© 1993 Oxford University Press
Nucleic Acids Research, 1993, Vol. 21, No. 19 4491–4498

Efficient method for constructing comprehensive murine Fab antibody libraries displayed on phage

Henrik Ørum[1], Peter S.Andersen[1,2], Anne Øster[1], Lene K.Johansen[1], Erik Riise[1], Mads Bjørnvad[2], Ivan Svendsen[2] and Jan Engberg[1,*]

[1]The Royal Danish School of Pharmacy, Department of Biology, Universitetsparken 2, DK 2100 Copenhagen and [2]The Monoclonal Antibody Laboratory, 6B3.61, Novo Nordisk A/S, Novo Alle, DK 2880 Bagsværd, Denmark

Received June 24, 1993; Revised and Accepted August 11, 1993

ABSTRACT

We have developed efficient methodologies for construction and expression of comprehensive phage display libraries of murine Fab antibody fragments in *E.coli* cells. Our methods optimize several critical steps of the polymerase chain reaction (PCR) amplification of transcripts of the re-arranged immunoglobulin genes and of their subsequent assembly and expression: Firstly, we have designed exhaustive sets of PCR primers of low degeneracy for the amplification of transcripts of the Fab region of the heavy and light-chain genes. These primers proved effective in amplification of Fab gene fragments from a large panel of hybridoma cell lines of different specificity and family sub-type. Secondly, we have developed a 'jumping PCR' technique that effectively assembled and recombined the amplified heavy and light-chain gene fragments into a bi-cistronic operon. Thirdly, we have constructed expression vectors for insertion of the combinatorial Fab gene-cassette in fusion with a truncated version of the phage surface protein, gIIIp. The heavy chain and the light chain-gIII fusion are transcribed as a polycistronic mRNA from the *lacZ* promoter and efficient transcriptional control is provided by wildtype *lacI* present on the vector. The utility of the system was demonstrated by isolating several antigen-binding clones from hybridomas and libraries made from immunized mice.

INTRODUCTION

Bacteriophage display systems have been developed that link a protein or peptide of interest to the DNA that encodes it. Such display systems have been used to screen peptide libraries for binding to selected target molecules (1–5) and to display functional proteins with the potential of screening these proteins for desired properties (6,7). Recently, improvements of the display approach have made it possible to express libraries of antibody fragments on the surface of bacteriophage thus allowing for selection of specific antibody-expressing phages by screening with antigen (reviewed in 8–9). The antibody-display approach uses phagemids that co-express combinations of randomly assembled pairs of heavy(H) and light(L)-chain genes. The membrane anchorage domain of the phage surface protein, gIIIp, is fused to the H(or L)-chain and the resulting protein becomes anchored in the *E.coli* membrane. Expression of the L(or H)-chain from the phagemid DNA as a soluble periplasmic protein permits Fab assembly within the periplasmic space and upon infection with helper phage, the Fab-gIIIp fusion proteins become displayed on the phage surface by displacing some of the wild-type gIII proteins (10–13). Many of the published protocols for phage antibody display (14) are based on model systems where previously cloned antibody gene fragments were tested. Although these protocols have demonstrated the general feasibility of the approach, they often include technical steps which lead to limited or biased libraries; for example, restriction enzymes are being used that recognize a significant portion of the known antibody gene repertoire. We have been concerned with the design of an improved method of generating comprehensive and unbiased antibody libraries from mouse spleen mRNA preparations and have addressed the limitations of the different steps involved in the current phage display technology. As a result we have improved the system by designing comprehensive sets of PCR primers corresponding to the N-terminus of the variable regions, by devising an assembly reaction that links the H and L-chain gene fragments into a composite bi-cistronic fragment without the use of restriction enzymes, and by constructing cloning vectors that keep the expression of the Fab antibodies tightly regulated.

The system has been used successfully for cloning of Fab-expressing phage from two monoclonal hybridoma cell lines (SAL10F68 and HUI018) that recognize a *Salmonella* surface antigen and human insulin, respectively, and for cloning of several highly antigen-specific phages from a Fab library made from mice immunized with human blood-clotting factor VII.

* To whom correspondence should be addressed

MATERIAL AND METHODS

E.coli strains

JM 103: *thi*$^-$*,strA,supE,endA,sbcB,hsdR*$^-$*Δ(lac-proAB)*, *F*'{traD36,proAB,lacI^q,lacZΔ15}
TOP10F': *mcrA,Δ(mrr-hsdRMS-mcrBC)θ80ΔlacX74,deo R,recA1,araD1,Δ(ara.leu), 7697, galU,galK,*λ$^-$,rpsL,endA1, nupG,F'{tet^r}. (British Biotechnology).
XL-1 Blue: *recA1,endA1,gyrA96,thi,hsdR17(r_k^-,m_k^+),supE44, relA1.*λ$^-$,(lac),F'{proAB,lacI^q,lacZΔ*M15,Tn10(tet^r*}. (Stratagene).

PCR primers

An overview of the different primers used in this study is presented in Figure 1. Detailed sequence information on primers and expression vectors can be obtained upon request from the corresponding author via E-mail (JE@CHARON.DFH.DK).

Heavy chain primers

MVH 1–25: These primers consist of an equimolar mixture of 25 individually synthesized oligos with the sequences:

```
MVH-1   G  CCG GCC ATG GCC GAG GTR MAG CTT CAG GAG TCA GGA C
MVH-2   G  CCG GCC ATG GCC GAG GTS CAG CTK CAG CAG TCA GGA C
MVH-3   G  CCG GCC ATG GCC CAG GTG CAG CTG AAG SAS TCA GG
MVH-4   G  CCG GCC ATG GCC GAG GTG CAG CTT CAG GAG TCS GGA C
MVH-5   G  CCC GCC ATG GCC GAR GTC CAG CTG CAA CAG TCY GGA C
MVH-6   G  CCC GCC ATG GCC CAG GTC CAG CTK CAG CAA TCT GG
MVH-7   G  CCG GCC ATG GCC CAG STB CAG CTG CAG CAG TCT GG
MVH-8   G  CCC GCC ATG GCC CAG GTY CAG CTG CAG CAG TCT GGR C
MVH-9   G  CCG GCC ATG GCC GAG GTY CAG CTY CAG CAG TCT GG
MVH-10  G  CCG GCC ATG GCC GAG GTC CAR CTG CAA CAA TCT GGA CC
MVH-11  G  CCG GCC ATG GCC CAG GTC CAC GTG AAG CAG TCT GGG
MVH-12  G  CCG GCC ATG GCC GAG GTG AAS STG GTG GAA TCT G
MVH-13  G  CCG GCC ATG GCC GAV GTG AAG YTG GTG GAG TCT G
MVH-14  G  CCG GCC ATG GCC GAG GTG CAG SKG GTG GAG TCT GGG G
MVH-15  G  CCG GCC ATG GCC GAK GTG CAM CTG GTG GAG TCT GGG
MVH-16  G  CCG GCC ATG GCC GAG GTG AAG CTG ATG GAR TCT GG
MVH-17  G  CCG GCC ATG GCC GAG GTG CAR CTT GTT GAG TCT GGT G
MVH-18  G  CCG GCC ATG GCC GAR GTR AAG CTT CTC GAG TCT GGA
MVH-19  G  CCG GCC ATG GCC GAA GTG AAR STT GAG GAG TCT GG
MVH-20  G  CCG GCC ATG GCC GAA GTG ATG CTG GTG GAG TCT GGG
MVH-21  G  CCG GCC ATG GCC CAG GTT ACT CTR AAA GWG TST GGC C
MVH-22  G  CCG GCC ATG GCC CAG GTC CAA CTV CAG CAR CCT GG
MVH-23  G  CCG GCC ATG GCC CAG GTY CAR CTG CAG CAG TCT G
MVH-24  G  CCG GCC ATG GCC GAT GTG AAC TTG GAA GTG TCT GG
MVH-25  G  CCG GCC ATG GCC GAG GTG AAG GTC ATC GAG TCT GG
```

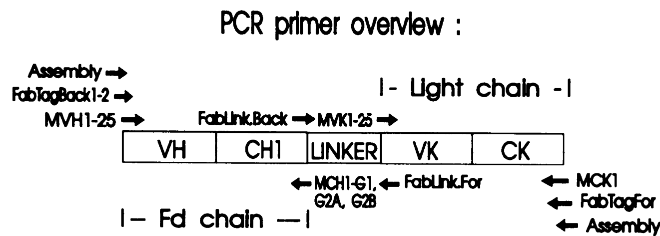

Figure 1. Oligonucleotide primers used to generate Fab gene fragments by PCR. The boxes delineate the different gene-segments included in the construction of the Fab expression-cassette: Fd, includes the H-chain from the N-terminal amino acid to the cysteine residue of the hinge region which forms the di-sulphide bridge to the C-terminal cysteine of the L-chain; Light chain, corresponds to the entire variable and constant parts of the light (Kappa) chain; LINKER, a 117 bp fragment containing a translational stop codon for Fd translation, a ribosome binding site for L-chain expression and the coding region corresponding to the N-terminal part of the *Pel*B leader. Primers depicted below the boxes are forward primers and complementary to mRNA. Primers above the boxes are back-primers and complementary to 1st strand cDNA.

(The one-letter nucleotide symbols are used according to IUB nomenclature).

The mixture represents 88 variants. The nucleotides in bold correspond to the N-terminal part of the variable heavy (VH) sequences starting with amino acid number 1. Sequence numbers 1 to 13 correspond to the C-terminal part of the *pelB* leader. The size of the individual primers range from 35 to 39 nucleotides and have Tm's (melting temperatures of duplex DNA) between 62°C and 80°C with their target sequence. Tm's were estimated using the formula: Tm = (4(G+C) + 2(A+T))°C, where G,C,A and T represent the overall number of the corresponding bases.

FabTag.BACK-1: 5' CA GTC ACA GATCCT CGC GAA TTG GCC CAG CCG GCC **ATG GCC SAN G** 3'
FabTag.BACK-2: 5' CA GTC ACA GAT CCT CGC GAA TTG GCC CAG CCG GCC **ATG GCC SAN C** 3'. The nucleotides in bold overlap with the MVH primers with a Tm of 60°C. The *Sfi*I recognition site (cf Figure 2) is underlined. The primers were used as an equimolar mixture.
ASSEMBLY-1: 5' **CA GTC ACA GAT CCT CGC GAA TTG G** 3'. The nucleotides in bold overlap with FabTag.BACK-1 and 2, pos. 1 to 24, with a Tm of 72°C. MCH1-G1: 5'CGACTAGTTTAGAATTCAAGCTGTCGAC TCA **ACA ATC CCT GGG CAC AAT TTT CTT GTC CAC C** 3'
MCH1-G2A: 5'CGACTAGTTTAGAATTCAAGCTGTCGAC TCA **ACA GGG CTT GAT TGT GGG CCC TCT GGG** 3'
MCH1-G2B: 5'CGACTAGTTTAGAATTCAAGCTGTCGAC TCA **ACA GGG GTT GAT TGT TGA AAT GGG CCC G** 3'

Nucleotides in bold are complementary to the CH1 and hinge junction regions of the γ1, γ2a and γ2b H-chain isotypes with Tm's ranging between 86 and 92°C. The stop codon is underlined and the sequence pos. 1 to 26 are complementary to the 5' end of the LINK-D fragment, pos. 1 to 26, with a Tm of 88°C. The primers were used as an equimolar mixture.

Light chain primers

MVK 1–25: These primers consist of an equimolar mixture of 25 individually synthesized oligos with the sequences:

```
MVK-1   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTT CTC ACC CAG TCT CC
MVK-2   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG CTS ACC CAG TCT CC
MVK-3   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACT CAG TCT CC
MVK-4   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG CTM ACT CAG TCT CC
MVK-5   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG YTR ACA CAG TCT CC
MVK-6   TTGGCTGCACAACCAGCAATGGCA GAC ATT GTR ATG ACA CAG TCT CC
MVK-7   TTGGCTGCACAACCAGCAATGGCA GAC ATT MAG ATR ACC CAG TCT CC
MVK-8   TTGGCTGCACAACCAGCAATGGCA GAC ATT CAG ATG AMC CAG TCT CC
MVK-9   TTGGCTGCACAACCAGCAATGGCA GAC ATT CAG ATG ACD CAG TCT CC
MVK-10  TTGGCTGCACAACCAGCAATGGCA GAC ATT CAG ATG ACA CAG ACT AC
MVK-11  TTGGCTGCACAACCAGCAATGGCA GAC ATT CAG ATG ATT CAG TCT CC
MVK-12  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTT CTC AWC CAG TCT CC
MVK-13  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTT CTC TCC CAG TCT CC
MVK-14  TTGGCTGCACAACCAGCAATGGCA GAC ATT GWG CTS ACC CAA TCT CC
MVK-15  TTGGCTGCACAACCAGCAATGGCA GAC ATT STG ATG ACC CAR TCT C
MVK-16  TTGGCTGCACAACCAGCAATGGCA GAC ATT KTG ATG ACC CAR ACT CC
MVK-17  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACT CAG GCT AC
MVK-18  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACB CAG GCT GC
MVK-19  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATA ACY CAG GAT G
MVK-20  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACC CAG TTT GC
MVK-21  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACA CAA CCT GC
MVK-22  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACC CAG ATT CC
MVK-23  TTGGCTGCACAACCAGCAATGGCA GAC ATT TTG CTG ACT CAG TCT CC
MVK-24  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTA ATG ACC CAA TCT CC
MVK-25  TTGGCTGCACAACCAGCAATGGCA GAC ATT GTG ATG ACC CAC ACT CC
```

The mixture represents 50 variants. Nucleotides in bold correspond to the N-terminal part of the L-kappa sequences starting with amino acid number 1. The individual primers range in size between 46 to 47 nucleotides with Tm's of 50 to 68°C with their target sequences. Sequence pos. 1 to 24, corresponding

to the C-terminal part of the *pelB* leader, overlaps the LINK-D fragment, pos. 94 to 117, with a Tm of 84°C.

MCK1:5′ TGC GGC CGC **ACA CTC ATT CCT GTT GAA GCT CTT GAC** 3′. Sequences in bold are complementary to the C-terminus of the constant part of the L-kappa chain with a Tm of 78°C. The *Not*I recognition site (cf Figure 2) is underlined.

FabTAG.FORW-1: 5′CAG TCA CAG ATC CTC GCG AAT TGG **TGC GGC CGC ACA CTC ATT CCT G** 3′. The sequences in bold are complementary to MCK1, pos. 1 to 22, with a Tm of 72°C. The *Not*I site is underlined.

ASSEMBLY-1: 5′ CAG TCA CAG ATC CTC GCG AAT TGG 3′. This sequence is complementary to the 5′ end of FabTAG.FORW-1 (and to FabTAG.BACK-1 and 2) with a Tm of 74°C.

LINKER primers

FabLINK.FORW: 5′ GTC TGC CAT TGC TGG TTG TGC AGC CAA 3′. This sequence is complementary to the 3′ end of the LINK-D fragment.

FabLINK.BACK: 5′ CGA CAG CTT GAA TTC TAA ACT AGT CGA AGG CGC GCC AAG GAG ACA GTC AT 3′. This sequence overlaps the 5′ end of the LINK-D fragment.

LINK-D fragment: 5′**CGACAGCTTGAATTCTAAACTAG-TCG**AAGGCGCGCCAAGGAGACAGTCATA ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA **TTG GCT GCA CAA CCA GCA ATG GCA** 3′. The highlighted sequence at the 5′ end overlaps with the MHC-primers, while the highlighted sequence at the 3′ end is complementary to the MVK-primers. The ribosome binding site and the ATG triplet marking the start of the *pelB* leader sequence is underlined. The LINK-D fragment originates from the λc2 vector (16) which was taken through several PCR cloning steps using different sets of tagged primers in order to introduce the sequence changes necessary for the present assembly system. The LINKER fragment was cloned into the *Sma*I site of the Bluescript KS+ vector (Stratagene) and subsequently amplified via PCR using the primers FabLINK.BACK and FabLINK.FORW.

All primers in the present study were synthesized on an Applied Biosystems oligonucleotide machine, model 394, and subsequently checked by electrophoresis on sequencing gels. Only high quality preparations containing substantial amounts of full-size product were used for further applications.

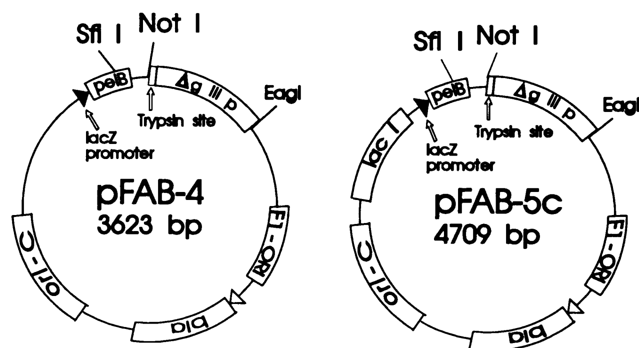

Figure 2. Phagemid vectors used for expression of antibody fragments. The Fab antibody-cassette is introduced as a *SfiI-NotI* fragment and is expressed from the inducible *LacZ* promoter. For an explanation of their construction, see the text.

Vector constructions

The two expression vectors used in this study are depicted in Figure 2. The vector pFAB4 was derived from pBluescript KS+: After deletion of the unique *Not*I site by cutting and filling-in with Klenow polymerase, the modified vector was used for cloning of the *Hind*III-*Eco*RI fragment obtained from pHEN1 (kind gift of Dr H.R. Hoogenboom, cf Ref. 17). This fragment has internal *Sfi*I and *Not*I sites allowing for fragments to be inserted in fusion with the *pelB* signal peptide and the gIII protein, respectively. The *Not*I-*Eco*RI fragment of this new vector was replaced with a *Not*I-*Eco*RI fragment containing a truncated version of the gIIIp gene obtained by PCR of the helper phage R408 (Stratagene) using primers recognizing the sequence from Pro 198 to the C-terminal amino acid of gIIIp. Furthermore, the primers used for this amplification were designed to introduce a trypsin cleavage site between the *Not*I site and the truncated gIIIp gene. Finally, the 131 bp region between the start codons of the *lacZ* and the *pelB* leader was deleted by site-directed mutagenesis.

The pFAB5c vector is a derivative of pFAB4 with the *lacI* operon inserted: *Sac*I and *Spe*I sites were introduced by site-directed mutagenesis around the truncated *lacI* gene which is present in pBluescript vectors and derivatives thereof. Primers tagged with these restriction sites were used to PCR amplify the wild type *lacI* operon from *E.coli* DH5α (Stratagene) and the fragment was cloned into the modified pFAB4 vector. Finally, an adaptor containing several restriction sites for further applications was inserted between the *Sfi*I and *Not*I sites.

cDNA synthesis, PCR and cloning

Spleens of BALB/c mice boosted six weeks after primary immunization or samples (10^7 cells) of the hybridoma cell lines SAL10F68 and HUI018 were used for RNA isolation according to the method of Chirgwin *et al* (18). First-strand cDNA was primed with oligo-$dT_{n=18}$ using SuperScript Plus (Gibco BRL) reverse transcriptase in a reaction mixture incubated at 42°C for one hour. Following heat inactivation of the transcriptase, aliquots were used directly for primary PCR amplifications of Fd and L-chains. PCR reactions were performed in 100μl volumes containing dNTP's (0.2 mM), and reaction-buffer supplied by the manufacturer (Cetus) and primers and cDNA. Fd reactions contained MVH 1−25 (0.227 pmoles of each of the 88 variants) and 6.7 pmoles of each of the MCH γ1, γ2A and γ2B primers. L-chain reactions contained MVK 1−25 (0.4 pmoles each of the 50 variants) and 20 pmoles of the MCK primer. Reactions mixtures were overlaid with mineral oil and kept at 94°C for 2 min. Then 1.5U of Supertaq (HT Biotechnology) or Amplitaq (Cetus) was added and the mixtures cycled 30 times (94°C 1 min, 55°C 1 min, 72°C 2 min) followed by incubation at 72°C for 10 min using a PREM III thermocycler. The primary PCR amplification products were purified by preparative agarose gel electrophoresis in combination with GeneClean (Bio101, Inc.) procedures.

'Jumping' PCR amplifications

Assembly of Fd genes with LINK-D: 100 μl reactions contained buffer and dNTPs as above and five ng purified Fd gene fragments, two ng purified LINK-D, FabTAG.BACK 1−2 (25 pmoles each) and 50 pmoles FabLINK.FORW.

Assembly of L-chain genes with LINK-D: 100 μl reactions contained buffer and dNTPs as above, five ng of purified L-chain

genes, two ng of purified LINK-D and 50 pmoles of FabLINK.BACK and FabTAG.FOR respectively. The thermocycling program was initiated as above, cycled 25 times (94°C 2 min, 55°C 1 min, 72°C 2 min) and the assembled product was gel-purified.

Final assembly

100 μl reactions contained buffer and dNTPs as above and five ng purified Fd-LINK-D fragment, five ng purified L-chain-LINK-D fragment and 20 pmoles ASSEMBLY-1 primer. The thermocycling program was initiated as above, run 25 times (94°C 1½ min, 69°C 1 min, 72°C 2 min) and the assembled product was gel-purified.

Cloning

DNA fragments from final assembly reactions were digested with *Not*I using 10 U per μg of DNA at 37°C for 2 hours. Following phenol extractions and ethanol precipitation, the DNA was dissolved in *Sfi*I buffer and incubated under oil at 50°C for 2 hours with 10 U enzyme per μg DNA. The DNA was purified using GeneClean procedures and ligated to GeneClean-purified *Not*I and *Sfi*I-cut pFAB4 or pFAB5c. The 20μl ligation reaction included 0.5 μg of digested vector, 0.5 μg of insert DNA and was incubated overnight at 15°C with 1.5 U T_4-DNA ligase (Amersham). The ligation mix was then purified by phenol extraction and ethanol precipitation followed by resuspension in 20 μl water. Portions of two μl were electrophorated into appropriate *E.coli* strains using a Bio-Rad *E.coli* pulser set at 25 μF, 2.5 kV and 200 Ohms. Immediately after the pulse, one ml of freshly made SOC medium (15) was added and the cells were shaken for one hour at 37°C and plated on selective plates or processed directly for superinfection.

Rescue of phagemids from individual clones or libraries

100 ml flasks containing 10 ml 2×TY broth (15) and ampicillin (100 μg/ml) or tetracycline (10μg/ml) were inoculated with 10^8 transformed cells and shaken at 37°C at 250 rpm. At an OD_{600} of 0.5, helper phage, R408 (Stratagene) was added at a multiplicity of 50 and infection proceeded for 20 min at 37°C without shaking. The cells were then added isopropyl-β-D-thiogalactopyranoside (IPTG) and incubation was continued overnight at room temperature with shaking (250 rpm). Cells were pelleted and phage particles were used directly from the supernatant or purified and concentrated by PEG precipitations (15) followed by resuspension in PBS (20 mM KH_2PO_4, 8 mM Na_2HPO_4, 130 mM NaCl, pH 7.3). Phage titrations were done with male *E.coli* strains using standard procedures (15). When producing libraries, the above volumes were scaled up by a factor of 10.

ELISA

Binding of Fab-phages or free Fabs to human insulin (Achtrapid, 40000 IU/ml, Novo Nordisk), human factor VII (prepared at Novo Nordisk), *Salmonella* flagellin (prepared at Novo Nordisk) was detected using standard ELISA techniques. Briefly, wells of microtiter plates (Maxisorp, NUNC) were coated with 100 μl antigen solutions (10 μg/ml i PBS) overnight at 4°C followed by blocking with PBS containing 2% skimmed milk powder (Difco) for 2 hrs. Serial dilutions of phages or cell homogenates in blocking solution were added to the wells and incubated with rocking between one and two hrs at room temperature. Washings between the different steps were done with PBS containing 0.05% Tween-20. Cell homogenates were prepared from 20 times concentrated cell suspensions by sonication using a microtip equipped Branson Sonifier operated at 60 W at 0°C for 10 consecutive 10 sec pulses. Cell debris were spun down and the supernatant used directly. Binding was assayed using biotinylated goat anti-mouse-Fab (Sigma) and peroxidase-conjugated streptavidin or peroxidase-conjugated rabbit anti-M13 (Pharmacia) followed by detection with ortho-phenylen-diamine (DAKO) and H_2O_2 reacting for 10−30 min before stopping the reaction with H_2SO_4. Plates were read at OD_{490}.

Selection of antigen binders by panning

Microtiter wells coated with antigen were added 10^{10} phages in blocking solution and incubated for two hours with rocking at room temperature. After washings, phages were eluted using one of two methods: 1) 10 min incubation at room temperature with 100 μl 0.1 M glycine-HCl (pH 2.2) containing 1mg/ml BSA (2), or 2) 10 min incubation at room temperature with 100 μl PBS containing 0.125% trypsin (Difco) and 1 mM EDTA. Eluates from 1) were neutralized, eluates from 2) were diluted and both were then propagated for re-infection, titration and storage at 4°C.

Affinity determinations

The apparent binding constants of the soluble Fab fragments and the Fab-phages were determined by inhibition ELISAs (20). Briefly, dilutions of Fab preparations corresponding an ELISA OD_{490} reading of about 2.0 were chosen and mixed with increasing amounts of antigen before transfer to antigen-coated ELISA plates, incubated for 2 hrs with rocking at room temperature and processed as for normal ELISAs. Apparent affinities were determined as the reciprocal of the antigen concentration required to inhibit 50% maximal binding.

RESULTS

Design of amplification primers

A collection of 90 H and 70 L-chain murine IgG sequences was extracted from the Kabat and Genebank/EMBL databases and used for our design of consensus PCR primers. The collection consists of antibodies with different specificities and represents all known VH groups and VL subgroups as defined by Kabat (21). The variable regions were aligned with respect to their N-terminal 23 to 25 nucleotides and checked for sequence homology. From this analysis, we identified 25 subgroups which satisfied the following criteria: members of a subgroup differed from the other members at no more than two positions, subgroup members were identical at their 3′ end for a stretch of at least five nucleotides, and the primers ended at a position corresponding to a conserved amino acid sequence. Primers for the H-chain isotypes γ1, γ2A and γ2B and the kappa L-chain were designed based on their established unique sequences. The L-chain primers include the C-terminal part of the L-chain, whereas the heavy-constant (HC) primers straddle the CH1 and hinge regions and include the cysteine residue that forms part of the di-sulfide bridge with the C-terminal cysteine of the kappa chain. When properly assembled, the chains generated from these gene fragments constitute complete Fab fragments.

To test our primer design, we used the H or L-chain sets of primers to amplify the Fab sequences of mRNA isolated from a series of 15 hybridoma cell lines of different specificity and family sub-type. A clear difference in quality was observed in

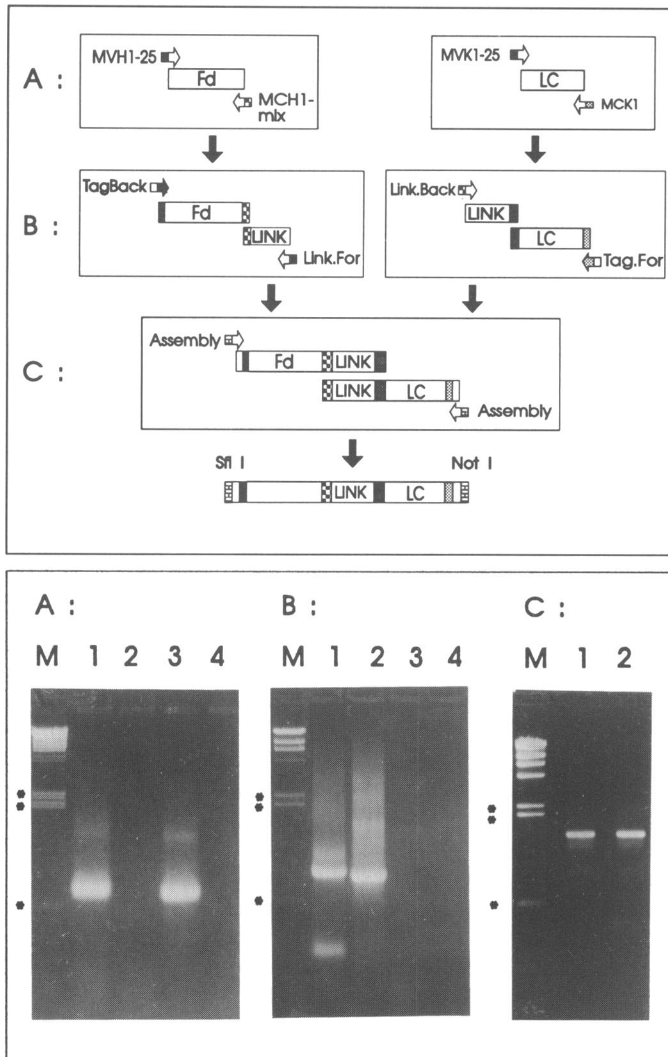

Figure 3. Construction of combinatorial Fab libraries using the 'jumping-PCR-assembly' method. **Upper panel**: Diagram exemplifying the primary amplification PCR (**A**), the linker assembly (**B**) and the final assembly (**C**), respectively. The frames indicate separate reactions. The boxes and the horizontal arrows are explained in Figure 1. Marked zones within boxes represent regions of overlapping complementarity. **Lower panel**: Analysis by electrophoresis on 1% agarose gels of PCR products generated as explained above. The different M-lanes are λ-DNA digested with *Hind*III. The positions corresponding to the 2.2, 2.0 and 0.56 kb marker fragments are indicated with asterisks. **A**: Lanes 1 and 3, Fd and L-chain repertoire PCR products from immunized mice. Lanes 2 and 4 are PCR reactions without templates (negative controls). **B.** Lane 1, PCR assembly product of purified Fd fragment (from A) and LINK-D fragment (cf Materials and Methods). Lane 2, as in lane 1 but with purified L-chain gene fragments (from A). Lanes 3 and 4, negative controls. **C**: PCR assembly products of gel-purified Fd-LINK-D and L-chain-LINK-D fragments obtained from hybridoma (Lane 1) or spleen (Lane 2) material.

the two sets of reactions when 45°C (Tm minus 15–20°C) was used as the annealing temperature: the L-chain amplifications resulted in distinct products of predicted size (about 650 bp) whereas the H-chain amplification reactions frequently resulted in fragments of odd sizes or no fragments at all. However, when amplifications were attempted using the MVH-primers and a set of forward primers corresponding to the J-region, distinct PCR products of the expected size were obtained (results not shown) suggesting that the difficulty observed in amplifying the Fd gene fragments was caused by some kind of strong secondary structure formation within the CH1-region that prevented the Taq polymerase from producing full size products. Therefore, the reaction involving the H-chain was optimized by testing the HC-variable primers individually in a systematic series of PCR reactions together with a well-characterized cloned cDNA copy of a H-chain (kind gift of Poul Baad Rasmussen). These studies demonstrated that the primers worked effectively, and very discriminatively, at annealing temperatures close to and above the estimated Tm of the duplex structure involved (results not shown). Therefore, subsequent PCR reactions used annealing temperatures close to the Tm of the involved primers (cf. Materials and Methods).

Jumping PCR assembly

The gene-fragments corresponding to the heavy chain of the Fab fragment (Fd) and the L-chains were paired randomly and assembled via a DNA-linker before cloning into an expression vector as a *SfiI-NotI* fragment. The DNA-linker contains a translational stop codon for the Fd gene followed by a ribosome binding site placed at an optimal distance from the start codon of the *pelB* leader sequence. Expression of the cloned Fab-cassette results in Fd and L-chains both having the *pelB* leader at their N-terminus.

Procedures for the assembly of H and L-gene fragments corresponding to a limited portion of the Fab region, namely the variable parts (Fv), via a linker fragment has been described by Clackson *et al* (10). This procedure consists of submitting the H and L-chain Fv and linker fragments to several cycles of annealing-denaturations followed by PCR amplification of the mixture in the presence of added oligo-primers complementary to the opposite ends of the Fv-fragments. This protocol, however, does not work effectively in the analogous assembly of the larger Fab fragments (results not shown) possibly because the Fab-region sequences form inter- and intra-molecular duplexes which can act as primer-templates during the initial PCR temperature cycles when oligo-primers are not present to drive the desired reaction. Therefore, a new assembly strategy was designed: Firstly, the Fd and L-chain gene fragments were amplified (the primary amplification, Figure 3A). Secondly, the LINK-D fragment was joined to the above fragments in two separate PCR reactions by submitting the fragments to a PCR protocol in the presence of primers complementary to the opposite ends of the two fragments (the LINKER assembly, Figure 3B). A similar strategy was used in the final assembly , this time using the linker-assembled Fd and L-chain fragments (Figure 3C).

As illustrated by the gel-pictures of Figure 3, lower panel, the jumping PCR procedure resulted in high yields of distinct PCR products of the expected sizes. Each of the primers used in the assembly reactions were designed with a 100% match to the constant 5′ portion of the primers used in the preceding reaction so that no biased amplification was introduced during these assembly reactions. Furthermore, to improve the cutting efficiencies of the *Sfi*I and *Not*I restriction enzymes, the primers used in the final assembly extended the assembled fragments with 22 and 25 nucleotides beyond the recognition sequences of these enzymes. Since these extensions are eliminated in the cloning step and do not appear in the vector, the final assembly primers will not be able to pick up previously cloned Fab fragments, in case they should be present as contaminants. The above methods were tested by using mRNA prepared from two well-characterized hybridoma cell lines, SAL10F68 and HUI018, as

Table 1. Superinfection titers (cfu/ml) for different phagemid/host combinations.

Phagemid	Host cell	
	TOP10F' lacI^c	*XL-1 Blue* /F'lacI^q
pFAB4	1.3×10^9	2.6×10^{10}
pFAB4/HUI018	4.4×10^9	5.0×10^{10}
pFAB5c/HUI018	5.7×10^{10}	8.0×10^{10}
pFAB5c	1.1×10^{11}	1.0×10^{11}

Cells growing exponentially at 37°C in 2×TY were infected at an OD_{600} of 0.5 with R408 helper phage and processed as described in Materials and Methods. 100 μM IPTG was used for the induction. Titers are expressed as cfu (ampicillin resistant colonies) per ml of phage supernatant. *LacI*^{q} is a promoter-up mutant producing 20 times more repressor molecules per cell than the wild-type *lacI* gene (26). The listed titer values are mean values from three separate experiments.

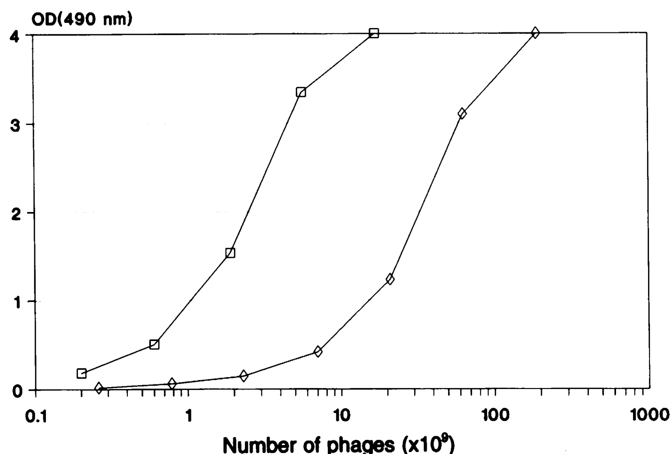

Figure 4. Functional binding characteristics of Fab-phages produced in different host cells. Phage stocks obtained after growth of pFAB5c/HUI018 in TOP10F'(squares) or XL1-Blue(diamonds) were titered (cfu/ml) and dilution series were monitored by ELISA assays using anti-M13 as detection antibody.

starting material. The purified products of the final assembly reactions were cloned into the expression vector pFAB3, a predecessor of the pFAB4 vector described in the following section, and phages were rescued from individually superinfected transformants and tested by ELISA for binding to antigen. About 25% of the clones showed specific binding with apparent binding constants (cf legend to Figure 5) in the nanomolar range as determined by inhibition ELISA assays (results not shown). Sequence analysis of some of the negative clones revealed that they had been generated from non-functional IgG-mRNAs originating from the myeloma cell-line used for the hybridoma fusion.

Design of improved expression vectors

In the vectors used here, the polycistronic unit corresponding to the Fd and the L-chain-gIIIp fusion is under transcriptional control of the wild type *lac* promoter-operator (cf Figure 2). This allows for transcriptional repression through the action of the *lac* repressor and for induction through the addition of IPTG. It appears that tight regulation of Fab expression is very important for several reasons: firstly, antibody fragments (correctly folded or not) may be toxic for *E.coli* (22). Secondly, even small amounts of gIIIp or fusion products of gIIIp have been reported to make the host cell resistant to superinfection (23). Both of the above effects will lead to very limited and/or biased antibody libraries. We have introduced two features in our expression vectors to improve their usefulness in cloning antibody fragments: The region corresponding to the N-terminal 197 amino acids was deleted as it is the N-terminal domain of gIIIp that influences the formation of sex pili which in turn is necessary for superinfection (24). This feature has been used by others in the construction of similar expression vectors (7,13,25). To provide sufficient overproduction of *lac* repressor and to make our expression vector independent of the host genotype, we introduced a functional *lacI*^{c} gene in the vector (pFAB5c, cf Figure 2). Our different constructs were tested in different host cell backgrounds and phagemid titers were monitored after superinfection with helper phage and induction with IPTG (Table 1).

The data in Table 1 show that the highest titer values (cfu/ml) were obtained with cells having the highest production of the *lac* repressor. This observation stresses the importance of keeping the Fab gene-cassette tightly repressed during growth. In evaluating the data of Table 1 it should be noted, that pFAB4 produces a truncated gIIIp fragment because the reading frame starting at the *lacZ* promoter continues through the *Sfi*I-*Not*I 'stuffer' fragment, whereas pFAB5c has several translational stop codons inserted between the *Sfi*I and *Not*I sites so that no gIIIp-like product is produced in cells harboring these phagemids. It therefore becomes possible to evaluate the individual effects on superinfectivity of the Fab-fragment and of the ΔgIIIp-fragment, particularly in the TOP10F'-strain containing the pFAB4 constructs where these effects are not masked by effective transcriptional repression. From inspection of the TOP10F'-data presented in rows 1 and 2 of Table 1 it appears that production of the HUI018 Fab-fragment has no negative effect on superinfectivity. In contrast, comparison of the TOP10F'-data in rows 1 and 2 with those of rows 3 and 4 demonstrate a more than 10-fold difference in superinfectivity. Since the Fab-fragment does not seem to influence superinfectivity, it appears that gIIIp, even in the truncated version used in our vectors, has a negative effect on superinfection titers but that this effect, however, can be eliminated by effective repression of the *lacZ* promoter activity using *lacI*^{q} strains or phagemids containing a resident *lacI*^{c} gene (Table 1, second column and rows 3 and 4). In a separate set of experiments with pFAB5c/HUI018 in TOP10F', we varied the IPTG concentration used for induction and found no increase in titers or in Fab-production (as measured by ELISA) using concentrations above 100μM IPTG (results not shown). This concentration was therefore used in subsequent experiments. As the phage display system uses affinity chromatography (panning) for the selection of antigen-binding clones, it becomes important to relate phage titers to the production of functional (antigen-binding) phages. Therefore, the binding characteristics of the Fab-phages obtained after growth in the different host cells listed in Table 1 were tested and the results expressed as ELISA OD_{490} values obtained using a fixed number of phages. These values are measures of functionality in terms of antigen recognition of the phages. As illustrated by Figure 4, pFab5c/HUI018 gave significantly higher relative ELISA-values in TOP10F' compared to XL1-Blue. pFab4/HUI018 and a selected Factor VII binding clone (cf Figure 5) gave similar and more pronounced differences, respectively, when tested in the same two strains (results not shown). The molecular basis for the superiority of TOP10F' to produce functional Fab-phages is presently unknown.

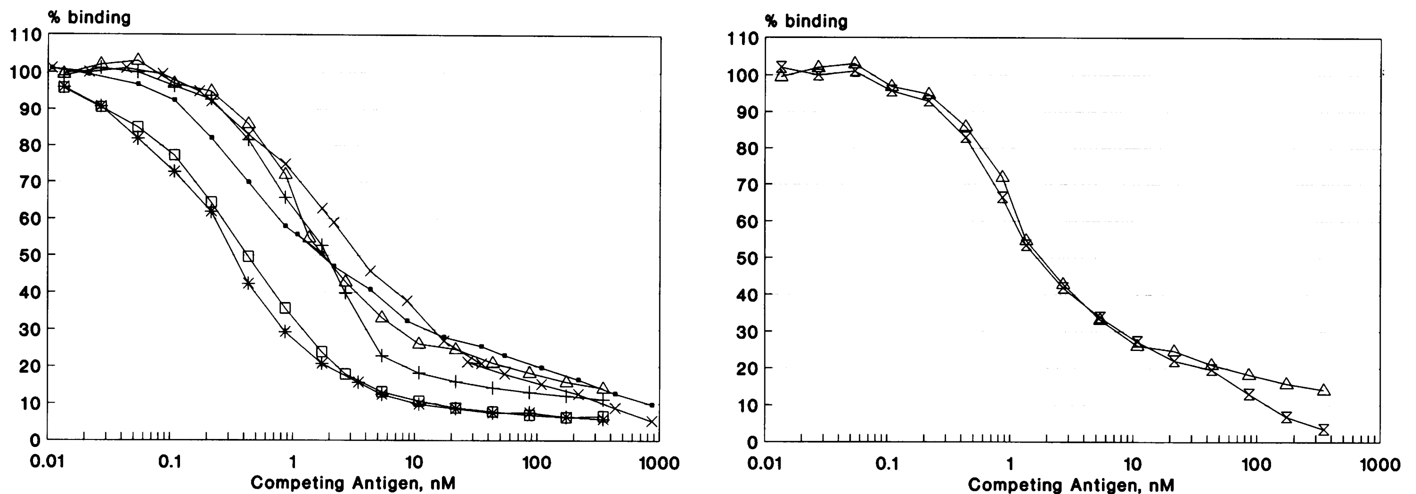

Figure 5. Binding characteristics of selected Fab-phages or free Fabs as measured by inhibition ELISA assays. **Left**: Twenty Factor VII positive Fab-phage clones were divided into seven groups according to restriction enzyme analysis. The affinity of one member from each group was tested by inhibition ELISA protocols as described in Materials and Methods. **Right**: Inhibition ELISA curves of a selected Fab-phage (one of the clones of panel A) and its cognate free Fab-antibody. Free Fab-antibodies (Fab-ΔgIIIp) were obtained from homogenates of cells transformed with the selected clone and induced with IPTG without the addition of helper phage.

Two additional features of our vector design should be noted: firstly, we have introduced a trypsin cleavage site at the junction between the L-chain and the ΔgIIIp domain to facilitate gentle and efficient elution of Fab-phages during the panning procedures in a manner that is independent of their binding strength. The phage proteins are not affected by the trypsin treatment (results not shown) and infectivity during subsequent cultivation of the eluted phage is not effected by the intactness of the Fab-antibody domain (in case it becomes destroyed by the trypsin treatment). Secondly, an *Eag*I site, which contains a subset of the *Not*I recognition sequence, was introduced downstream of the ΔgIIIp coding region (cf Figure 2). *Eag*I-digestion of a phagemid, encoding a selected Fab, thus removes the ΔgIIIp fragment. Religation of the isolated vector fragment will yield phagemids that produce soluble Fabs. These features have been used successfully in our cloning and characterization of highly antigen-specific Fab-phages from libraries made from immunized mice.

Screening libraries

To test our cloning and panning procedures we generated Fab libraries from spleen mRNA prepared from mice previously immunized with human blood-clotting factor VII. Libraries of 5×10^6 members were taken through three rounds of antigen selection (panning) before supernatants of individually superinfected clones were tested for binding to factor VII-coated microtiter plates. The panning procedure resulted in a step-wise enrichment for positive binders from 0% (of 192 tested) in the pre-selected library to more than 60% after three rounds of panning and rescue. In model panning experiments using defined mixtures of Fab phages of different binding specificities, we have demonstrated that our panning procedure enriches about 100-fold for the correct clones during each round of panning (results not shown). Twenty factor VII binding clones could be divided into seven subgroups based on restriction fragment patterns after *Pst*I and *Eco*RI digestions. However, this grouping did not correlate with binding behaviour of the individual clones as estimated from the shape and slope of the ELISA dilution curves. To determine the binding affinities of the positive clones, we performed inhibition ELISAs. Apparent affinities were determined as the reciprocal of the antigen concentration required to inhibit 50% maximal binding (cf Materials and Methods). Inhibition ELISA curves of selected clones from the seven sub-groups are depicted in Figure 5 (left) which demonstrates that all clones had high affinities to the antigen with apparent binding constants (K_a) between 5.3×10^8 and 3.4×10^9 M^{-1}. One of the factor VII clones was used to compare the binding properties of the soluble Fab molecule to that of the corresponding Fab-phage particle. As depicted in Figure 5 (right), the two ELISA inhibition curves are virtually superimposable demonstrating that the binding specificity of the Fab fragment is not changed by being displayed as a fusion protein on the surface of a phage particle.

DISCUSSION

The Fab-phage system described in this report exhibits some characteristics that are important in advancing the antibody-phage technology. First of all, we want to emphasize the importance of using an extended series of primers for the primary PCR amplifications. The initial proposal that consensus primers can be designed for the amplification of Fv and Fab regions (27) has resulted in several protocols (14,28,29) recommending the use of a limited set (4 to 8) of H and L-chain variable region primers. Comparison of these primers with the database information on immunoglobulin gene sequences reveals that only a sub-set of the relevant sequences can be matched by these primers under stringent annealing conditions. The task of designing primers to match the N-terminal portions of the VH and VL-chains cannot be solved by using highly degenerate primers: Firstly, highly degenerate primers introduce amino acid sequences not normally found there. This could influence the overall antigen binding capacity in an unpredictable manner. Secondly, and more importantly, highly degenerate primers are likely to generate very biased libraries. This is because highly degenerate primers, with sufficient specificity to serve as PCR primers in the amplification of Fab gene fragments, are unlikely to represent all variable gene families equally well. We suggest that the primer sets for the variable regions that we have designed in this study, combined with a systematic analysis of the optimal temperature conditions

for the annealing reaction in the primary PCR reactions, are minimal requirements for the successful generation of comprehensive libraries.

Our method of combining the H and L-chain Fab gene segments into a bi-cistronic transcription unit (the 'jumping PCR method') exhibits some novel features: Firstly, the method is of general applicability since we were able to obtain high yields of the final assembly product from both spleen and hybridoma material (cf Figure 3). Secondly, the splitting-up of the three-fragment assembly reaction into separate two-fragment reactions can generally be used for circumventing primer interference problems that may occur in three-fragment PCR assembly reactions. For instance, our VH-primers contain sequences that code for a portion of the *pelB* leader that is also found in front of the VL-coding region and this situation would lead to truncated PCR products in a three-fragment assembly reaction. By performing separate PCR reactions for the H and L-chains using primers tagged with unique extensions, the two sets of chains can now be combined and assembled using new sets of primers that match the introduced extensions.

It should be mentioned that either of the two chains of the Fab molecule can be used for fusion to the gIII protein. Our reason for choosing the L-chain relates to our assay conditions for detection of antigen-bound Fab-phage in which we used anti-M13. Since the H-chain alone has been reported to bind antigen to a significant extent (30), we designed the L-chain fusion to ensure that our binding assay would detect only Fab-phage which displayed both the L- and H-chains.

In our efforts to optimize each of the technical steps involved in the construction of comprehensive libraries, we also focused on superinfection (cf Table 1). The efficiency of superinfection is of crucial importance as it determines the size and complexity of the library. Our analyses demonstrated the importance of producing sufficient amounts of *lacI* to keep the Fab-gIIIp fusion product under tight expressional control. This was accomplished effectively by our pFAB5c expression vector, but the analyses further suggested that even more truncated versions of gIIIp should be tested as fusion partners for the Fab fragments.

Based on our efforts in generating comprehensive libraries we expected to isolate a variety of clones with different binding characteristics from our Factor VII library, but individual clones selected from each of the seven sub-groups, as defined by restriction enzyme analysis, all had rather similar apparent binding constants (cf Figure 5). It is possible that the panning conditions used select for binders with the fastest on-rates which might not correlate strictly with affinities (cf ref. 13), or that antibodies with similar on-rates (or affinities) were present in the mice used to generate the library. The identical binding characteristics of the free Fab molecules compared to the Fab-phages (Figure 5, right) is an important observation as it shows that our systems probably displays Fab fragments in a monovalent fashion. Two (or more) Fab fragments displayed on the tip of a phage is likely to result in co-operativity in binding to antigen and such binding behaviour will result in inhibition curves that are different from those obtained from monovalent Fab-fragments. Multi-valent display system have been reported to allow for isolation of Fab-phages with binding affinities in the micromolar range (31). It is conceivable that expression vectors with independent regulation of the wild-type gIIIp gene and the Fab-gIIIp fusion could be used to switch between mono- and multi-valent display modes. This proposal remains to be tested.

ACKNOWLEDGEMENTS

We want to thank Drs Andrew D.Griffiths, Hennie R.Hoogenboom and Ron Jackson, Cambridge, UK, and Dr Melwin Little and colleagues, DKFZ, Heidelberg for their help and advice during the initial phases of this work. Likewise, the gifts of λc2 from the Scripps Research Institute, plasmid pBR155 from Dr Poul Baad Rasmussen, hybridomas SAL10F68 from Flemming Stig Larsen and HUI018 from Per Norbert Jørgensen, Novo Nordisk, are gratefully acknowledged. This work was supported by the Danish Medical and Natural Science Research Councils, The Royal Danish School of Pharmacy and the P. Carl Petersen's Foundation.

REFERENCES

1. Smith,G.P. (1985) *Science*, **228**, 1315–1317.
2. Parmley,S.F. and Smith,G.P. (1988) *Gene*, **73**, 305–318.
3. Scott,J.K. and Smith,G.P. (1990) *Science*, **249**, 386–390.
4. Devlin,J.J., Panganiban,L.C. and Devlin,P.E. (1990) *Science*, **249**, 404–406.
5. Cwirla,S.E., Peters,E.A., Barrett,R.W. and Dover,W.J. (1990) *Proc.Natl.Acad.Sci USA*, **87**, 6378–6382.
6. Bass,S., Greene,R. and Wells,J.A.(1990) *Protein Struct.Funct.Genet.*, **8**, 309–314.
7. Lowman,H.B., Bass.S.H. and Wells,J.A. (1991) *Biochemistry*, **30**, 10832–10838.
8. Criswell,D.J. and McCafferty,J. (1992) *Tibtech*, **10**, 80–84.
9. Lerner,R.A., Kang,A.S., Bain,J.D., Burton,D.R. and Barbas,C.F. **Science** 1992. **258**:1313–1314.
10. Clackson,T., Hoogenboom,H.R., Griffiths,A.D. and Winter,G.(1991) *Nature*, **352**, 624–628.
11. Kang,A.K., Barbas,C.F., Janda,K.D., Benkovic,S.J. and Lerner,R.A. (1991) *Proc. Natl.Acad.Sci. USA*, **88**, 4363–4366.
12. Breitling,F., Dübel,S., Seehaus,T., Klewinghaus,I. and Little,M. (1991) *Gene*, **104**, 147–153.
13. Garrard,L.J., Yang,M., O'Connell,M.P., Kelly,R.F. and Henner,D.J.(1991) *BIO/TECHNOLOGY*, **9**, 1373–1377.
14. A companion to **Methods in Enzymology** (Academic Press) 1991, **2(2)**, 88–124.
15. Sambrook, J., Fritsch, E.F. and Maniatis, T.(1989) Molecular Cloning. A Laboratory Manual, 2nd. ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
16. Huse,W.D., Sastry,L., Iverson,S.A., Kang,A.S., Alting-Mees,M., Burton,D.R., Benko vic,S.J. and Lerner,R.A.(1989) *Science*, **246**, 1275–1281.
17. Hoogenboom,R.H., Griffiths,A.D., Johnson,K.S., Chriswell,D.J., Hudson,P. and Winter,G.(1991) *Nucl.Acids.Res.*, **19**, 4133–4137.
18. Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J.(1979) *Biochemistry*, **18**, 5294–5299.
20. Rath,S.,Stanley,C.M. and Steward,M.W. (1988) *J.Immunol.Methods* **106**,245–249.
21. Kabat,E.A., Wu, T.T., Reid-Miller, M., Perry,H.M. and Gottesman, K.S.(1987) In 'Sequences of Proteins of Immunological Interest', US Dept. of Health and Human Services, US Government Printing Office.
22. Pluckthun,A. and Skerra,A.(1989) *Meth.Enzymol.*, **178**, 497–515.
23. Boeke,J.D., Model,P. and Zinder,N.D.(1982) *Mol.Cell Genet.*, **186**, 185–192.
24. Crissman,J.W. and Smith,G.P.(1984) *Virology*, **132**, 445–455.
25. Barbas III,C.F.,Kang,A.S.,Lerner,R.A.,and Bencovic,S.J.(1991) *Proc.Natl.Acad.Sci USA*, **88**, 7978–7982.
26. Müller-Hill,B.(1975) *Proc.Biophys.Mol.Biol.*, **30**,227.
27. Orlandi,R., Güssow,D.H., Jones,P.T. and Winter, G.(1989) *Proc.Natl.Acad. Sci. USA*, **86**, 3833–3837.
28. Stratacyte, IMMUNOZAP^{TM} CLONING KIT. Catalog #2000.
29. Pharmacia, Recombinant Phage Antibody System. Catalog #27–9400–01
30. Ward,E.S., Gussow,D., Griffiths,A.D., Jones,P.T. and Winther,G.(1989) *Nature*,**341**,544–546.
31. Gram,H., Marconi,L-A., BarbasIII,C.F., Collet,T.A., Lerner,R.A. and Kang,A.S. (1992) *Proc.Natl.Acad.Sci.USA*, **89**, 3576–3580.