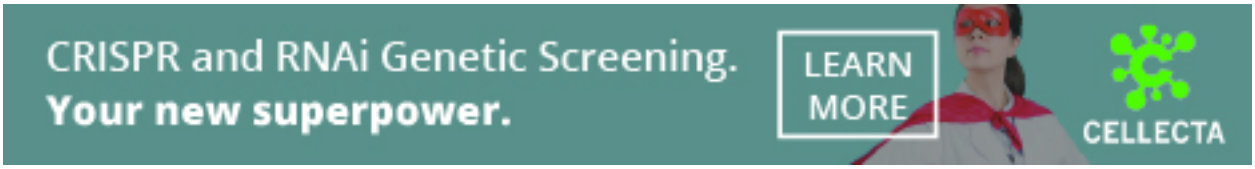
Direct PCR sequencing of murine immunoglobulin genes using E. coli single-stranded DNA-binding protein.

R Rapley, S Flora and M R Walker

Genome Res. 1992 2: 99-101
Access the most recent version at doi:10.1101/gr.2.1.99



Direct PCR Sequencing of Murine Immunoglobulin Genes Using *E. coli* Single-stranded DNA-binding Protein

Ralph Rapley, Singh Flora, and Matthew R. Walker

Molecular Biology Research Group,
Wolfson Research Laboratories,
Birmingham University Medical School,
Birmingham B15 2TH, UK

The polymerase chain reaction (PCR) has become a widely established method for producing large amounts of DNA from limited initial template sources.[1] When combined with direct nucleotide sequencing of PCR-amplified products, this becomes a powerful procedure that allows sequence information to be obtained within a very short timescale. Direct dideoxy sequencing of PCR products is, however, far from straightforward. Many of the problems associated with the sequencing of PCR products appear to be due to the rapid reannealing of the short linear DNA strands following heat denaturation. This reannealing may prevent the extension of a sequencing primer and may result in multiple termination or four tracking, where the two strands begin to reassociate. Many methods and technical modifications have been developed to overcome this problem, including the use of denaturants such as dimethyl sulfoxide (DMSO),[2] inclusion of detergents such as nonidet P-40,[3] and using *Taq* polymerase sequencing at elevated temperatures.[4] During our study of rearranged murine immunoglobulin genes, we wished to analyze heavy- and light-chain variable regions from a number of hybridoma cell lines by PCR amplification and direct sequencing. We found that in some cases sequencing these PCR products directly gave terminations that were either (1) unable to be determined unambiguously, (2) weak in signal intensity (despite long autoradiographic exposures), or (3) only shorter extended fragments were preferentially terminated (up to 40–60 bp from the sequencing primer). Therefore, we investigated the use of *E. coli* single-stranded binding protein in an attempt to prevent the rapid reannealing of PCR strands following duplex denaturation, and to allow uninterrupted extension and specific termination over the increased lengths of the PCR fragment.

AMPLIFICATION STRATEGY

A number of murine hybridoma cell lines secreting IgG1 monoclonal antibodies to a synthetic peptide representing the 1–34 amino acid region of parathyroid hormone-related protein (PTHrP) were used in this study.[5] Initially the mRNA encoding the heavy-chain variable region from one of these hybridomas (cell line E2) was isolated and subjected to first-strand cDNA synthesis with reverse transcriptase (United States Biochemicals, Cambridge, UK) under conditions described in ref. 6. PCR was then carried out using a degenerate reverse primer D750 (/ indicates either base at that position) (5′-AGGTC/GA/CAA/GCTGCAGC/GAGTCA/TGG-3′) and a forward primer D751 (5′-GGCCAGTG-GATAGAC-3′) for the times and temperatures described in ref. 6. These primers annealed to heavy-chain framework region 1 and constant region 1 (26 bp from the D–J junction), respectively, and amplified a predominant 400-bp fragment.

PURIFICATION AND SEQUENCING OF PCR PRODUCTS

PCR products were separated through 8% polyacrylamide and excised from the gel with a scalpel. The bands were eluted in 1 ml of 0.1×) TE buffer (pH 8.0) at 37°C overnight. The acrylamide was removed and the eluate concentrated by centrifugation at 3000×*g* in a Centricon 10 microconcentrator (Amicon, Basingstoke, UK) for 15 min. The concentrated sample was freeze-dried in a Univapo 100 freeze drier (Jones Chromatography, Cardiff, UK) and resuspended in 12 μl of sterile distilled water.

In the annealing reaction, 2 μl of the concentrated fragment, 1 μl of 6× reaction buffer (240 mM Tris-HCl, pH 7.5, 150 mM $MgCl_2$, 300 mM NaCl, 60% DMSO), 2 μl of sequencing primer (10 μM) (either of the PCR primers or an internal primer), and 1 μl of sterile distilled water were added together with 0.5 μl of *E. coli* single-stranded binding protein (2 μg/μl) (United States Biochemicals, Cambridge, UK). Next, 1 μl of 100 mM dithiothreitol was added, and the reaction mixture was heated to 75°C for 3 min and immediately snap-frozen in liquid nitrogen.

The labeling reaction consisted of adding 1 μl of [^{35}S]dATPαS (250 μCi) (Amersham International, Amersham, UK), 2 μl (2 units) of Sequenase V2.0 (United States Biochemicals, Cambridge, UK) directly to the annealed reaction mix. Aliquots (2 μl) of this labeling mix were terminated in the termination mixes supplied with Sequenase II kit (United States Biochemicals, Cambridge, UK) made 10% for DMSO. The reaction mix was incubated for 5 min at 20°C and a further 5 min at 37°C. Subsequently, 2 μl of chase mix (0.25 mM dNTPs, 50 mM

2:99–101©1992 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/92 $3.00

NaCl, and 10% DMSO) was added to each of the four termination reactions and the incubation continued for a further 15 min. Stop buffer (4 μl) was added to complete the sequencing reaction (supplied with Sequenase II kit). Prior to loading the terminated reactions on a denaturing polyacrylamide gel, the *E. coli* single-stranded binding protein was digested by incubation with Proteinase K (0.1 μg) for 20 min at 60°C. The result of direct PCR sequencing of the rearranged immunoglobulin heavy-chain variable region (cell line E2) gene in the presence or absence of *E. coli* single-stranded binding protein is represented in Figure 1.

DISCUSSION

The inclusion of *E. coli* single-stranded binding protein in the direct sequencing of rearranged immunoglobulin gene PCR products has allowed a further seven heavy-chain variable regions and four light chains to be analyzed (in prep.). The signals obtained were more intense and in general produced extension and specific terminations further from the primer (up to 200 bp in some cases) than those observed in the absence of *E. coli* single-stranded binding protein using identical template preparations. The use of other helix destabilizing proteins, such as the gene *32* protein from bacteriophage T4, have been used to improve the sequencing of denatured plasmid DNA template.[7] The mechanism of action of *E. coli* single-stranded binding protein is probably similar to that of gene *32* protein, although the renaturation of circular plasmid DNA is slower than that of the short linear PCR fragments such as those used in this study (400 bp). It is also probable that the binding of *E. coli* single-stranded binding protein to linear denatured PCR strands may also prevent or reduce the formation of any significant secondary structure, a problem known to cause polymerase pausing and premature termination. This particular protocol has also been applied with similar benefits to sequencing PCR products generated from the G:C-rich human apolipoprotein E gene (250 bp),[8] human immunoglobulin heavy-chain variable regions (700 bp), and a major grass pollen allergen (700 bp) (in prep.). The use of *E. coli* single-stranded binding protein seems to be generally applicable in improving the sequencing of a wide range of PCR products amplified from various sources. Presently, we are investigating other methods of using helix destabilization

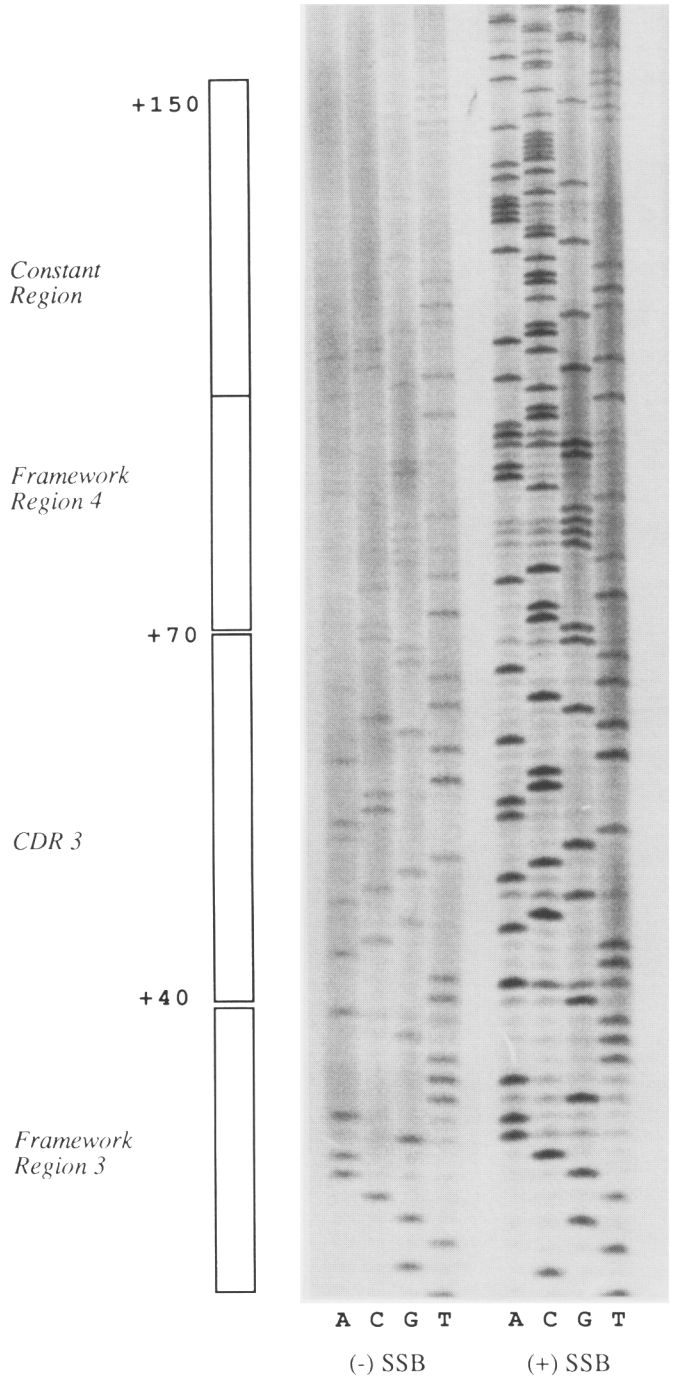


FIGURE 1. Identical murine immunoglobulin heavy chain-variable regions amplified by PCR, directly sequenced in the absence of *E. coli* single-stranded binding protein (−) and in the presence of *E. coli* single-stranded binding protein (+). The termination reactions were separated on 20% polyacrylamide 6 M urea denaturing sequencing gel run for 3 hr at 12,500 V, fixed in 10% methanol/10% acetic acid (45 min), and subsequently dried onto Whatman 3M paper. Autoradiography was carried out overnight using Hyperfilm-MP (Amersham International plc, UK). The sequencing primer for this reaction was an internal primer specific for the VH2 gene family, annealing within framework region 3 (5′-ACATGCAGCTCAGCAGCCTGACATCT-3′). Figures denote distance from the sequencing primer. Also indicated are the main elements that may be defined within the immunoglobulin heavy-chain variable region, including complementarity determining region 3 (CDR3) and the flanking framework regions.

in an attempt to enhance further the length of extension from the sequencing primer and to reduce signal background.

REFERENCES

1. Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **239:** 487–491.
2. Winship, P.R. 1989. An improved method for direct sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.* **17:** 1266.
3. Bachmann, B., W. Lucke, and G. Hunsmann. 1990. Improvement of PCR amplified DNA sequencing with the aid of detergents. *Nucleic Acids Res.* **18:** 1309.
4. Innis, M.A., K.B. Myambo, D.H. Gelfield, and M.A.D. Brown. 1988. DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci.* **85:** 9436–9440.
5. Ratcliffe, W.A., S. Hughes, M.G. Gilligan, D.A. Heath, and J.G. Ratcliffe. 1990. Production and characterization of monoclonal antibodies to parathyroid hormone-related protein. *J. Immunol. Meth.* **122:** 109–116.
6. Huse, W.D., L. Sastry, S.A. Iverson, A. Kang, M. Alting-Mees, D.R. Burton, S.J. Benkovic, and R.A. Lerner. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246:** 1275–1281.
7. Kaspar, P., S. Zadrazil, and M. Fabry. 1989. An improved double-stranded DNA sequencing method using gene 32 protein. *Nucleic Acids Res.* **17:** 3616.
8. Breslow, J.L. 1988. Apolipoprotein genetic variation and human disease. *Physiol. Rev.* **68:** 85–132.

Received October 28, 1991; accepted in revised form April 29, 1992.