGHKA	Vκ2	DPK-18	MQGTHWPAT	Im	4	1
NIP–BSA	D02	VH3	DP-51	SRGDS	Vλ1	DPL-3	AAWDDSLRSV	Im	4	1
NIP–BSA	F01	VH3	DP-47	TFSPQ	Vκ2	DPK-18	MQGTHRRAT	M	3	1
NIP–BSA	B03	VH3	DP-47	SFRRNL	Vλ1	DPL-3	AAWDDSLLV	Im	3	1
NIP–BSA	A11	VH3	DP-58	SFRRNL	Vκ3	DPL-16	DSWDNSLVSPV	Im	3	1
NIP–BSA	C04	VH3	DP-38	PGYRGTR	Vκ2	DPK-15	MQALQSPT	Im	4	2
NIP–BSA	D07	VH3	DP-38	PGYRGTR	Vκ2	DPK-12	MQSIQLPT	Im	4	1
NIP–BSA	D01	VH3	DP-38	PGYRGTR	Vκ2	DPK-15	MQALQSPAT	Im	4	1
NIP–BSA	C10	VH3	DP-38	PGYRGTR	Vκ2	DPK-15	MQALQTPVT	Im	4	1
NIP–BSA	C11	VH3	DP-38	PGYRGTR	Vλ1	DPL-3	AAWDDSLSAYV	Im	4	1
NIP–BSA	F04	VH3	DP-45	RAINGQR	Vλ3	DPL-16	NSRDSSGRVNV	M	3	1
NIP–BSA	B04	VH3	DP-47	RRGSTRY	Vκ2	DPK-15	MQALRTRT	Im	3	1
NIP–BSA	F05	VH3	DP-38	VNSRFAT	Vλ3	DPL-16	NSRDSSGVSRV	M	3	1
NIP–BSA	E11	VH4	DP-67	IKFRSSSI	Vκ2	DPK-19	MQGTHWPFT	M	3	1
NIP–BSA	H06	VH4	DP-67	SFAKAFDY	Vλ1	DPL-3	AAWDDSLPYV	M	4	1

Table III. *Continued*

Antigen[a]	Clone	Heavy chains[b]			Light chains[c]			Selection method[d]	Round[e]	No. of copies[f]
		Family	Segment	CDR3[g]	Family	Segment	CDR3[g]			
NIP–BSA	E09	VH4	DP-67	SFAKAFDY	Vλ3	DPL-16	NSRDSSGSVRV	M	3	1
NIP–BSA	C02	VH1	DP-7	SKRTSFDY	Vκ2	DPK-18	MQGTHWHPT	Im	4	1
NIP–BSA	G08	VH3	DP-47	SLFSKFDY	Vλ3	DPL-16	NSRDSSGVSRV	M	4	1
NIP–BSA	A07	VH3	DP-47	SVLSLFDY	Vλ1	DPL-3	AAWDDSLFYPV	Im	3	1
NIP–BSA	C03	VH3	DP-45	SYMRGMRN	Vλ3	DPL-16	NSRDSSGNHRV	Im	4	1
NIP–BSA	A02	VH3	DP-42	HRRAYYMIP	Vκ2	DPK-18	MQGTHWPVT	Im	3	1
NIP–BSA	A09	VH4	DP-67	IGKLSQPTS	Vκ2	DPK-18	MQGTHWRPT	Im	3	1
NIP–BSA	E03	VH3	DP-47	RSGVRMLID	Vκ2	DPK-18	MQGTHWRT	M	3	1
NIP–BSA	A03	VH3	DP-42	HLRWASGGPR	Vκ2	DPK-18	MQGTHWRT	Im	3	1
NIP–BSA	G11	VH3	DP-47	PLNSKKNTTTQ	Vλ1	DPL-3	AAWDDSLFYV	M	4	1
NIP–BSA	C12	VH3	DP-49	GRTWSPSLPPLR	Vκ2	DPK-12	MQSIQLPLT	Im	4	1
NQ11 (Fc)	NML7	VH3	DP-47	KWGG	Vλ1	DPL-2	AAWDDSLLGSV	Im	2	1
NQ11 (Fc)	NML9	VH1	DP-14	GTGLDG	Vλ2	DPL-10	CSYAGSSYV	Im	2	1
NQ11 (Fc)	NML8	VH3	DP-47	KFGNNM	Vλ3	DPK-23	QQDYNLLT	Im	2	1
NQ11 (Fv)	NML1	VH3	DP-47	ASSPFVLQ	Vλ8	DPL-21	VLYMGSGSAV	Im	2/3/4	25
NQ11	NML3	VH1	DP-7/3[g]	YKSLSFDY	Vκ2	DPK-13	MQRIEFPNT	Im	2	1
NQ11	NML5	VH1	DP-10	AANYSKAHI	Vλ1	DPL-2	AAWDDSLACAV	Im	2/3	4
NQ11 (Fv)	NML2	VH3	DP-47	RSWDGGMVD	Vκ1	DPK-5	QQANSFRT	Im	2	1
NQ11 (Fc)	NML11	VH3	DP-3	SKLWVTFDY	Vλ1	DPL-8/2[g]	AAWDDSLSRPV	Im	3	1
NQ11 (Fc)	NML6	VH3	DP-3	SKLWVTFDY	Vλ1	DPL-2	AAWDDSLSRPV	Im	2/3/4	34
NQ11	NML4	VH3	DP-3	AKQSGVECLT	Vλ1	DPL-3	AAWDDSLYNV	Im	2	2
NQ11 (Fc)	NML10	VH3	DP-3	SKYPLAWTLS	Vλ1	DPL-2	AAWDDSLNRNV	Im	2	1
FITC–BSA	B01	VH3	DP-47	ALRR	Vκ2	DPK-15	MQVLQTRT	Im	3	1
FITC–BSA	B06	VH3	DP-47	GGRV	Vκ2	DPK-15	MQALQTRT	Im	3	1
FITC–BSA	A03	VH3	DP-47	IGQF	Vλ1	DPL-3	AAWDDSLAFV	Im	3/4	4
FITC–BSA	D10	VH3	DP-47	KAKT	Vκ2	DPK-15	MQALQTRT	Im	4	1
FITC–BSA	G06	VH3	DP-47	KSAI	Vκ2	DPK-15	MQALQTRT	M	4	1
FITC–BSA	H03	VH3	DP-47	KSRW	Vκ1	DPK-9	QQSYSTRM	M	4	1
FITC–BSA	D12	VH3	DP-47	KSTV	Vκ2	DPK-15	MQALRTRT	Im	4	1
FITC–BSA	A08	VH3	DP-47	LNRK	Vκ2	DPK-15	MQALQTRT	Im	3/4	4
FITC–BSA	D08	VH3	DP-47	RHGS	Vκ2	DPK-15	MQALRTRT	Im	4	1
FITC–BSA	G07	VH3	DP-47	RKRH	Vκ2	DPK-15	MQALQTLT	M	4	1
FITC–BSA	H05	VH3	DP-47	RSKT	Vκ2	DPK-15	MQALQTRT	M	4	1
FITC–BSA	H02	VH3	DP-47	RWSF	Vλ1	DPL-3	AAWDDSLV	M	4	1
FITC–BSA	E06	VH3	DP-47	AKFRL	Vκ2	DPK-15	MQALRTRT	M	3	1
FITC–BSA	E11	VH3	DP-47	AYHGR	Vκ2	DPK-15	MQALQTRT	M	3	1
FITC–BSA	C02	VH3	DP-47	GKVLG	Vκ2	DPK-15	MQALQTPT	Im	4	1
FITC–BSA	B02	VH3	DP-47	GKVLG	Vκ2	DPK-15	MRALQTPT	Im	3/4	2
FITC–BSA	E07	VH3	DP-47	GSSRT	Vλ1	DPL-3	AAWDDSLPGYV	M	3	1
FITC–BSA	E08	VH3	DP-47	KRMDG	Vκ2	DPK-15	MQALQTRT	M	3	1
FITC–BSA	A10	VH1	DP-10	LKRGH	Vλ1	DPL-3	AAWDDSLGFV	Im	3	1
FITC–BSA	D09	VH3	DP-47	LRREY	Vκ2	DPK-15	MQALRTRT	Im	4	1
FITC–BSA	G12	VH3	DP-47	RAGRD	Vλ1	DPL-3	AAWDDSLFLV	M	4	1
FITC–BSA	D03	VH3	DP-47	LKSAYK	Vκ2	DPK-15	MQALQTPT	Im	4	1
FITC–BSA	C12	VH3	DP-47	LNVRPK	Vκ2	DPK-15	MQALQTRT	Im	4	1
FITC–BSA	B10	VH3	DP-47	SRGKSM	Vκ2	DPK-15	MQALRTRT	Im	3	1
FITC–BSA	E09	VH3	DP-47	IRFRNAT	Vκ2	DPK-15	MQALRTRT	M	3	1
FITC–BSA	B09	VH3	DP-47	LKTSTPV	Vκ2	DPK-15	MRALQTPT	Im	3/4	2
FITC–BSA	G10	VH3	DP-47	LSRAFTM	Vκ2	DPK-15	MQALRTRT	M	4	1
FITC–BSA	E03	VH3	DP-47	LSRAFTM	Vκ2	DPK-15	MQALQTRT	M	3	1
FITC–BSA	B07	VH4	DP-67	AQRKYFDY	Vκ2	DPK-12	MQSIQLRT	Im	3	1
FITC–BSA	D01	VH4	DP-67	DLRKHFDY	Vκ1	DPK-9	QQSYSTRT	Im	4	1
FITC–BSA	E10	VH4	DP-67	DRWRVFDY	Vλ1	DPL-2	AAWDDSLSIV	M	3	1
FITC–BSA	A09	VH1	DP-14	KGLRLFDY	Vλ1	DPL-3	AAWDDSLV	Im	3	1
FITC–BSA	G08	VH3	DP-58	KKYQSAAR	Vκ2	DPK-19	MQGTHWPT	M	4	1
FITC–BSA	B03	VH4	DP-67	KTRRRFDY	Vκ2	DPK-15	MQALHTRT	Im	3	1
FITC–BSA	C03	VH4	DP-67	KTRRRFDY	Vκ2	DPK-15	MQALQTRT	Im	4	1
FITC–BSA	G11	VH3	DP-47	PYAKRFDY	Vκ2	DPK-15	MQALQTRT	M	4	1
FITC–BSA	G03	VH3	DP-47	RFARSFDY	Vλ3	DPL-16	NSRDSSGSV	M	4	3
FITC–BSA	A04	VH4	DP-67	RSFVGYEI	Vλ1	DPL-3	AAWDDSLV	Im	3	2
FITC–BSA	D06	VH4	DP-67	RWGRTFDY	Vκ2	DPK-15	MQALQTRT	Im	4	1
FITC–BSA	C06	VH1	DP-7	SQKRLITG	Vκ2	DPK-15	MQALQTRT	Im	4	1
FITC–BSA	C01	VH1	DP-7	SQKRLITG	Vκ3	DPK-22	QQYGSSPYT	Im	4	1
FITC–BSA	H04	VH4	DP-67	SRKRAFDY	Vκ2	DPK-15	MQALQTRT	M	4	1
FITC–BSA	C08	VH4	DP-67	SWVSGFDY	Vκ1	DPK-9	QQSYSTRT	Im	4	2
FITC–BSA	D04	VH1	DP-7	SYHRTFDY	Vκ1	DPK-5	QQANSFAAT	Im	4	1
FITC–BSA	C05	VH1	DP-7	SYHRTFDY	Vκ1	DPK-5	QQANSFPAT	Im	4	1
FITC–BSA	D05	VH1	DP-7	THSKTFDY	Vλ8	DPL-21	VLYMGSGVYV	Im	4	1
FITC–BSA	B12	VH3	DP-47	TRSSSYGE	Vκ2	DPK-15	MQALRTRT	Im	3	1
FITC–BSA	B04	VH4	DP-66	WSRETNYS	Vλ1	DPL-3	AAWDDSLWSAV	Im	3	1
FITC–BSA	A07	VH3	DP-47	RTRGALPRN	Vλ1	DPL-3	AAWDDSLPRRLV	Im	3	1
FITC–BSA	A02	VH3	DP-47	YRFSAPPRD	Vλ1	DPL-3	AAWDDSLPSGV	Im	3	1
FITC–BSA	E04	VH3	DP-47	RFNRLSPRRA	Vκ2	DPK-15	MQALQTRT	M	3	1
FITC–BSA	B05	VH1	DP-25	SSVMGRVPVM	Vκ2	DPK-15	MQALQTLT	Im	3	1
FITC–BSA	E05	VH3	DP-47	TSGKLHSPRT	Vλ1	DPL-3	AAWDDGLLRV	M	3	1
FITC–BSA	D11	VH5	DP-73	GRGRPSMAYDV	Vλ1	DPL-3	AAWDDSLALV	Im	4	1
FITC–BSA	B08	VH3	DP-47	RSGVSRKVYTI	Vκ2	DPK-15	MQALRTRT	Im	3	1
Plasmin	MP01	VH3	DP-47	MTPQFFDY	Vκ2	DPK-15	MRALQTPT	Im	4	1
Plasmin	MP02	VH4	DP-67	SAYSYFDY	Vλ3	DPL-16	NSRDSSGFQLV	Im	4	1

Table III. *Continued*

Antigen[a]	Clone	Heavy chains[b]			Light chains[c]			Selection method[d]	Round[e]	No. of copies[f]
		Family	Segment	CDR3[g]	Family	Segment	CDR3[g]			
t-PA	MT09	VH1	DP-7	DSGLGDPAL	Vλ2	DPL-11	SSYTSSSTLV	Im	4	1
t-PA	MT03	VH1	DP-7	DSGLGEPAL	Vλ2	DPL-11	SSYTSSSTLG	Im	4	1
t-PA	MT06	VH1	DP-7	DSGLGEPAL	Vλ2	DPL-11	SSYTSSSTLV	Im	4	1
t-PA	MT01	VH1	DP-7	ESGLGDPAL	Vλ2	DPL-11	SSYTSSSTLV	Im	4	1
t-PA	MT07	VH3	DP-47	TSRLKAHPS	Vλ1	DPL-8	QSYDSNLRV	Im	4	1
u-PA	MU02	VH3	DP-47	TSRLEAHPR	Vκ2	DPK-15	MRALQTPT	Im	4	1
u-PA	MU01	VH3	DP-47	TSRLKAHPS	Vκ1	DPK-8	QQLNSYPT	Im	4	1
u-PA	MU03	VH3	DP-47	TSRLKAHPS	Vλ3	DPL-16	NSRDSSGFQLV	Im	4	1
HGF/SF	MH10	VH3	DP-47	GRQSRL	Vκ1	DPK-5	QQANSFPIT	Im	4	1
HGF/SF	MH19	VH3	DP-42	KFPHFGD	Vκ1	DPK-8	QQLNSYPT	Im	4	1
HGF/SF	MH22	VH3	DP-42	KFPHFGD	Vκ1	DPK-5	QQANSFPIT	Im	4	4

[a]The region of the monoclonal antibody NQ11/7.22 (NQ11) bound by the Fab is indicated (Fv or Fc region); three Fabs bound neither fragment and therefore probably interacted with the CH1, Cκ or hinge region. Hapten-BSA binding clones listed did not bind BSA alone.
[b]Human germline V_H gene segments are assigned to families as in Tomlinson *et al.* (1992).
[c]Human germline V_κ gene segments (Cox *et al.*, 1994) are assigned to subgroups as in Kabat *et al.* (1991) and human germline V_λ gene segments (Williams and Winter, 1993) are assigned to families as in Chuchana *et al.* (1990).
[d]Im, selected using antigen-coated immunotubes; M, selected using biotinylated antigen and streptavidin-coated paramagnetic beads. Numbers refer to how many rounds of selection the library had undergone when Fabs with the sequence indicated were isolated.
[e]The number of independent clones which were isolated with the same sequence.
[f]CDR3 (complementarity determining region 3) for both heavy and light chains are as defined by Kabat *et al.* (1991).
[g]These genes appear to have been created by cross-overs between two V genes during PCR amplification.

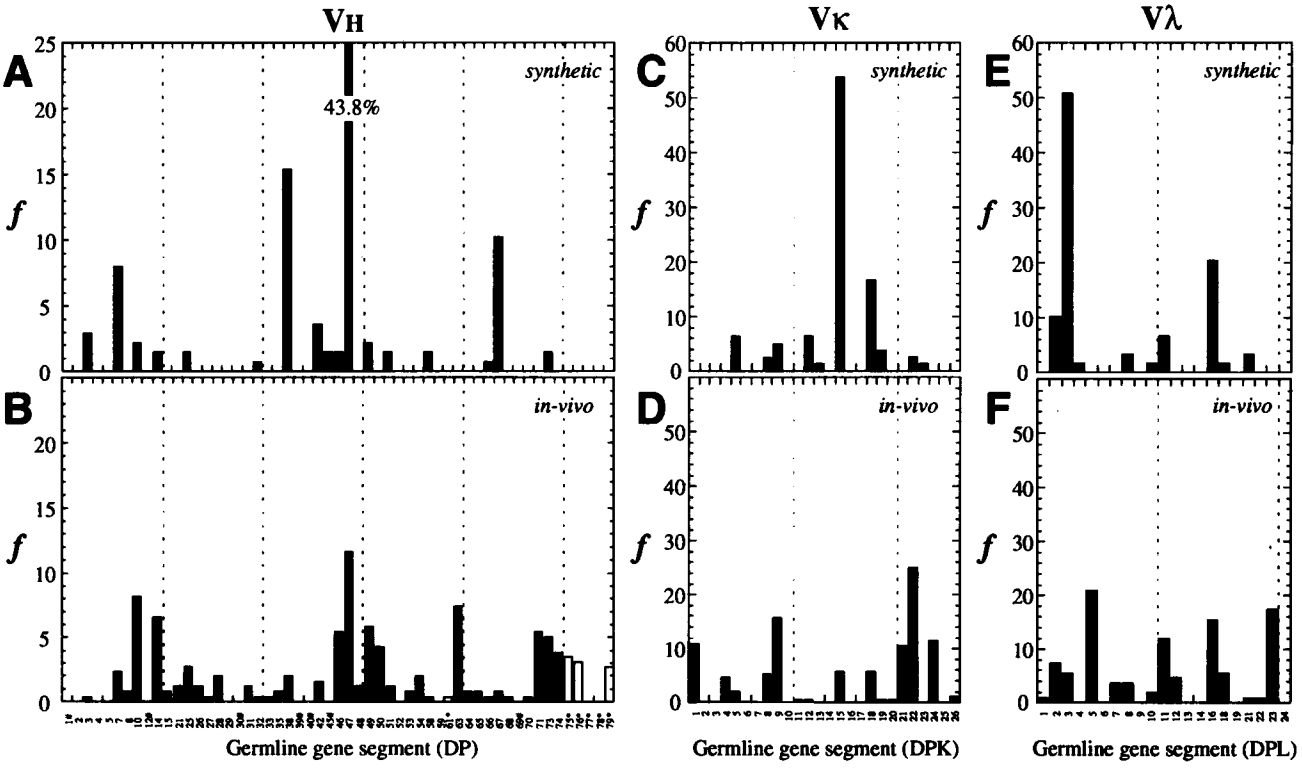

Fig. 4. Use of human germline V gene segments. Frequencies of use of human V_H, V_κ and V_λ segments from the synthetic repertoire (A, C and E), or from natural antibodies (B, D and F). Frequencies (f) are plotted as % of total. V gene usage was compiled for the synthetic antibodies from Table III, and for natural antibodies, from the 292 rearranged V_H genes in the database described in Tomlinson *et al.* (1992), from the 236 rearranged V_κ genes in the database described in Cox *et al.* (1994) and from a database of 110 rearranged V_λ genes taken from the 'Entrez' sequence database (release 8.0; National Center for Biotechnological Information). V_H segments are listed by DP numbers (Tomlinson *et al.*, 1992), V_κ segments by DPK numbers (Cox *et al.*, 1994) and V_λ segments by DPL numbers (Williams and Winter, 1993). All V gene segments listed were included in the synthetic repertoire except those marked (*). V_H gene segments (located on chromosome 15 or 16) which are not used *in vivo* (Tomlinson *et al.*, 1994), but which were included in the synthetic repertoire are indicated (#).

over the limited number of antigens, appears to differ from the usage *in vivo*.

Distribution of CDR3 lengths

We also analysed the lengths and sequences of the CDR3 loops in the selected Fab fragments (Table III; Figure 5). All the heavy chain CDR3 lengths included in the synthetic repertoire were present. The four residue loops almost all include a glycine residue, presumably to make the tight turn. However, the distribution of the heavy chain CDR3 lengths, biased towards short lengths, contrasts with natural antibodies. This appears to be especially characteristic of

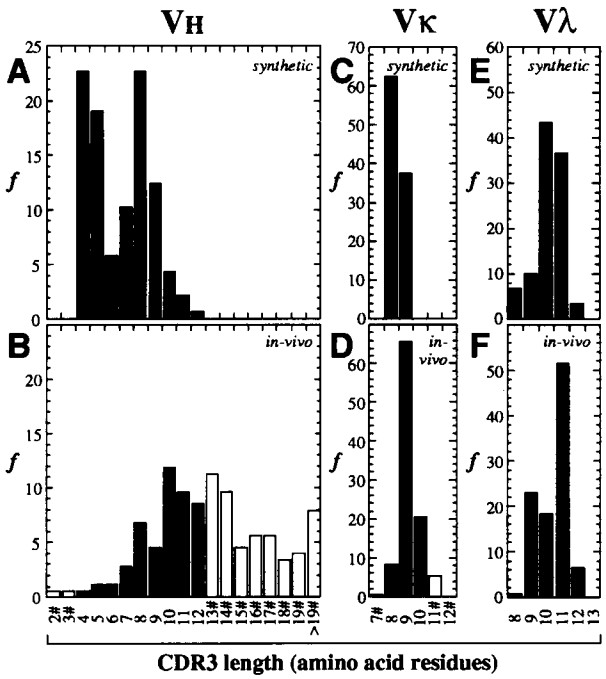

Fig. 5. Distribution of CDR3 lengths. Length distribution of CDR3 loops in human V_H, V_κ and V_λ chains from the synthetic repertoire (A, C and E), or from natural antibodies (B, D and F). Frequencies (f) are plotted as % of total. Data were compiled as in Figure 4 legend, except that for natural rearranged V_H genes the data was taken from the 177 human genes described by Wu *et al.* (1993). All CDR lengths listed were included in the synthetic repertoire except for those marked (#).

the synthetic Fab fragments binding to hapten–BSA conjugates; those fragments binding to protein antigen appear to have longer CDR3 lengths. The distribution of CDR3 lengths for the synthetic λ light chains was similar to natural antibodies, but differed for the κ light chains, with no loops of 10 residues seen in the synthetic chains.

Pairings of heavy and light chains

The distribution of heavy and light chain pairings (Figure 6) identified some 52 different segment pairings among the 137 unique clones. Several V_H gene segments were found paired with several V_κ and V_λ gene segments, for example DP-7, DP-38, DP-47 and DP-67 were each found in combination with several different light chain segments. Likewise, the light chain segments DPK-15 and DPL-3 were found in combination with several heavy chain segments. Not surprisingly these segments correspond to those used with higher frequency (see above). The usage of V gene segments differs for antibodies of different specificity, and the pattern of pairings provides a 'fingerprint', readily distinguishing the 'response' to different antigens. Some pairings, for example DP-47 with DPK-15, and DP-47 with DPL-16, were also present in Fab fragments of different specificities.

There were also several examples of 'promiscuous' pairings (Clackson *et al.*, 1991), chains that bind to the same antigen with any of several partner chains (Table III). For example, in the Fab fragments binding to NIP–BSA, the same heavy chain sequence (DP-38 segment with CDR3 of sequence AGTL) was paired with six different light chains of $V_\kappa 2$, $V_\lambda 1$ and $V_\lambda 3$ families (segments DPK-12, DPL-3 and DPL-16, respectively). Likewise in Fab fragments binding to FITC–BSA, the same light chain sequence (DPK-15 segment with CDR3 of sequence MQALQTRT) was paired with 15 different heavy chains of $V_H 1$, $V_H 3$ and $V_H 4$ families (DP-7, DP-47 and DP-67 segments, respectively).

Affinities of selected antibodies

Soluble Fab fragments were produced and purified via their hexahistidine tag by immobilized metal chelate affinity chromatography (IMAC), with typical yields of 100–500 μg/l. From the large (6.5×10^{10} clones) repertoire we measured the affinities of several Fab fragments binding to soluble NIP–CAP or fluorescein by fluorescence quench titration (Eisen, 1964). The affinities (K_d) of the Fab fragments ranged from 3.8 to 217 nM (Table IVA). This shows that high affinity anti-hapten antibodies can be isolated directly from large antibody repertoires. We also characterized three Fab fragments binding to the haptens NIP–CAP and fluorescein after selection of a small fraction (10^7 clones) of the repertoire on NIP–BSA or FITC–BSA. In contrast these affinities (K_d) ranged from 0.8 to 12 μM (Table IVB).

We measured kinetics and affinity of Fab fragments (from the 6.5×10^{10} clone repertoire) binding to immobilized monoclonal antibody NQ11/7.22 and HGF/SF by surface plasmon resonance (Table IVA). For the Fab fragments against the variable region (NML1) and the Fc portion (NML9) of antibody NQ11/7.22, the binding affinities were determined both by on- and off-rate analysis and by Scatchard plots of equilibrium binding (see Figure 7) to be in the range 30–60 nM. For NML1, the on-rate was calculated as 6.4×10^5 M^{-1} s^{-1} and the off-rate as 2.2×10^{-2} s^{-1}; for NML9 as 5.2×10^5 M^{-1} s^{-1} and 3×10^{-2} s^{-1}, respectively. However for the Fab fragment (MH22) against HGF/SF, Scatchard analysis indicated several classes of binding sites, with affinities from micro- to nanomolar. At low Fab concentrations (<40 nM), where high affinity interactions predominate, the affinity could be estimated as 7 nM from an initial on-rate of $1.7–1.9\times10^6$ M^{-1} s^{-1} and off-rate of 1.3×10^{-2} s^{-1}.

Discussion

In the immune system, antibodies with moderate affinities are selected from primary repertoires, and their affinities improved step-wise by rounds of somatic mutation and selection. However, theoretical arguments based on the idea of 'shape space' have suggested that larger and more diverse repertoires should give rise to higher affinity antibodies (Perelson and Oster, 1979). The probability (P) that an epitope is recognized by at least one antibody in a repertoire depends on the probability (p) that an antibody recognizes a random epitope with an affinity above a threshold value, and on the number of antibodies (N) according to the equation $P = 1-e^{-Np}$ (Perelson and Oster, 1979). This predicts, as expected, that the larger the repertoire, the greater the chances of finding a high affinity antibody. However it does not provide an explicit relationship between repertoire size and affinity.

Here we used a repertoire of phage antibodies as an 'artificial immune system' to explore the possibility of

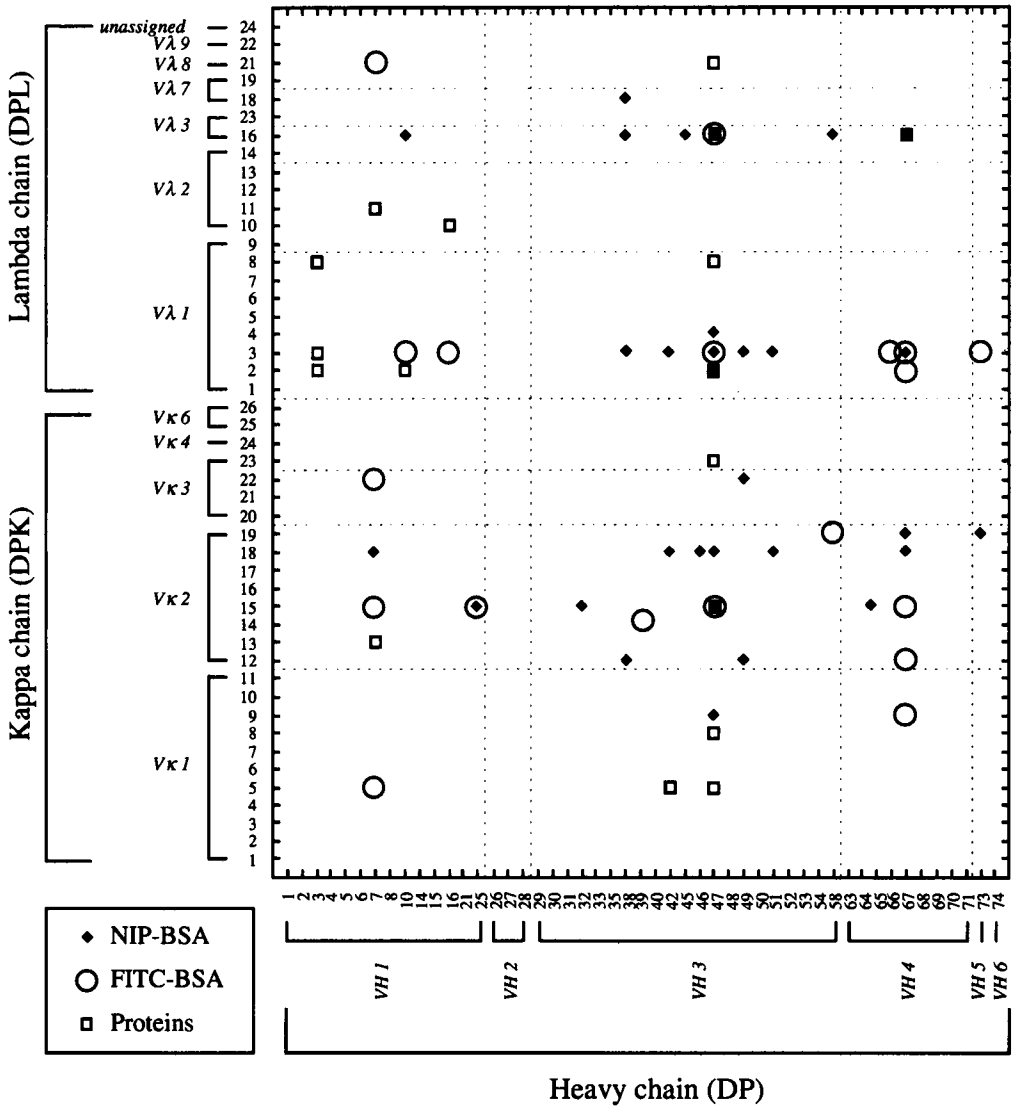

Fig. 6. Pairings of heavy and light chain V gene segments in the synthetic repertoire. Data were compiled and are listed as described in Figure 4 legend. Fab fragments binding NIP–BSA are indicated by black diamonds; Fab fragments binding FITC–BSA by open circles; and Fab fragments binding antibody NQ11/7.22, plasmin, u-PA, t-PA or HGF/SF by open squares.

isolating high affinity antibodies directly from a very large primary antibody repertoire. Previously the size of phage antibody repertoires has been limited to $<10^9$ clones by the efficiency of transfection of DNA into bacteria: here we have used the process of combinatorial infection and *in vivo* recombination (Waterhouse *et al.*, 1993) to overcome this. Bacteria harbouring a repertoire of heavy chains (encoded on a plasmid replicon) were infected with phage encoding a repertoire of light chains, and the heavy chain genes translocated to the phage replicon by recombination within the bacterium. By this means we were able to make a repertoire of 6.5×10^{10} clones and obtained antibodies to a range of antigens and haptens (Table II) with affinities <10 nM (Table IVA). With a smaller repertoire (10^7 clones) we only found antibodies of moderate affinities (>800 nM), in agreement with earlier work in which antibody fragments isolated from smaller repertoires (10^7-10^8 clones) were found to have affinities of 700 nM for the hapten NIP (Hoogenboom and Winter, 1992) and 140 nM for the hapten fluorescein (Barbas *et al.*, 1992c, 1993). The characterization of repertoires of different sizes (and of known diversity) should help in defining the explicit relation between repertoire size and affinity.

As shown in Figure 8, the binding affinities of secondary response (hypermutated) mouse monoclonal antibodies to the haptens NIP and fluorescein were similar to those of human Fab fragments isolated directly from the large repertoire. Only a single mouse anti-NIP antibody has been described (Lucisano-Valim and Lachmann, 1991) with an affinity (K_d = 5.6 nM) similar to the best human Fab fragment (K_d = 4.0 nM; Table IVA). However, eight mouse monoclonal antibodies have been described with higher affinities for fluorescein than the best human Fab fragment (K_d = 3.8 nM; Table IVA), the best of these (4-4-20) with an affinity of 0.19 nM (Bedzyk *et al.*, 1986).

It is quite possible that antibody fragments with even higher affinities are present in the repertoire, as we had designed the selection process to capture antibody fragments with a range of binding affinities to antigen,

Table IV. Sequences and affinities of Fab fragments

(**A**) Fabs from 6.5×10^{10} repertoire

Antigen	Clone	Heavy chain		Light chain		K_d (nM)
		segment	CDR3	segment	CDR3	
NIP–CAP	NIP-G6	DP-38	AGTL	DPK-12	MQSIQL**PT**	4.0 (±0.1)
NIP–CAP	NIP-G10[c]	DP-38	AGTL	DPK-12	MQSIQL**PAT**	5.4 (±0.2)
NIP–CAP	NIP-H1[c]	DP-38	KGSE	DPL-3	AAWDDSL**AWFV**	11.3 (±0.4)
NIP–CAP	NIP-C11	DP-38	PGYRGTR	DPL-3	AAWDDSL**SAYV**	16.5 (±0.5)
NIP–CAP	NIP-H3[c]	DP-38	HGQH	DPL-3	AAWDDSL**CPEFV**	19.7 (±1.7)
NIP–CAP	NIP-G11[c]	DP-47	PLNSKKNTTTQ	DPL-3	AAWDDSL**FYV**	20.1 (±3.6)
NIP–CAP	NIP-G9[c]	DP-38	AGTL	DPL-3	AAWDDSLV	22.0 (±1.0)
NIP–CAP	NIP-E5[c]	DP-38	PNGDQ	DPL-3	AAWDDSL**AFV**	22.1 (±0.8)
NIP–CAP	NIP-E7[c]	DP-38	GGTQ	DPL-3	AAWDDSLV	29.8 (±1.1)
NIP–CAP	NIP-A4	DP-38	PATRS	DPK-15	MQALQT**PLT**	48 (±1.0)
NIP–CAP	NIP-C9	DP-38	GGKD	DPL-18	LLYYGGA**YV**	59 (±3.0)
Fluorescein	FITC-B4	DP-66	WSRETNYS	DPL-3	AAWDDSL**WSAV**	3.8 (±0.4)
Fluorescein	FITC-A4	DP-67	RSFVGYEI	DPL-3	AAWDDSLV	14.3 (±2.0)
Fluorescein	FITC-B11	DP-47	IGQF	DPL-3	AAWDDSL**AFV**	24.1 (±0.8)
Fluorescein	FITC-B7	DP-67	AQRKYFDY	DPK-12	MQSIQL**RT**	151 (±3.0)
Fluorescein	FITC-A2	DP-47	YRFSAPPRD	DPL-3	AAWDDSL**PSGV**	217 (±16.0)
NQ11 (Fv)	NML1	DP-47	ASSPFVLQ	DPL-21	VLYMGSG**SAV**	32[a] and 34[b]
NQ11 (Fc)	NML9	DP-14	GTGLDG	DPL-10	CSYAGSS**YV**	41[a] and 58[b]
HGF/SF	MH22	DP-42	KFPHFGD	DPK-5	QQANSF**PIT**	7[b]

(**B**) Fabs from 1×10^{7} repertoire

Antigen	Clone	Heavy chain		Light chain		K_d (μM)
		segment	CDR3	segment	CDR3	
NIP–CAP	sNIP-D10	DP-53	PWARGTD	DPK-21	QQYNNW**LST**	8 (±0.6)
NIP–CAP	sNIP-F3[c]	DP-47	NYNAAFDY	DPL-21	VLYMGSG**HRV**	12 (±1.3)
Fluorescein	sFITC-C2	DP-67	SGVRGLMT	DPK-9	QQSYST**RT**	0.82 (±0.14)

Affinities (K_d) for haptens were determined by fluorescence quench titration. Affinities (K_d) for protein antigens were determined by SPR, by Scatchard analysis[a] and from analysis of the rate constants[b]. All clones were derived from selections on immunotubes except for those marked[c] which were derived from selections using magnetic beads. The residues in the light chain CDR3 regions encoded by randomized codons are in bold.

including those with only moderate affinity. Thus, we took advantage of the binding avidity of multiple Fab fragments on the surface of the phage (by using phage rather than phagemid vectors) and short wash times to retain phage with fast dissociation kinetics. Also for selections on immunotubes we used a high coating density of antigen to favour rebinding of the phage, and for selections with soluble biotinylated antigen we used a relatively high concentration of antigen (50 nM) to try to include even those phage with moderate equilibrium constants (Hawkins *et al.*, 1992). Presumably it would be possible to favour the selection of higher affinity antibodies from this repertoire by more stringent selection.

The phage repertoire was not only large, but also highly diverse. It was assembled from the majority of V gene segments used *in vivo*, including all the major V_H and V_λ families, and V_κ subgroups. The segments included all the major heavy and light chain CDR1 and CDR2 loop conformations (Chothia and Lesk, 1987; Chothia *et al.*, 1989, 1992), and the CDR3 loops were made of diverse sequences and lengths. The chains were paired at random (Huse *et al.*, 1989), creating diverse pairings. The repertoire was sufficiently large that several pairings of a single heavy or light chain could be isolated from the repertoire. Such chain 'promiscuity' is characteristic of repertoires derived from the mRNA from immune sources (see for

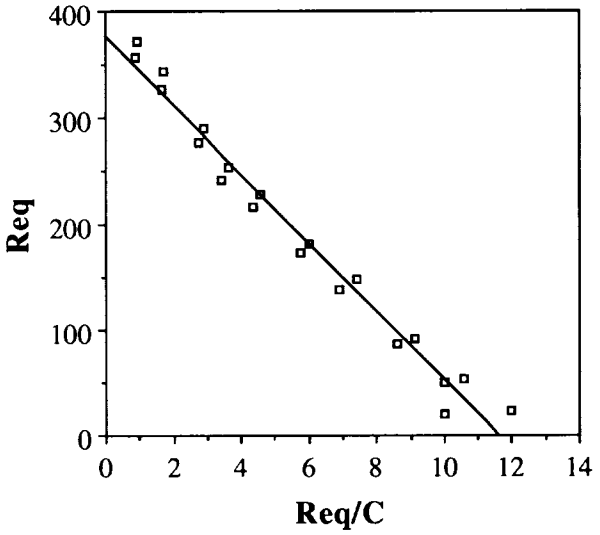

Fig. 7. Affinity of Fab NML1 by SPR. Binding of the Fab to a sensor chip surface coated with mouse monoclonal antibody NQ11/7.22. See Materials and methods for further details. For each concentration of antibody (C, nM), the equilibrium binding signal (Req, in resonance units) was plotted against the Req/C. The slope gives the binding affinity (K_d) as 32.3 nM.

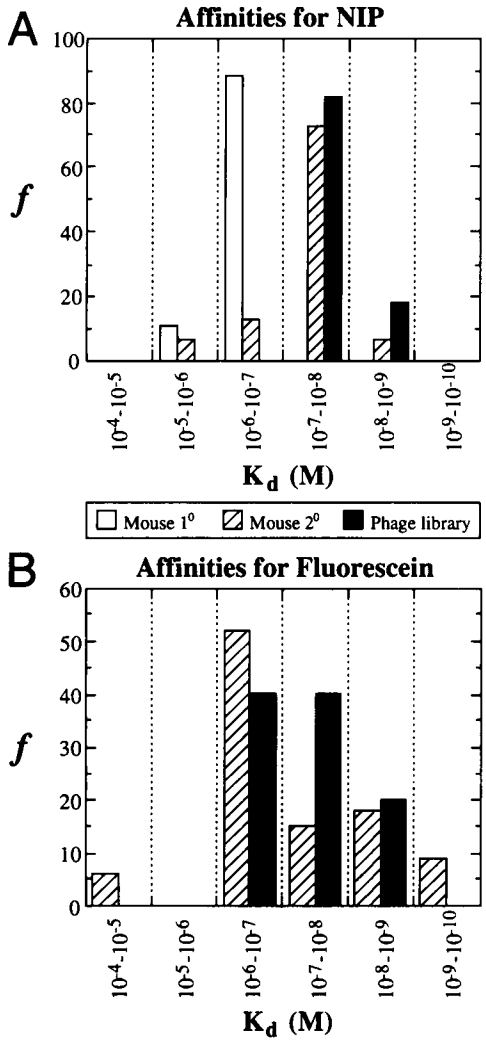

Fig. 8. Affinities for synthetic antibodies and mouse monoclonal antibodies. Affinity data for the Fab fragments from the synthetic repertoire were compiled from Table IVA. (**A**) Antibodies binding to NIP; data on the mouse immune response were taken from Mariuzza and Strand (1981), Cumano and Rajewsky (1986) and Lucisano-Valim and Lachmann (1991). (**B**) Antibodies binding fluorescein; data on the mouse immune response were taken from Kranz and Voss (1981), Kranz *et al.* (1982), Reinitz and Voss (1984), Bates *et al.* (1985), Bedzyk *et al.* (1986) and Denzin and Voss (1992).

example Clackson *et al.*, 1991) and reflects the chances of a chain making multiple pairings, which in turn depends on the frequency of the chain and the size of the repertoire.

In the antibodies binding to the haptens, the combinations of heavy and light chain segments appeared to be restricted. For example, the antibodies binding to soluble hapten NIP–CAP (Table IVA) mainly utilized the heavy chain segment DP-38 and the light chain segment DPL-3, and included a four residue heavy chain CDR3 loop with a distinctive motif, X-Gly-X-X. This is reminiscent of the restricted response seen with mouse antibodies raised by immunization with the hapten 4-hydroxy-3-nitrophenyl acetate (NP): the immune response in C57BL/6 mice is dominated by antibodies with λ1 light chains paired with heavy chains encoded by the V186.2 V_H segment and the DFl16.1 D segment (Bothwell *et al.*, 1981; Cumano and Rajewsky, 1985).

As *in vivo*, only a few segments contributed to most of the 'response'. However, the usage of V gene segments found in the selected Fab fragments from the artificial repertoire differed from that of antibodies from natural repertoires (Figure 4). This might reflect a different representation of segments in the primary repertoire, or the high frequency of sequences from phage antibodies directed against hapten conjugates. Nevertheless we would expect the artificial repertoire to be shaped by different selection pressures to those that operate *in vivo*, leading to a different spectrum of binding sites. If 'holes' exist in the natural repertoire, they may be absent in an artificial repertoire and vice versa. Despite the differences in segment usage between the natural and synthetic antibodies, there is one striking similarity: the heavy chain segment DP-47 is the most commonly used segment in both. *In vivo*, it has been suggested that pre-B cells displaying DP-47 heavy chains are selected by binding to autoantigen (Schwartz and Stollar, 1994). We suggest that the usage of DP-47 reflects a more fundamental property of this segment; the structure of the binding site may be more capable of accommodating a wider diversity of antigens than other segments (Figure 6).

We conclude that human antibodies with affinities in the nanomolar range, and specific for protein antigens and haptens, can be derived directly from large and diverse synthetic phage antibody repertoires. The binding affinities are typical of somatically mutated mouse antibodies produced *in vivo*, and presumably could be improved further through rounds of 'chain shuffling' (Marks *et al.*, 1992b) or point mutagenesis (Hawkins *et al.*, 1992) to create binding specificities and affinities outside the reach of the immune system. In addition, the use of synthetic phage antibody repertoires based on V gene segments to simulate natural immune systems may provide further insights into immune strategy, for example in helping to define the relationship between affinity and repertoire size, or the relationship between the structures of the antigen binding sites and the usage of V, D and J segments, CDR3 lengths and junctional diversity.

Materials and methods

Vectors

The 'acceptor' vector, into which light chain repertoires are cloned, fdDOG-2loxVκdel, is identical to fdDOG-2lox (Waterhouse *et al.*, 1993), except that the light chain variable region gene has been deleted (Figure 2A). The 'donor' vector, into which heavy chain repertoires are cloned, pUC19-2loxVHdel, is identical to pUC19-2lox (Waterhouse *et al.*, 1993), except that the heavy chain variable region gene has been deleted (Figure 2B). The vector for expression of soluble Fab fragments, pUC119His6mycXba, is a derivative of pUC119 (Vieira and Messing, 1987) in which the polylinker has been replaced by the sequence shown in Figure 2C.

Construction of synthetic heavy chain repertoires

A diverse repertoire of rearranged V_H genes has previously been built *in vitro* (Nissim *et al.*, 1994) from a bank of 49 cloned V_H gene segments (Tomlinson *et al.*, 1992) [one of the 50 segments (DP-20) included in the repertoire of Nissim *et al.* (1994) was a pseudogene]. To these segments completely randomized CDR3 regions (Kabat *et al.*, 1991) were appended, varying in size between four and 12 residues (Figure 1A). This cloned repertoire, with $>10^8$ different clones, was re-amplified using PCR with primers pUC-reverse and JH-Xho-FOR (Table VA), the DNA was cut with *Nco*I and *Xho*I, and ligated into pUC19-2loxVHdel. The ligation mixture was electroporated (Dower *et al.*, 1988) into *E.coli*

Table V. Oligonucleotides

A. ***Re-cloning of synthetic human VH repertoires into pUC19-2loxVHdel***

pUC-reverse	5'- AGC GGA TAA CAA TTT CAC ACA GG	
JH-Xho-FOR	5'- GCC TGA ACC GCC TCC ACC ACT CGA GAC GGT GAC CAG GGT ACC TTG GCC CCA	

B. ***Construction of synthetic human kappa chain repertoires***

1. Amplification of human Cκ

CκFOR	5'- CTG CTA TTA TCG GGC GCG CCT TTA TTA ACA CTC TCC CCT GTT GAA GCT CTT TGT GAC GGG	
CκLink	5'- ACG TTC GGC CAA GGG ACC AAG STG GAA ATC AAA CGT ACT GTG GCT GCA CCA TCT GTC	

2. Human Vκ back primers

SYNκB1	5'- CAT GAC CAC AGT GCA CTT GAC ATC CAG WTG ACC CAG	DPK1,4,5,6,7,8,9,11
SYNκB2	5'- CAT GAC CAC AGT GCA CTT GTC ATC TGG ATG ACC CAG	DPK10
SYNκB3	5'- CAT GAC CAC AGT GCA CTT GCC ATC CAG ATG ACC CAG	DPK3
SYNκB4	5'- CAT GAC CAC AGT GCA CTT GAT RTT GTG ATG ACT CAG	DPK15,18,19
SYNκB5	5'- CAT GAC CAC AGT GCA CTT GAR ATT GTG ATG ACC CAG	DPK12,13,14,16,17
SYNκB6	5'- CAT GAC CAC AGT GCA CTT GAA ATT GTG TTG ACG CAG	DPK20,22
SYNκB7	5'- CAT GAC CAC AGT GCA CTT GAA ATA GTG ATG ACG CAG	DPK21
SYNκB8	5'- CAT GAC CAC AGT GCA CTT GAC ATC GTG ATG ACC CAG	DPK24
SYNκB9	5'- CAT GAC CAC AGT GCA CTT GAT GTT GTG ATG ACA CAG	DPK25
SYNκB10	5'- CAT GAC CAC AGT GCA CTT GAA ATT GTG CTG ACT CAG	DPK26
SYNκB11	5'- CAT GAC CAC AGT GCA CTT AAC ATC CAG ATG ACC CAG	DPK2
SYNκB12	5'- CAT GAC CAC AGT GCA CTT GAA ATT GTA ATG ACA CAG	DPK23

3. Human Vκ forward primers encoding synthetic CDR3s

DPK1FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GAG ATT ATC ATA CTG TTG AC	DPK1
DPK2FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GTA ACT ATT ATG CTG TAG AC	DPK2
DPK3FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GTA ATT GTA ATC TTG TAG AC	DPK3
DPK4FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GGC ACT GTT ATA CTT TTG AC	DPK4
DPK5/6FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GAA ACT GTT AGC CTG TTG AC	DPK5,6
DPK7FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GTA ACT ATT ATA CTG TTG GC	DPK7
DPK8FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GTA ACT ATT AAG CTG TTG AC	DPK8
DPK9FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GGT ACT GTA ACT CTG TTG AC	DPK9
DPK10FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GAA ACT ATA ATA CTG TTG AC	DPK10
DPK11FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GGC ATT GTA AGT CCG TTG AC	DPK11
DPK12FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AAG CTG TAT ACT TTG CAT GC	DPK12
DPK13FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AAA CTC TAT ACG TTG CAT GC	DPK13
DPK14FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG ATC TTG TGC ATC TTG CAT GC	DPK14
DPK15FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AGT TTG TAG AGC TTG CAT GC	DPK15
DPK16FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AAA TTG TGT AGC TTG CAT GC	DPK16
DPK17FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AAA TTG TGT AGC TTG CGT GC	DPK17
DPK18/19FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG CCA GTG TGT ACC TTG CAT GC	DPK18,19
DPK20/22FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG TGA GCT ACC ATA CTG CTG AC	DPK20,22
DPK21FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG CCA GTT ATT ATA CTG CTG AC	DPK21
DPK23FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG TAA GTT ATA ATC CTG CTG AC	DPK23
DPK24FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG AGT ACT ATA ATA TTG CTG AC	DPK24
DPK25FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG GTG CTT ATT GCC CTG CTG AC	DPK25
DPK26FOR	5'- CTT GGT CCC TTG GCC GAA CGT (MMM)$_{0-2}$ NNG TAA ACT ACT ACT CTG ATG AC	DPK26

C. ***Construction of synthetic human lambda chain repertoires***

1. Amplification of human Cλ2

CL2FOR1	5'- GGA ATT CGG GGT GAG GGT TGA GAA	
CL2BACK1	5'- CCC CCA AGC TTC TGC CCC TCA TCC	
CL2BACK2	5'- GTA TTC GGC GGA GGG ACC AAG CTG ACC GTC CTA GGT CAG CCC AAG GCT GCC CCC TCG GTC ACT	
HuCλFORCYSASCNOT	5'- GAG TCA TTC TCG ACT TGC GGC CGC CTG CTA TTA TCG GGC GCG CCT TTA TTA TGA AGA TTC TGT AGG GGC CAC TGT CTT	

2. Human Vλ back primers

DPVλ1a	5'- CAT GAC CAC AGT GCA CTT CAG TCT GTG YTG ACG CAG CCG CC	DPL4,5,8,9
DPVλ1b	5'- CAT GAC CAC AGT GCA CTT CAG TCT GTC GTG ACG CAG CCG CC	DPL6,7
DPVλ1c	5'- CAT GAC CAC AGT GCA CTT CAG TCT GTG CTG ACT CAG CCA CC	DPL1,2,3
DPVλ2	5'- CAT GAC CAC AGT GCA CTT CAR TCT GCC CTG ACT CAG CCT	DPL10,11,12,13,14
DPVλ3a	5'- CAT GAC CAC AGT GCA CTT TCT TCT GAG CTG ACT CAG GAC CC	DPL16
DPVλ3b	5'- CAT GAC CAC AGT GCA CTT TCC TAT GAG CTG ACT CAG CCA CC	DPL23
DPVλ7/8	5'- CAT GAC CAC AGT GCA CTT CAG RCT GTG GTG ACY CAG GAG CC	DPL22,24
DPVλ9	5'- CAT GAC CAC AGT GCA CTT CWG CCT GTG CTG ACT CAG CCM CC	DPL18,19,21

3. Human Vλ forward primers encoding synthetic CDR3s

DPL1/2/3V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-5}$ --- CAG GCT GTC ATC CCA TGC TGC ACA GTA	DPL1,2,3
DPL4V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- CGG GCT GGT ATC CCA TGC	DPL4
DPL4V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-2}$ TCT CGG GCT GGT ATC CCA TGC TAA GCA GTA	DPL4
DPL5V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- CAG GCT GCT ATC CCA TGT	DPL5
DPL5V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-4}$ ACT CAG GCT GCT ATC CCA TGT TCC GCA CTA	DPL5
DPL6/7/8V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- CAG GCT GCT GTC ATA GGA	DPL6,7,8
DPL6/7/8V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-2}$ ACT CAG GCT GCT GTC ATA GGA CTG GCA GTA	DPL6,7,8
DPL9V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- CAG GCT GTT ATC CCA TGC	DPL9
DPL9V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-1}$ ATT CAG GCT GTT ATC CCA TGC TTT GCA GTA	DPL9
DPL10V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- GCT ACT ACC TGC ATA TGA	DPL10
DPL11/13V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- GCT GCT GCT TGT ATA TRA GCT GCA GTA	DPL11,13
DPL11/13V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-2}$ AGT GCT GCT GCT TGT ATA TRA GCT GCA GTA	DPL11,13
DPL12V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- GTA GCT GCC TGC ATA TGA	DPL12
DPL14V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- GTA ACT ACT TGA ATA TAA	DPL14
DPL14V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-2}$ AGT GTA ACT ACT TGA ATA TAA GCT GCA GTG	DPL14
DPL16V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- --- ACC ACT GCT GTC CCG GGA	DPL16
DPL16V/JFORa	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-2}$ ATG GTT ACC ACT GCT GTC CCG GGA GTT ACA	DPL16
DPL18V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ AGC ACC ACC ATA GTA GAG CAG	DPL18
DPL19V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- AGC ACC ACT ATA GGA GAG	DPL19
DPL21V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ GCC ACT ACC CAT ATA CAG	DPL21
DPL22V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-1}$ GAA GTT GCT CCC ACT GCC ATG GTC TGC CCC	DPL22
DPL23V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-3}$ --- AGT GCT GCT GTC CCA CGC CTG ACA GTA	DPL23
DPL24V/JFOR	5'- CTT GGT CCC TCC GCC GAA TAC (MMM)$_{0-1}$ GAC TTG GCC ATC AAT CGT GTG GCT CTC TCC	DPL24

D. ***Sub-cloning of selected repertoires for expression of soluble Fab fragments***

fdSEQ1	5'- GAA TTT TCT GTA TGA GG	
G3LXbaGTGBack	5'- GTC CTC GCA ACT TGC TCT AGA CAA TTT CAC AGT AAG GAG GTT TAA CTT GTG AAA AAA TTA TTA TTC GCA ATT	

E. ***Sequencing and probing***

TNFCDR3PRB	5'- CCT TGG AAG GCA GCA GC
Cκ.lib.seq	5'- CAA CTG CTC ATC AGA TGG CG
Cλ.lib.seq	5'- GTG GCC TTG TTG GCT TGA AGC
CH1.lib.seq	5'- GGT GCT CTT GGA GGA GGG TGC
pelBback	5'- GAA ATA CCT ATT GCC TAC GG
LMB3	5'- CAG GAA ACA GCT ATG AC
fdPCRback	5'- GCG ATG GTT GTT GTC ATT GTC GGC

A = adenosine; C = cytosine; G = guanine; T = thymidine; Y = C or T; R = A or G; W = A or T; S = G or C; K = T or G; M = C or A
Restriction sites are underlined. Primer names are to the left of the sequences and the V-gene segments amplified to the right.

TG1 (Gibson, 1984) to create the library pUC19-2loxVHlib, and in total 5×10^8 clones were obtained. Diversity was confirmed by *Bst*NI fingerprinting (Clackson *et al.*, 1991) and sequencing of 24 independent clones. All these clones were found to be different.

Construction of synthetic kappa chain repertoires

The human C_κ gene was amplified from the vector pSW1/FabD1.3 (Skerra *et al.*, 1991) by PCR with Taq polymerase using primers CκFOR and CκLink (Table VB1) which introduce a consensus human J_κ segment at the 5'-end of the C_κ gene and two stop codons (TAA) and an *Asc*I site at the 3'-end. The reaction mixture (50 µl) was cycled 25 times (94°C for 1 min, 60°C for 1 min, 72°C for 1 min). The amplified C_κ gene was purified using Magic PCR Preps (Promega) and resuspended in 50 µl water.

In parallel, 26 human germline V_κ gene segments with open reading frames (DPK1–26), which had been cloned from the genomic DNA of a single individual (Cox *et al.*, 1994), were individually amplified with Back primers that introduce an *Apa*LI site at the 5'-end (see Table VB2) and Forward primers that append a portion of the J_κ segment to the 3'-end (see Table VB3). For each segment three independent PCRs were performed with different Forward primers to construct CDR3 regions (Kabat *et al.*, 1991) of length eight, nine or 10 residues that included one, two or three residues of random sequence (Figure 1B). For each gene, a 50 µl PCR was performed using Taq polymerase and toothpicked frozen glycerol stocks of *E.coli* infected with the appropriate M13 clone as template. Reactions were cycled 25 times (94°C for 1 min, 60°C for 1 min, 72°C for 1 min). The Forward primers (Table VB3) introduced length and sequence diversity into CDR3 corresponding to that observed *in vivo* (Kabat *et al.*, 1991).

The amplified V_κ genes were each joined to the amplified C_κ gene using PCR (Horton *et al.*, 1989). Assembly PCRs (25 µl) used Taq polymerase, 1 µl of amplified C_κ and 0.8 µl of the V_κ gene PCR from above. The appropriate V_κ Back primer was used for each gene (Table VB2) together with C_κ FOR (Table VB1) and the reaction cycled 30 times (94°C for 1 min, 55°C for 1 min, 72°C for 2 min).

The PCR assembly reactions for each V_κ gene were checked by agarose gel electrophoresis, then pooled with the other V_κ genes according to CDR3 length, and the V_κ–C_κ genes purified using Magic PCR Preps. The pooled DNA was then cut with *Apa*LI and *Asc*I and digested DNA (~6 µg) purified from a 1.5% low melting-point agarose gel using Magic PCR Preps. Approximately 1 µg of the purified and cut V_κ–C_κ DNA from each pool was ligated in a 60 µl volume with 1200 U of T4 DNA ligase (New England Biolabs) to ~5 µg of digested fdDOG-2loxVκdel vector [previously electroeluted from a 0.8% agarose gel (Sambrook *et al.*, 1990)]. DNA was purified from the ligation mixture using Geneclean II (Bio 101), resuspended in 30 µl water, and electroporated (Dower *et al.*, 1988) into four 50 µl aliquots of *E.coli* TG1. Cells were grown in 1 ml 2×TY broth containing 1% glucose for 1 h and then plated in 243 mm×243 mm dishes (Nunc) on TYE (Miller, 1972) medium with 12.5 µg/ml tetracycline (TYE–TET). After overnight incubation at 37°C, colonies were scraped off the plates into 7 ml of 2×TY broth (Miller, 1972) containing 15% (v/v) glycerol for storage at −70°C.

The frequency of inserts was checked by PCR for each of the three pools and found to be 90% for CDR3 of eight residues, 100% for CDR3 of nine residues and 87% for CDR3 of 10 residues. The number of clones with light chains could then be calculated as 9.9×10^3 (CDR3 of eight residues), 1.5×10^4 (CDR3 of nine residues), 6.5×10^4 (CDR3 of 10 residues). Sequence diversity was confirmed by sequencing eight clones of each CDR3 length; all clones were found to be different. The pools were then combined, to create the library fdDOG-2loxVκlib, corresponding to 9.0×10^4 light chains.

Construction of synthetic lambda chain repertoires

The human Cλ2 gene (Vasicek and Leder, 1990) was amplified from genomic DNA by PCR with Taq polymerase using primers CL2BACK1 and CL2FOR1 (Table VC1) based in the regions flanking the Cλ2 exon. The *Eco*RI and *Hin*dIII sites in CL2FOR1 and CL2BACK1, respectively, were used to clone the PCR product into M13mp19 (Yanisch-Perron *et al.*, 1985).

Twenty-one V_λ germline gene segments with open reading frames, previously cloned in M13mp19 (Williams and Winter, 1993), were individually amplified using PCR and Taq polymerase with back primers (Table VC2) which anneal to framework 1 (FR1) and introduce a 5' *Apa*L1 site, and forward primers (Table VC3) which append a portion of the Jλ2 gene (Vasicek and Leder, 1990) to CDR3. CDR3 loops (Kabat *et al.*, 1991) of eight to 13 residues that included zero, one, two, three, four or five residues of random sequence were encoded by the forward primers (Figure 1C). The number of residues of random sequence included was designed to match the pattern of V gene rearrangement seen *in vivo* and varied with the different V_λ segments used.

Each synthetically rearranged V_λ gene was individually joined to the human Cλ2 gene by PCR with Taq polymerase (Horton *et al.*, 1989). Each 50 µl PCR assembly reaction contained ~1 ng of M13mp19 containing the Cλ2 gene, ~0.1 µg of the V_λ gene, the appropriate (FR1) back primer (25 pmol) (Table VC2), the back primer CL2BACK2 (2.5 pmol) (Table VC1), which contains the 3' sequence of the Jλ2 gene linked to the 5' sequence of the Cλ2 gene, and the forward primer HUCλFORCYSASCNOT (25 pmol) (Table VC1) which appends two stop codons (TAA) followed by an *Asc*I site to the 3' of the Cλ2 gene. Reactions were cycled 30 times (94°C for 1 min, 65°C for 1 min, 72°C for 2 min).

The PCR assembly reactions were combined into a single pool and the V_λ genes digested and ligated into fdDOG-2loxVκdel as described for the V_κ–C_κ gene pools (see above), thus creating the library fdDOG-2loxVλlib. Ninety-two percent of clones were found to carry inserts of the correct size, corresponding to a repertoire size of 7.4×10^5 λ light chains. Thirty-three clones were sequenced to confirm the presence of each V_λ segment. All the sequences were different.

Combinatorial infection and in vivo recombination

To create a large combinatorial repertoire of heavy and light chains on an fd phage vector we used the strategy of combinatorial infection and *in vivo* recombination (Waterhouse *et al.*, 1993). This system uses the *lox*–Cre site-specific recombination system of bacteriophage P1 (Sternberg and Hamilton, 1981; Hoess *et al.*, 1982) to bring together heavy and light chain genes onto the same replicon.

Phage P1 lysates were made by thermal induction (Rosner, 1972). *E.coli* C600 Su^- (Appleyard, 1954) harbouring phage P1Cm *c*1.100 r^-m^- (Yarmolinsky *et al.*, 1989) were grown in a 2 l baffled flask containing 1 l of 2×TY, 25 µg/ml chloramphenicol, 10 mM $MgSO_4$ with vigorous shaking at 30°C to OD_{600} of 0.6. The temperature was then raised quickly to 42°C by shaking in a 70°C water bath and then shaking continued for a further 35 min in a 40°C water bath. Shaking was then continued at 37°C until lysis was visible (usually ~1.5–2 h). The culture was then spun at 5000 *g* for 15 min at 4°C and 100 µl chloroform added to the supernatant. P1 phage titres were measured by adding serial dilutions of the lysate to mid-log phase *E.coli* TG1 (Gibson, 1984) grown in 2×TY broth containing 5 mM $CaCl_2$, the mixture incubated for 30 min at 30°C to allow infection and then plated on TYE medium (Miller, 1972) containing 30 µg/ml chloramphenicol. Chloramphenicol resistant colonies were counted after 24 h incubation at 30°C. The P1 titre of the lysate used for this library was 3×10^9 transducing units (t.u.) per ml.

10^9 *E.coli* TG1, harbouring the library of synthetic V_κ genes (or the library of synthetic V_λ genes) cloned in fdDOG-2loxVκdel, were used to inoculate 1 l of 2×TY broth containing 12.5 µg/ml tetracycline (2x TY–TET) and the culture shaken for 20 h at 30°C in two 500 ml aliquots in 2 l baffled Erlenmeyer flasks. Phage were purified from the supernatant by precipitation with polyethylene glycol as in McCafferty *et al.* (1990), resuspended in PBS (phosphate buffered saline: 25 mM NaH_2PO_4, 125 mM NaCl, pH 7.0) and filtered through a 0.45 µm sterile filter (Minisart, Sartorius). Phage were titred by infecting exponential phase *E.coli* TG1 (30 min, 37°C) and plating on TYE–TET. Yields were typically 10^{10} t.u. per ml of culture.

At various points during the recombination procedure aliquots of bacteria were removed and serial dilutions plated on TYE plates supplemented with 1% glucose and containing a variety of different antibiotics (100 µg/ml ampicillin; 15 µg/ml tetracycline; 30 µg/ml chloramphenicol). From the number of colony forming units (c.f.u.) the overall repertoire size could be calculated. These points are indicated in the protocol below and the results are summarized in Table I (also see Results).

Approximately 10^9 *E.coli* TG1 harbouring the library of synthetic heavy chain genes cloned in pUC19-2loxVHdel (pUC19-2loxVHlib; see above) were used to inoculate 100 ml 2×TY broth containing 100 µg/ml ampicillin and 1% (w/v) glucose (2×TY–AMP–GLU). An aliquot of bacteria was plated for c.f.u. determination (see Table I, sample point 1) and the rest of the culture grown overnight at 30°C. An aliquot of bacteria was then plated for c.f.u. determination (see Table I, sample point 2). Two 5 ml aliquots of the overnight culture were then used to inoculate two 500 ml aliquots of 2×TY–AMP–GLU in 2 l Erlenmeyer flasks and the cultures grown, shaking, at 37°C to an OD_{600} of 0.5.

2×10^{12} t.u. of V_κ library in fdDOG-2λoxVκdel were added to one

of the above cultures and 2×10^{12} t.u. of V_λ library in fdDOG-2loxVκdel were added to the other culture. Each culture was immediately split into 5× 100 ml aliquots and each aliquot mixed with 1 l of 2×TY−AMP−GLU, pre-warmed to 37°C. These cultures were then incubated at 37°C, without shaking for 30 min, and then with shaking until an OD_{600} of 0.4 was reached (~30 min). An aliquot of bacteria from the kappa infection and another from the lambda infection were plated for c.f.u. determination (see Table I, sample point 3).

$CaCl_2$ was then added to a final concentration of 5 mM and 200 ml phage P1Cm *c*1.100 r^-m^- lysate (6×10^{11} t.u, see above) were added to each 1 l flask (giving a multiplicity of infection of ~1). Incubation was continued at 30°C for 1 h, with a short burst of shaking every 15 min. The culture was then centrifuged at 5000 *g* for 15 min and the pellets resuspended in the original volume of 2×TY broth containing 100 μg/ml ampicillin, 12.5 μg/ml tetracycline, 25 μg/ml chloramphenicol and 1% glucose (i.e. 5 l for the V_κ library and 5 l for the V_λ library). An aliquot of bacteria was plated for c.f.u. determination (see Table I, sample point 4). The repertoire size was determined to be 6.5×10^{10} from the number of ampicillin, tetracycline and chloramphenicol resistant c.f.u. at this point (see Results). An aliquot of each culture was also centrifuged at 12 000*g* for 5 min and the supernatant filtered through a 0.45 μm sterile filter (Minisart, Sartorius). The fd phage in the supernatant were titred by infecting exponential phase *E.coli* TG1 (30 min, 37°C) plating on TYE−TET (see Table I, sample point 5).

The cultures were incubated overnight at 30°C with shaking for 24 h in 2 l baffled flasks (1 l medium per flask). An aliquot of bacteria from each culture was plated for c.f.u. determination (see Table I, sample point 6). The fd phage in the supernatant were also titred by infection of exponential phase *E.coli* TG1 as above (see Table I, sample point 7). The total yield of fd phage was 4.1×10^{13} t.u. and therefore >99.9% were propagated from bacteria containing the pUC 'donor' vector and phage P1. The cultures were then centrifuged at 5000 *g* for 15 min at 4°C and the fd phage precipitated from the supernatant using polyethylene glycol (McCafferty *et al.*, 1990) and resuspended in a final volume of 10 ml PBS.

Ten 2 l flasks, each containing 1 l 2×TY broth were inoculated with *E.coli* TG1 and grown, shaking, at 37°C until an OD_{600} of 0.4 (~4×10^{12} bacteria) was reached. Two microlitres of the above recombined V_κ fd phage (8×10^{11} t.u.) were added to 5 l of *E.coli* and 2 ml of the above recombined V_λ fd phage (2×10^{12} t.u.) were added to the other 5 l of *E.coli* and the cultures held at 37°C for 30 min without shaking and then for 30 min with shaking. The number of *E.coli* infected with fd phage was determined by plating bacteria on TYE−TET plates to be 1.7×10^{12} (V_κ repertoire) and 1.1×10^{12} (V_λ repertoire). This exceeds the estimated repertoire size by >10-fold, hence maintaining library diversity. Tetracycline was then added to 12.5 μg/ml and the culture shaken for 16 h at 30°C. The V_κ culture and the V_λ cultures were then centrifuged at 5000 g for 10 min and the pellet from each repertoire resuspended in 250 ml 2×TY broth containing 15% glycerol and stored in 15 ml aliquots at −70°C.

Aliquots of the two libraries were also spread on TYE−TET in 243 mm×243 mm dishes (Nunc). After overnight incubation at 30°C the number of colonies on the large plates was calculated from the number of colonies on small TYE−TET plates on which serial dilutions had been spread. Two plates, one containing 3.5×10^6 colonies of the V_κ library and the other containing 6.4×10^6 colonies of the V_λ library were selected, and the bacteria scraped into 10 ml 2×TY broth containing 15% glycerol. This stock therefore corresponded to a repertoire of 10^7 clones.

Assaying the efficiency of in vivo recombination

To test the efficiency of replacement of the anti-TNF heavy chain in the 'acceptor' vectors (fdDOG-2loxVκlib and fdDOG-2loxVλlib) with synthetic heavy chain from the 'donor' vector (pUC19-2loxVHlib), 250 individual colonies from each of the V_λ and V_κ recombined libraries were picked onto TYE−TET plates and grown overnight at 30°C. Colony hybridization was then performed as in Tomlinson *et al.* (1992), with a primer (TNFCDR3PRB; Table VE) complementary to the CDR3 region of the anti-TNF heavy chain gene found in the 'acceptor' vectors (fdDOG-2loxVλlib and fdDOG-2loxVκlib). Where recombination is successful, the anti-TNF heavy chain gene should be replaced by the synthetic heavy chain from the 'donor' vector. Probing of the colonies indicated that only 12 V_λ colonies (5%) and 39 V_κ colonies (16%) retained the original heavy chain. Probing of 250 colonies from each of the V_λ and V_κ phage libraries before recombination indicated that, as expected, all colonies harboured the original heavy chain.

Clones lacking the anti-TNF heavy chain gene (42 V_λ, 48 V_κ) were screened by PCR (Güssow and Clackson, 1989) for the presence of heavy chains with the primers pelBback and CH1.lib.seq (see Table VE) and for the presence of light chains with the primers fdPCRback and Cκ.lib.seq (or Cλ.lib.seq). The probing and PCR screening indicated that in the recombined V_λ library, 28% of clones had acquired a heavy chain from the donor vector and also had a lambda light chain gene; 5% were unrecombined fdDOG-2loxVλlib; and 67% had deletions of the light chain, heavy chain or both. For the V_κ library, 28% of clones had acquired a heavy chain from the donor vector and also had a kappa light chain gene; 16% were unrecombined fdDOG-2loxVκlib; and 56% had deletions of the light chain, heavy chain or both. Further cycles of infection (without selection) led to a further decrease in the frequency of phage harbouring heavy and light chain genes, presumably due to competition with deletion phage. Nevertheless, immediately after recombination, for both V_λ and V_κ repertoires, we can calculate that 28% of all fd phage clones had both heavy and light chain genes and that the heavy chain gene derived from the pUC 'donor' vector.

Propagation of phage from the recombined library

Five litres of 2×TY−TET were inoculated with a 15 ml aliquot of the recombined V_κ library glycerol stock (5×10^{10} c.f.u.) and a further 5 l 2×TY−TET inoculated with a 15 ml aliquot of the recombined V_λ library glycerol stock (1×10^{11} c.f.u.). The cultures were grown, shaking, overnight at 30°C in baffled flasks (1 l medium per flask). The cultures were centrifuged at 5000 *g* for 15 min at 4°C and the fd phage precipitated from the supernatant using polyethylene glycol as in McCafferty *et al.* (1990) and each repertoire resuspended in a final volume of 10 ml PBS. Total phage yields (from 10 l) were typically around 10^{14} t.u.

Selection of the recombined library on immunotubes

The phage repertoire was panned using immunotubes (Nunc; Maxisorp) coated with each antigen (Marks *et al.*, 1991; Griffiths *et al.*, 1993). A range of antigens was used, as described in Table II. Here we have focused on five protein antigens: a mouse monoclonal antibody (NQ11/7.22; Griffiths *et al.*, 1984); and four proteins belonging to the kringle-serine proteases family (HGF/SF, plasmin, t-PA and u-PA). In addition, selection was performed on two haptens conjugated to BSA (FITC Isomer I and NIP). FITC conjugated to BSA (FITC−BSA; 11.2 FITC groups per BSA molecule) was purchased from Sigma; NIP conjugated to BSA (NIP−BSA) was synthesized by coupling NIP−caproate-*O*-succinimide to BSA (Brownstone *et al.*, 1966) to give 27.9 NIP groups per BSA molecule. Tubes were coated with 10 μg/ml protein or 100 μg/ml hapten−BSA conjugates in PBS overnight at room temperature.

For the first round of selection 0.5 ml (6.4×10^{12} t.u.) of the recombined V_κ library and 0.5 ml (7.5×10^{12} t.u.) of the recombined V_λ library were used per immunotube. For the first two rounds of selection tubes were washed 10 times with PBS, 0.1% (v/v) Tween 20 and 10 times with PBS. For subsequent rounds of selection tubes were washed 20 times with PBS, 0.1% (v/v) Tween 20 and 20 times with PBS (Griffiths *et al.*, 1993). Phage were eluted with 100 mM triethylamine (Marks *et al.*, 1991). Eluted phage were used to infect 10 ml log phase *E.coli* TG1 cells and plated on TYE−TET medium in 243×243 mm dishes (Nunc). After incubation overnight at 30°C the colonies were scraped off the plate into 200 ml 2×TY−TET and incubated, shaking, at 30°C for ~6 h. The culture was centrifuged at 5000 *g* for 15 min at 4°C and the fd phage precipitated from the supernatant using polyethylene glycol (McCafferty *et al.*, 1990), each repertoire being resuspended in a final volume of 2 ml PBS. The cell pellet was resuspended in 20 ml 2×TY broth containing 15% glycerol and a 2 ml aliquot stored at −70°C.

One millilitre of these phage (~10^{12} t.u.) was used per immunotube for the next round of selection. The library was subjected to four or five rounds of growth and selection for each antigen.

Selection of the recombined library using streptavidin-coated paramagnetic beads

The library was also selected using soluble biotinylated, hapten−BSA conjugates and streptavidin-coated paramagnetic beads (Hawkins *et al.*, 1992), but with some modifications. FITC−BSA (11.2 FITC:BSA) and NIP−BSA (27.9 NIP−BSA) (see above) were biotinylated using Immunopure NHS-SS-Biotin (sulfosuccinimidyl 2-(biotinamido) ethyl-1,3′-dithiopropionate; Pierce) according to the manufacturer's instructions.

For the first round of selection 0.5 ml (6.4×10^{12} t.u.) of the recombined V_κ library and 0.5 ml (7.5×10^{12} t.u.) of the recombined V_λ library were made up to 2.5 ml with PBS and mixed with 2.5 ml PBS containing 4% skimmed milk powder, 50 μl Tween 20, and biotinylated hapten−BSA added to give a final concentration of 50 nM. The mixture

was then gently rotated on an inclined wheel for 1 h at room temperature. 1.5 ml of Dynabeads M-280 coated with Streptavidin (Dynal) [and previously blocked by incubating for 2 h at 37°C with PBS containing 2% skimmed milk powder (2% MPBS)] were then added and mixing continued for a further 15 min at room temperature. The Dynabeads were then washed a total of 15 times, using a Dynal MPC (Magnetic Particle Concentrator); each wash was with 1 ml PBS or with 1 ml 2% MPBS (every third wash). Phage were eluted from the beads by incubating 5 min at room temperature in 300 µl PBS, 50 mM dithiothreitol (DTT) and the eluate used to infect 10 ml log phase *E.coli* TG1 cells and plated on TYE−TET in 243×243 mm dishes (Nunc). Phage were harvested from the plates as above and each repertoire resuspended in a final volume of 2 ml PBS. The cell pellet was resuspended in 20 ml 2×TY broth containing 15% glycerol and a 2 ml aliquot stored at −70°C (see above).

For the second round (and subsequent rounds) of selection 1 ml of phage (~10^{12} t.u.) were mixed with 0.5 ml PBS containing 6% skimmed milk powder, 10 µl Tween 20, and biotinylated hapten−BSA added to give a final concentration of 50 nM. Selection was then as above, except that only 300 µl of streptavidin-coated Dynabeads M-280 were used. The library was subjected to four or five rounds of growth and selection for each antigen.

ELISA screening of repertoire selections

'Polyclonal' mixtures of phage produced by re-propagation of the library after each round of selection were screened for binding to the antigen used for selection and to other control antigens by ELISA. The phage ELISA was performed essentially as McCafferty *et al.* (1990) using 10 µl PEG precipitated phage (~10^{10} t.u.), but using horseradish peroxidase conjugated anti-sheep antibody (Sigma) and 3',3',5',5'-tetramethylbenzidine (TMB) for detection. Reactions were stopped by the addition of H_2SO_4 after 10 min and readings taken by subtracting the A_{650} from the A_{450}. All antigens were coated at 10 µg/ml in PBS.

Single tetracycline resistant colonies from infection of *E.coli* TG1 with eluted phage were also screened to identify those producing antigen-binding phage by ELISA essentially as in Clackson *et al.* (1991) except that phage were grown at 30°C, and detection was as for the polyclonal phage ELISA above.

Sub-cloning of selected repertoires for expression of soluble Fab fragments

Approximately 10^8 bacteria harbouring phage fd were taken from the stocks frozen down after the appropriate round of selection (5 µl of a 10-fold dilution of the frozen stock). In general, these bacteria were used as template in a 50 µl PCR reaction and the heavy chain genes amplified by pre-soaking at 94°C for 10 min and then cycling 30 times (94°C for 1 min, 50°C for 1 min, 72°C for 2.5 min) using the primers fdSEQ1 and G3LXbaGTGBack (Table VD). The products were run on a 1.3% low melting point agarose gel and purified from the gel using 'Magic PCR Preps' (Promega). The DNA was then cut with *Xba*I and *Not*I and ligated into pUC119His6mycXba. The ligation mixture was electroporated (Dower *et al.*, 1988) into *E.coli* TG1 and plated on TYE medium containing 100 µg/ml ampicillin and 1% glucose (TYE−AMP−GLU) and incubated at 37°C overnight.

Individual ampicillin resistant colonies were grown in 96-well plates and soluble Fab production induced with isopropyl β-D-thiogalactoside (IPTG) as in Marks *et al.* (1991). Soluble Fab fragments in supernatants were assayed for binding to antigen coated plates by ELISA. All antigens were coated at 10 µg/ml in PBS. Bound Fabs were detected with a mixture of rabbit anti-human λ light chain antibodies and rabbit anti-human κ light chain antibodies (Dako) followed by goat anti-rabbit IgG (whole molecule) peroxidase conjugate (Sigma), or by using a mixture of peroxidase conjugated sheep anti-human κ light chain (free and bound) antibodies and peroxidase conjugated sheep anti-human λ light chain (free and bound) antibodies (The Binding Site). ELISAs were developed with TMB as above.

Each antigen-binding clone was streaked on a TYE−AMP−GLU plate and two individual colonies picked and re-assayed for production of antigen-binding Fab fragments as above. Positive clones were stored in 2×TY, 15% glycerol at −70°C.

Sequencing of antibody V genes

Clones, toothpicked from frozen glycerol stocks, were amplified by PCR using Taq polymerase. Reactions (50 µl) were pre-soaked for 10 min at 94°C and then cycled 25 times (94°C for 1 min, 55°C for 1 min, 72°C for 30 s.). The primers were Cκ.lib.seq and LMB3 for kappa chain amplification; Cλ.lib.seq and LMB3 for lambda chain amplification; and CH1.lib.seq and pelBback for heavy chain amplification (Table VE). Aliquots of the product were analysed on a 1.3% agarose gel. The remaining product was purified using Magic PCR Preps (Promega).

PCR cycle sequencing reactions with fluorescent dideoxy chain terminators (Applied Biosystems) (20 µl in 25 cycles: 96°C for 30 s, 50°C for 15 s, 60°C for 4 min) were carried out according to the manufacturer's instructions. The purified PCR product from above (200 ng) was used as template with the following primers: Cκ.lib.seq for kappa chains, Cλ.lib.seq for lambda chains and CH1.lib.seq for heavy chains (Table VE).

The sequencing reactions were analysed on an Applied Biosystems 373A Automated DNA Sequencer. Sequence analysis was performed using SeqEd (Applied Biosystems) and MacVector (IBI Kodak, New Haven, CT).

Epitope mapping of Fab fragments binding monoclonal antibody NQ11/7.22

The repertoire selected on the mouse monoclonal antibody NQ11/7.22 (γ1,κ; Griffiths *et al.*, 1984) was sub-cloned after the second, third and fourth round of selection for production of soluble Fab fragments. 161 out of 384 clones bound to NQ11/7.22 (see above), and were further mapped by binding to MOPC21 (an unrelated mouse γ1, κ monoclonal antibody; Sigma); mouse Fc fragment (Jackson); and NQ11/7.22 zero-linker diabody (containing only the heavy and light chain variable domains; Holliger *et al.*, 1993). All proteins were coated at 10 µg/ml in 50 mM $NaHCO_3$ (pH 8.3). Eleven different Fab fragments identified by sequencing (Table III) were thereby directly mapped to binding to the Fv (idiotypic) or Fc regions, or to the remaining portions of the monoclonal antibody.

Purification of Fabs and affinity determination

Several pUC119His6mycXba clones encoding NIP and fluorescein specific Fab fragments were chosen at random for affinity determination. A 1 l culture of *E.coli* TG1 (Gibson, 1984) harbouring each plasmid was grown and Fab expression induced with IPTG (De Bellis and Schwartz, 1990). After induction, the culture was shaken for 3 h at 25°C and the Fab fragments harvested from the periplasm, essentially as in Breitling *et al.* (1991).

The antibodies were purified by IMAC (Hochuli *et al.*, 1988; Hoffmann and Roeder, 1991). The pooled 'periplasmic fraction' and 'osmotic shock fraction' were passed over a 5 ml Ni−NTA resin (Diagen) according to the manufacturer's instructions. The column was washed with 50 mM sodium phosphate buffer pH 7.5, 500 mM NaCl, 35 mM imidazole and the protein was eluted by applying 50 mM sodium phosphate buffer pH 7.5, 500 mM NaCl and 100 mM imidazole. The eluted protein was dialysed against 2× 3 l PBS for ~24 h. The dialysed fractions were then analysed by SDS−PAGE (Laemmli, 1970) under non-reducing conditions and the concentration determined spectrophotometrically (assuming A_{280} of 1.0 = 0.7 mg/ml).

Affinities of the purified Fabs were determined by fluorescence quench titration with free hapten (Eisen, 1964), essentially as described by Foote and Milstein (1991). The haptens used were fluorescein (Sigma) or NIP−CAP. All measurements were made with a Hitachi F-4500 spectrofluorimeter, using an excitation wavelength of 280 nm and monitoring emission at 340 nm. Antibody (1.0 ml) in PBS was placed in a 10 mm×10 mm cuvette in the instrument, mixed continually using a magnetic stir-bar, and held at 20°C. Hapten additions were made automatically using a 50 µl gastight syringe (Hamilton) driven by a Microlab M syringe controller (Hamilton). The titrations and data collection were performed automatically using a Dell 433/L computer interfaced with the syringe controller (through an RS-232 interface) and the spectrofluorimeter (through an IEEE interface and an instrument driver supplied by the manufacturer). The computer was programmed in QuickBasic (Microsoft). Data were averaged from two to five runs.

The binding of three of the Fab fragments binding protein antigens were analysed by SPR using the BIAcore system (Pharmacia Biosensor). The active concentrations of purified Fab fragments NML1 (anti-NQ11/7.22 Fv); NML9 (anti-NQ11/7.22 Fc) and MH22 (anti-HGF/SF) were determined by the measurement of mass transport-limited binding slopes (Karlsson *et al.*, 1993). The antibody NQ11/7.22 was purified from ascites by affinity chromatography (Mäkelä *et al.*, 1978) and 990 resonance units (RU) immobilized on the biochip by chemical coupling with NHS/EDC (Johnsson *et al.*, 1991; Chaiken *et al.*, 1992). HGF/SF was immobilized on the biochip after biotinylation. Thus 3.8 kRU of immobilized streptavidin (O'Shannessy *et al.*, 1992) were used to capture

2800 RU HGF which had been biotinylated with biotin-LC-hydrazide (Pierce) after periodate oxidation of the carbohydrate using a protocol based on O'Shannessy (1990), but with 20 mM Na_2SO_3 to quench excess periodate (Weber and Hof, 1975). After capture the hydrazone bond between the biotin and the HGF was stabilized by reduction with sodium cyanoborohydride (0.1 M in 0.1 M Na acetate pH 4: 40 µl at 2 µl/min). The fragments were passed over the surface at 5 µl/min at 25°C in concentrations ranging from 2 to 400 nM, and the dissociation constant determined by Scatchard analysis of the equilibrium binding, and on- and off-rate constants by a kinetics analysis software (Pharmacia Biosensor) (Karlsson *et al.*, 1991; Chaiken *et al.*, 1992).

Acknowledgements

We thank M.Yarmolinsky and N.Sternberg for helpful advice and the gift of several phage P1 strains. Also R.A.Mariuzza, S.Songsivilai, M.Bowles, V.V.Mesyanzhinov, K.P.Holliger, R.Hoet, R.de Wildt, C.de Lalla and D.Neri for practical help with the library construction and selection. A.D.G. was supported by the Cancer Research Campaign, O.H. and S.C.W. by MRC studentships, W.L.C. by the Plant Biotechnology Institute, National Research Council of Canada, P.W. by CSIRO Australia, I.M.T. by the Human Genome Mapping Programme and by Trinity College Cambridge, R.E.K. by Deutsche Forschungsgemeinschaft, N.M.L. by The Cambridge Commonwealth Trust, A.N. by the Human Frontier Science Programme Organization (HFSPO), J.P.L.C. by the Nuffield Foundation and M.Z. by the CEC BIOTECH programme.

References

Appleyard,R.K. (1954) *Genetics*, **39**, 440–452.
Barbas,C.F., Kang,A.S., Lerner,R.A. and Benkovic,S.J. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 7978–7982.
Barbas,C.F. *et al.* (1992a) *Proc. Natl Acad. Sci. USA*, **89**, 9339–9343.
Barbas,C.F., Crowe,J.E., Cababa,D., Jones,T.M., Zebedee,S.L., Murphy,B.R., Chanock,R.M. and Burton,D.R. (1992b) *Proc. Natl Acad. Sci. USA*, **89**, 10164–10168.
Barbas,C.F., Bain,J.D., Hoekstra,D.M. and Lerner,R.A. (1992c) *Proc. Natl Acad. Sci. USA*, **89**, 4457–4461.
Barbas,C.F., Amberg,W., Simoncsits,A., Jones,T.M. and Lerner,R.A. (1993) *Gene*, **137**, 57–62.
Bates,R.M., Ballard,D.W. and Voss,E.W. (1985) *Mol. Immunol.*, **22**, 871–877.
Bedzyk,W.D., Reinitz,D.M. and Voss,E.W. (1986) *Mol. Immunol.*, **23**, 1319–1328.
Bothwell,A.L.M., Paskind,M., Reth,M., Imanishi-Kari,T., Rajewsky,K. and Baltimore,D. (1981) *Cell*, **24**, 625–637.
Breitling,F., Dübel,S., Seehaus,T., Klewinghaus,I. and Little,M. (1991) *Gene*, **104**, 147–153.
Brownstone,A., Mitchison,N.A. and Pitt-Rivers,R. (1966) *Immunology*, **10**, 465–481.
Chaiken,I., Rosé,S. and Karlsson,R. (1992) *Anal. Biochem.*, **201**, 197–210.
Chothia,C. and Lesk,A.M. (1987) *J. Mol. Biol.*, **196**, 901–917.
Chothia,C. *et al.* (1989) *Nature*, **342**, 877–883.
Chothia,C., Lesk,A.M., Gherardi,E., Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) *J. Mol. Biol.*, **227**, 799–817.
Chuchana,P., Blancher,A., Brockly,F., Alexandre,D., Lefranc,G. and Lefranc,M.P. (1990) *Eur. J. Immunol.*, **20**, 1317–1325.
Clackson,T., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1991) *Nature*, **352**, 624–628.
Cox,J.P.L., Tomlinson,I.M. and Winter,G. (1994) *Eur. J. Immunol.*, **24**, 827–836.
Cumano,A. and Rajewsky,K. (1985) *Eur. J. Immunol.*, **15**, 512–520.
Cumano,A. and Rajewsky,K. (1986) *EMBO J.*, **5**, 2459–2468.
De Bellis,D. and Schwartz,I. (1990) *Nucleic Acids Res.*, **18**, 1311.
Denzin,L.K. and Voss,E.W. (1992) *J. Biol. Chem.*, **267**, 8925–8931.
Dower,W.J., Miller,J.F. and Ragsdale,C.W. (1988) *Nucleic Acids Res.*, **16**, 6127–6145.
Eisen,H.N. (1964) *Methods Med. Res.*, **10**, 115–121.
Foote,J. and Milstein,C. (1991) *Nature*, **352**, 530–532.
Garrard,L.J., Yang,M., O'Connell,M.P., Kelley,R.F. and Henner,D.J. (1991) *Bio/Technology*, **9**, 1373–1377.
Gibson,T.J. (1984) *PhD thesis*, University of Cambridge, UK.
Griffiths,G.M., Berek,C., Kaartinen,M. and Milstein,C. (1984) *Nature*, **312**, 271–275.
Griffiths,A.D. *et al.* (1993) *EMBO J.*, **12**, 725–734.
Güssow,D. and Clackson,T. (1989) *Nucleic Acids Res.*, **17**, 4000.
Hawkins,R.E., Russell,S.J. and Winter,G. (1992) *J. Mol. Biol.*, **226**, 889–896.
Hochuli,E., Bannwarth,W., Döbeli,H., Gentz,R. and Stüber,D. (1988) *Bio/Technology*, **6**, 1321–1325.
Hoess,R.H., Ziese,M. and Sternberg,N. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 3398–3402.
Hoess,R.H., Wierzbicki,A. and Abremski,K. (1986) *Nucleic Acids Res.*, **14**, 2287–2300.
Hoffmann,A. and Roeder,R.G. (1991) *Nucleic Acids Res.*, **19**, 6337–6338.
Holliger,P., Prospero,T. and Winter,G. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6444–6448.
Hoogenboom,H.R. and Winter,G. (1992) *J. Mol. Biol.*, **227**, 381–388.
Hoogenboom,H.R., Griffiths,A.D., Johnson,K.S., Chiswell,D.J., Hudson,P. and Winter,G. (1991) *Nucleic Acids Res.*, **19**, 4133–4137.
Horton,R.M., Hunt,H.D., Ho,S.N., Pullen,J.K. and Pease,L.R. (1989) *Gene*, **77**, 61–68.
Huse,W.D., Sastry,L., Iverson,S.A., Kang,A.S., Alting,M.M., Burton,D.R., Benkovic,S.J. and Lerner,R.A. (1989) *Science*, **246**, 1275–1281.
Johnsson,B., Löfås,S. and Lindqvist,G. (1991) *Anal. Biochem.*, **198**, 268–277.
Kabat,E.A., Wu,T.T., Perry,H.M., Gottesman,K.S. and Foeller,C. (1991) *Sequences of Proteins of Immunological Interest*. 5th edn, US Dept of Health and Human Services, Bethesda, USA.
Karlsson,R., Michaelsson,A. and Mattsson,L. (1991) *J. Immunol. Methods*, **145**, 229–240.
Karlsson,R., Fägerstam,L., Nilhans,H. and Persson,B. (1993) *J. Immunol. Methods*, **166**, 75–84.
Kranz,D.M. and Voss,E.W. (1981) *Mol. Immunol.*, **18**, 889–898.
Kranz,D.M., Herron,J.N. and Voss,E.W. (1982) *J. Biol. Chem.*, **257**, 6987–6995.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lucisano-Valim,Y.M. and Lachmann,P.J. (1991) *Clin. Exp. Immunol.*, **84**, 1–8.
Mäkelä,O., Kaartinen,M., Pelkonen,J.L.T. and Karjalainen,K.J. (1978) *J. Exp. Med.*, **148**, 1644–1660.
Mariuzza,R. and Strand,M. (1981) *Mol. Immunol.*, **18**, 847–855.
Marks,J.D., Hoogenboom,H.R., Bonnert,T.P., McCafferty,J., Griffiths,A.D. and Winter,G. (1991) *J. Mol. Biol.*, **222**, 581–597.
Marks,J.D., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1992a) *J. Biol. Chem.*, **267**, 16007–16010.
Marks,J.D., Griffiths,A.D., Malmqvist,M., Clackson,T., Bye,J.M. and Winter,G. (1992b) *Bio/Technology*, **10**, 779–783.
Marks,J.D., Ouwehand,W.H., Bye,J.M., Finnern,R., Gorick,B.D., Voak,D., Thorpe,S., Hughes-Jones,N.C. and Winter,G. (1993) *Bio/Technology*, **11**, 1145–1149.
McCafferty,J., Griffiths,A.D., Winter,G. and Chiswell,D.J. (1990) *Nature*, **348**, 552–554.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
Munro,S. and Pelham,H.R.B. (1986) *Cell*, **46**, 291–300.
Nissim,A., Hoogenboom,H.R., Tomlinson,I.M., Flynn,G., Midgley,C., Lane,D. and Winter,G. (1994) *EMBO J.*, **13**, 692–698.
Nossal,G.J. (1993) *Ann. N.Y. Acad. Sci.*, **690**, 34–41.
O'Shannessy,D.J. (1990) *Methods Enzymol.*, **184**, 162–166.
O'Shannessy,D.J., Brigham-Burke,M. and Peck,K. (1992) *Anal. Biochem.*, **205**, 132–136.
Perelson,A.S. and Oster,G.F. (1979) *J. Theor. Biol.*, **81**, 645–670.
Rathjen,D.A., Furphy,L.J. and Aston,R. (1992) *Br. J. Cancer*, **65**, 852–856.
Reinitz,D.M. and Voss,E.W. (1984) *Mol. Immunol.*, **21**, 775–784.
Rosner,J.L. (1972) *Virology*, **48**, 679–689.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1990) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
Schwartz,R.S. and Stollar,B.D. (1994) *Immunol. Today*, **15**, 27–32.
Skerra,A., Dreher,M.L. and Winter,G. (1991) *Anal. Biochem.*, **196**, 151–155.
Smith,G.P. (1985) *Science*, **228**, 1315–1317.
Sternberg,N. and Hamilton,D. (1981) *J. Mol. Biol.*, **150**, 467–486.
Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) *J. Mol. Biol.*, **227**, 776–798.
Tomlinson,I.M., Cook,G.P., Carter,N.P., Elaswarapu,R., Smith,S., Walter,G., Buluwela,L., Rabbitts,T.H. and Winter,G. (1994) *Human Mol. Gen.*, **3**, 853–860.

Vasicek,T.J. and Leder,P. (1990) *J. Exp. Med.*, **172**, 609–620.
Vieira,J. and Messing,J. (1987) *Methods Enzymol.*, **153**, 3–11.
Waterhouse,P., Griffiths,A.D., Johnson,K.S. and Winter,G. (1993) *Nucleic Acids Res.*, **21**, 2265–2266.
Weber,P. and Hof,L. (1975) *Biochem. Biophys. Res. Commun.*, **65**, 1298–1302.
Williams,S.C. and Winter,G. (1993) *Eur. J. Immunol.*, **23**, 1456–1461.
Winter,G., Griffiths,A.D., Hawkins,R.E. and Hoogenboom,H.R. (1994) *Annu. Rev. Immunol.*, **12**, 433–455.
Wu,T.T., Johnson,G. and Kabat,E.A. (1993) *Proteins*, **16**, 1–7.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Yarmolinsky,M.B., Hansen,E.B., Jafri,S. and Chattoraj,D.K. (1989) *J. Bacteriol.*, **171**, 4785–4791.

Received on March 2, 1994; revised on May 5, 1994