888–889 Nucleic Acids Research, 1994, Vol. 22, No. 5

Optimization of primer sequences for mouse scFv repertoire display library construction

Hong Zhou*, Robert J.Fisher and Takis S.Papas[1,+]
Laboratory of Cellular Biochemistry, PRI/DynCorp, Frederick, MD 21702 and [1]Laboratory of Molecular Oncology, NCI-FCRDC, NIH, Frederick, MD 21702, USA

Received November 17, 1993; Revised and Accepted January 28, 1994

In immunoglobulin repertoire library cloning, the homology between a particular primer sequence and that of its target template, and the diversity capacity of a primer pool are the two most important parameters which determine the cloning efficiency and the size of a resulting repertoire library. In this report we present an optimized set of primers, compatible with a commercially available vector, for mouse scFv repertoire phage display library construction.

We used the INMUNO-ZAP Fab cloning system (1) in an initial effort to clone monoclonal hybridomas for functional characterizations of the various domains of human proto-oncogene ETS1 and ETS2 by intracellular targeting. In our hands, however, the Immuno-ZAP compatible primers were only effective in PCR-cloning of the heavy chain of U-244 (an ETS2-specific MAb) (2) and the light chains of T7 (an ETS2-specific MAb) (2) and E44 (an ETS1-specific MAb) (3), but not their partner chains. This led us to try a scFv cloning system (4), which had been demonstrated effective in the construction of a human naive antibody repertoire display library (5). To develop a set of compatible primers for mouse antibody repertoire display library construction, we pooled and lined up the mouse antibody sequences collected in the Kabat-Wu data book (6). Similar sequences were grouped and a putative primer sequence was drawn from each group. All the putative primer sequences were then each compared against all the pooled sequences with the aid of a VAX computer and the best-fit primer sequences were selected. This process was repeated until all the sequences were covered. Flanking sequences compatible with the pHEN1 phagemid vector (4) and a linker (see notes for Table 1) were finally added to the V_H and V_L primers so that cloned V_H or V_L fragments can be used either in primary repertoire cloning or in affinity improvement by chain shuffling (7).

Computer-aided sequence homology analysis showed that the ten resulting MHV.BACK V_H repertoire-cloning 5′ primers cover 55% of the 147 listed mouse heavy chain sequences at 100% homology, 84% at > 95%, 97% at > 90% and 99% at > 86%. In contrast, the eight 5′ primers (H1–H8) from Immuno-ZAP system (1) only cover 58.5% of the listed mouse V_H sequences at a homology level equal or greater than 82%. The nine resulting MKV.BACK V_L kappa chain specific 5′ primers cover 67% listed sequences at 100% homology, 84% at > 95%, 91% at > 90% and 99% at > 86%. We also designed a single MLV.BACK 5′ primer for mouse lambda chain cloning, although lambda light chains constitute only less than 5% of the mouse repertoire antibodies (8). To maximize the possibility to reach the size of a naive mouse Ig repertoire, we further developed four MHV.FOR V_H cloning 3′ primers, four V_L kappa chain 3′ primers (three MKV.FOR primers derived from the V_L kappa chain FR4 3′ sequences and one MKC5.FOR primer from the C_L 5′ sequences) and one MLV.FOR V_L lambda chain 3′ primer. The MHV.FOR primers cover 94% of the listed V_H 3′ sequences at 100% match and 100% at 95% match. The three FR4 MKV.FOR primers cover 66% of the listed V_L kappa chain 3′ sequences with 100% homology, 97% at 95% match and the single MKC5.FOR primer

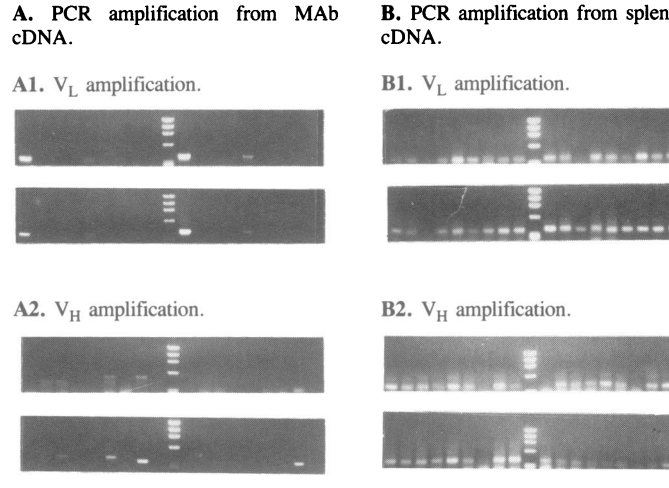

Figure 1. PCR results of the optimized mouse scFv repertoire cloning primers in cloning V_H and V_L of E44 MAb and ETS1-positive splenic Ig repertoire. **A1.** upper lanes 2–10: MKV.BACK1–9 with MKV.FOR1 and lanes 11–19: MKV.BACK1–9 with MKV.FOR2; lower lanes 2–10: MKV.BACK1–9 with MKV.FOR3 and lanes 11–19: MKV.BACK1–9 with MKC5.FOR. **A2.** upper lanes 2–10: MHV.BACK1–9 with MHV.FOR1 and lanes 11–19: MHV.BACK1–9 with MHV.FOR2; lower lanes 2–10; MHV.BACK1–9 with MHV.FOR3 and lanes 11–19: MHV.BACK1–9 with MHV.FOR4 (data from MHV.BACK10 with MHV.FORs not shown, see text). **B1** and **B2**. lane assignments are the same as those in A1 and A2, respectively. PCR reactions were performed following Perkin Elmer's GeneAmp PCR kit instruction (95°C, 1 min., 55°C 2 min., 72°C 3 min., 30 cycles).

*To whom correspondence should be addressed at: Department of Cancer Biology, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA
+Present address: Center for Molecular and Structural Biology, Medical University of South Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA

Table 1. Optimized primers for mouse scFv repertoire cloning

Primer	Oligonucleotide Sequence[1]
MHV.BACK1	5'-ttattactcgcggcccagccggccatggccGATGTGAAGCTTCAGGAGTC-3'
MHV.BACK2	5'-ttattactcgcggcccagccggccatggccCAGGTGCAGCTGAAGGAGTC-3'
MHV.BACK3	5'-ttattactcgcggcccagccggccatggccCAGGTGCAGCTGAAGCAGTC-3'
MHV.BACK4	5'-ttattactcgcggcccagccggccatggccCAGGTTACTCTGAAAGAGTC-3'
MHV.BACK5	5'-ttattactcgcggcccagccggccatggccGAGGTCCAGCTGCAACAATCT-3'
MHV.BACK6	5'-ttattactcgcggcccagccggccatggccGAGGTCCAGCTGCAGCAGTC-3'
MHV.BACK7	5'-ttattactcgcggcccagccggccatggccCAGGTCCAACTGCAGCAGCCT-3'
MHV.BACK8	5'-ttattactcgcggcccagccggccatggccGAGGTGAAGCTGGTGGAGTC-3'
MHV.BACK9	5'-ttattactcgcggcccagccggccatggccGAGGTGAAGCTGGTGGAATC-3'
MHV.BACK10	5'-ttattactcgcggcccagccggccatggccGATGTGAACTTGGAAGTGTC-3'
MHV.FOR1	5'-tgaaccgcctccaccTGCAGAGACAGTGACCAGAGT-3'
MHV.FOR2	5'-tgaaccgcctccaccTGAGGAGACTGTGAGAGTGGT-3'
MHV.FOR3	5'-tgaaccgcctccaccTGAGGAGACGGTGACTGAGGT-3'
MHV.FOR4	5'-tgaaccgcctccaccTGAGGAGACGGTGACCGTGGT-3'
MKV.BACK1	5'-tctggcggtggcggatcgGATGTTTTGATGACCCAAACT-3'
MKV.BACK2	5'-tctggcggtggcggatcgGATATTGTGATGACGCAGGCT-3'
MKV.BACK3	5'-tctggcggtggcggatcgGATATTGTGATAACCCAG-3'
MKV.BACK4	5'-tctggcggtggcggatcgGACATTGTGCTGACCCAATCT-3'
MKV.BACK5	5'-tctggcggtggcggatcgGACATTGTGATGACCCAGTCT-3'
MKV.BACK6	5'-tctggcggtggcggatcgGATATTGTGCTAACTCAGTCT-3'
MKV.BACK7	5'-tctggcggtggcggatcgGATATCCAGATGACACAGACT-3
MKV.BACK8	5'-tctggcggtggcggatcgGACATCCAGCTGACTCAGTCT-3'
MKV.BACK9	5'-tctggcggtggcggatcgCAAATTGTTCTCACCCAGTCT-3'
MKV.FOR1	5'-atgagttttgtctgcggccgcCCGTTTCAGCTCCAGCTTG-3'
MKV.FOR2	5'-atgagttttgtctgcggccgcCCGTTTTATTTCCAGCTTGGT-3'
MKV.FOR3	5'-atgagttttgtctgcggccgcCCGTTTTATTTCCAACTTTG-3'
MKC5.FOR	5'-atgagttttgtctgcggccgcGGATACAGTTGGTGCAGCATC-3'
MLV.BACK	5'-tctggcggtggcggatcgCAGGCTGTTGTGACTCAGGAA-3'
MLV.FOR	5'-atgagttttgtctgcggccgcCTTGGGCTGACCTAGGACAGT-3'
LINK.AMP5[2]	5'-GTCTCCTCAGGTGGAGGC-3'
LINK.AMP3	5'-CGATCCGCCACCGCCAGA-3'
LINK.AMP3T	5'-CGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCTGA-GGAGAC-3'

1. The 5' lower case letters represent sequences derived from pHEN1 vector and the 3' upper case letters represent sequences optimized from Kabat-Wu data, which were used for homology analyses. *The underlined sequences are the cloning sites (SfiI for 5' and NotI for 3' termini, respectively). **2.** This pair of primers PCR-amplifies from pαTEL clones (6) a $(Gly_4Ser)_3$ linker sequence (5'-GGTGG-AGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCG-3'), different from the one cited (6, 9), as revealed by sequencing analyses. LINK.AMP5 may also be used with LINK.AMP3T to generate a *ds* linker fragment for V_H and V_L PCR assembly.

for the C_L 5' sequence covers all the available kappa light chains at 100% match. The other single 3' primer MLV.FOR covers all the available mouse lambda chain sequences at 95% or higher homology.

To evaluate their cloning efficacy, which reflects the homology between each redesigned primer with those clones in a mouse immunoglobulin repertoire, we tested them with polyA$^+$ RNA from the MAb E44. All the 'FOR' primers were used individually to generate first strand cDNAs from E44 polyA RNA by reverse transcription. The resulting first strand cDNA were subsequently used for PCR amplifications using all combinations of 'FOR' and 'BACK' primers. The results with the V_L primers showed that MKV.BACK1 had the best match with E44 kappa light chain sequence, MKV.BACK5 to a lesser extent (Fig. A1), and all the 3' primers (MKV.FOR1, MKV.FOR2, MKV.FOR3 and MKC5.FOR) were effective, indicating that MKC5.FOR may be able to serve as a universal 'FOR' primer. In E44 V_H cloning, in combination with MHV.FOR1, MHV.BACK6 or MHV.BACK8 gave products of the expected size and those of MHV.BACK2 or MHV.BACK3 gave products of a smaller size; while in combination with MHV.FOR3, MHV.BACK3 and MHV.BACK6 gave products of the right size but MHV.BACK8 did not (Fig. A2). This implies that either the V_H primers may not 100% match with that of the E44 heavy chain sequence and therefore cover a larger Ig repertoire, or the smaller band represents a shortened cDNA by-product from reverse transcription.

To test their diversity capacity in polyclonal antibody repertoire cloning, female Balb/c mice were immunized and boosted once with HPLC purified recombinant human ETS1. Their tail-bleed were assayed by ELISA. Those mice with positive response were sacrificed and their spleens were taken for polyA RNA extraction, which were further subject to first strand cDNA syntheses and PCR amplifications. The results showed that all the MHV and MKV 'BACK' and 'FOR' primers were effective except MHV.BACK7 and MHV.BACK10 (Fig. B1 and B2). MHV.BACK10 was designed at 100% match to cover three hard-to-group sequences in the Kabat-Wu data (ref. 6, sequences #242, #243 and #244). Since the polyA RNA was extracted from ETS1 immunized mice after a single boost, it is unclear if the percentage of the B cells secreting immunoglobulins coded by these sequences was too low to be cloned in this ETS1-stimulated B cell repertoire or these sequences were resulted from somatic mutations during affinity maturation. Same explanation could be applied to the result with primer MHV.BACK7.

Taken together, these PCR amplification results and sequence homology analysis data demonstrate that the majority of the primers can be used for effective mouse V_H and V_L repertoire cloning. The MHV.BACK7 and MHV.BACK10 primers may be proven useful in some rare cases and in naive Ig repertoire cloning where diversity becomes a major concern. To make the optimized primers generally applicable, we decoded a $(Gly_4Ser)_3$ linker by sequencing the pHEN1-derived pαTEL9, pαTEL13, pαTEL14 and pαTEL16 constructs (5) and developed the linker amplification primers (Table 1). Thus, this work completes a set of optimized primers for mouse scFv repertoire display library construction, either using the pHEN1 vector or its recent commercial derivative, pCANTAB5 (Pharmacia).

ACKNOWLEDGMENTS

We thank the Advanced Scientific Computing Laboratory at FCRDC for assistance in sequence analyses and Dr Greg Winter for pαTEL clones. This work was supported by NCI, DHHS, under Contract No. NO1-CO-74102 with PRI/DynCorp, and NIH grant CP05667.

REFERENCES

1. Huse,W.D., Sastry,L., Iverson,S.A., Kang,A.S., Alting-Mees,M., Burton,D.R., Benkovic,S.J. and Lerner,R.A. (1989) *Science* **246**, 1275–1281.
2. Fugiwara,S., Koizumi,S., Fisher,R.J., Baht,N.K. and Papas,T.S. (1990) *Hybridoma* **9**, 559–571.
3. Koizumi,S., Fisher,R.J., Fujiwara,S., Jorcyk,C.L., Baht,N.K., Seth,A. and Papas,T.S. (1990) *Oncogene* **5**, 675–681.
4. Hoogenboom,H.R., Griffiths,A., Johnson,K.S., Chiswell,D.J., Hudson,P. and Winter,G. (1991) *Nucleic Acids Res.* **19**, 4133–4137.
5. Marks,J.D., Hoogenboom,H.R., Bonnert,T.P., McCafferty,J., Griffiths,A.D. and Winter,G. (1991) *J. Mol. Biol.* **222**, 581–597.
6. Kabat,E.A., Wu,T.T., Reid-Miller,M., Perry,H.M. and Gottesman,K.S. (1987) *Sequences of Proteins of Immunological Interest*, 4th ed. US Department of Health and Human Services, Public Service, NIH. pp. 508–553.
7. Marks,J.D., Griffiths,A.D., Malmqvist,M., Clarkson,T.P., Bye,J.M. and Winter,G. (1992) *BIO/TECHNOLOGY* **10**, 779–783.
8. Day,E.D. (1990) In, *Advanced Immunochemistry*, second edn, Wiley-Liss, Inc. pp. 20.
9. Huston,J.S., Levinson,D., Mudgett-Hunter,M., Tai,M.S., Novotny,J., Margolies,M.N., Ridge,R.J., Bruccoleri,R.E., Haber,E., Crea,R. and Oppertnann,H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5879–5883.