An efficient synthetic primer for the M13 cloning dideoxy sequencing system

Peter C. K. LAU and John H. SPENCER

Department of Biochemistry, Queen's University at Kingston, Ontario, Canada K7L 3N6

(Received 28 July 1982)

The deoxytetradecamer d(AAAACGACGGCCAG) has been shown to be an excellent universal primer for sequence determination of DNA cloned into the bacteriophage M13 mp7, mp8, and mp9 series. This new primer offers several advantages over others currently available and it has been used to define the cloning of Hinf I fragments of bacteriophage SI3 DNA into the Eco RI site of M13 mp7, utilizing the homologous complementary base pairing of the two restriction sites. Of the four possible sequence derivatives of the Hinf I GANTC recognition site~ only those corresponding to GAATC and GATTC have been found at cloning sites in chimeras.

Cloning of DNA fragments into the single-stranded DNA bacteriophage M13 lac system (Gronenborn & Messing, 1978; Messing et al., 1981) has greatly facilitated the determination of nucleotide sequences (Sanger et al., 1980; Anderson et al., 1981; Gardner et al., 1981; Fields & Winter, 1981) by the dideoxy sequencing method (Sanger et al., 1977). Theoretically only one primer is required for the multiple cloned templates (Schreier & Cortese, 1979) if it is complementary to a sequence flanking the cloning sites. Several 'universal' primers have been reported (Table 1). Two primers are present in chimeric plasmids, pHM 232 (Heidecker et al., 1980) and pSP 14 (Anderson et al., 1980), and their growth, isolation, and use are elegant, but they have the disadvantage of being double-stranded. They also have the inherent tendency to self-prime, resulting in artifactual bands in sequencing gels unless additional experimental manipulations are carried out (Sanger et al., 1980; Anderson et al., 1980). Other primers reported are synthetic oligonucleotides. Four of them (nos. 5, 7, 8, and 9, Table 1) are complementary to a sequence 21–38 bases 3’ to the Eco RI site, at position 6272 of M13 mp7. Two primers (nos. 3 and 4, Table 1) are complementary to a sequence adjacent to the same Eco RI site. We present data on a shorter efficient primer with exact base complementarity to a unique site on the vector which requires only simple annealing conditions and provides sequence information through the multipurpose cloning sites of M13 mp7 and derivatives (Messing, 1981) without inclusion of any superfluous vector sequence information. The synthetic tetradecamer has been used to demonstrate the cloning of Hinf I DNA fragments (GANTC) into the Eco RI (GAATTC) site of M13 mp7.

© 1982 The Biochemical Society
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Length</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AGTCGACCGGATCCGGGAAATTCACTGGCCGTCTGTTTACAAACGTCGTGACTGGAAA---(M13mp7)---3'</td>
<td>(96 bp) --- 5'</td>
<td>(1)</td>
</tr>
<tr>
<td>2</td>
<td>T TAAGTGACCGCGACAAATGTGGCAGCATGAGCT---(26 bp)---5'</td>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td>3</td>
<td>GTGACCGGCGACAAATGTGGCAGCTGGCAGCTAG-5'</td>
<td>(19-mer)</td>
<td>(3)</td>
</tr>
<tr>
<td>4</td>
<td>TGACCGGCGACAAATGTT-5'</td>
<td>(17-mer)</td>
<td>(4)</td>
</tr>
<tr>
<td>5</td>
<td>GTTGCAGCAGCTGACCT-5'</td>
<td>(17-mer)</td>
<td>(5)</td>
</tr>
<tr>
<td>6</td>
<td>GACGGCGACACAA-5'</td>
<td>(14-mer)</td>
<td>(6)</td>
</tr>
<tr>
<td>7</td>
<td>T TGCACTGACCTGGC -5'</td>
<td>(15-mer)</td>
<td>(7)</td>
</tr>
<tr>
<td>8</td>
<td>TGCAGCAGCTGACCT-5'</td>
<td>(15-mer)</td>
<td>(8)</td>
</tr>
<tr>
<td>9</td>
<td>TGTTGACGAGCT-5'</td>
<td>(12-mer)</td>
<td>(9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Length</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GTGGACCTGACAGCGCCAGTTGGCACTGGCCGTCTGTTTACAAACGTCGTGACTGGAAA---(M13mp8)---3'</td>
<td>(96 bp) --- 5'</td>
<td>(10)</td>
</tr>
<tr>
<td>2</td>
<td>HindII</td>
<td>PstI</td>
<td>HindIII</td>
</tr>
<tr>
<td>3</td>
<td>GTGCGACCGGATCCGGGAAATTCACTGGCCGTCTGTTTACAAACGTCGTGACTGGAAA---(M13mp9)---3'</td>
<td>(96 bp) --- 5'</td>
<td>(10)</td>
</tr>
</tbody>
</table>

1. Messing et al., 1981
2. Heidecker et al., 1980; BRL
3. Anderson et al., 1980; BRL
4. Wu et al., 1980
5. Duckworth et al., 1981
6. This study
7. Gardner et al., 1981; P.L.
9. Collaborative Research
Materials and Methods

Materials

Bacteriophage M13 mp7 and Escherichia coli JM103 were obtained from J. Messing (Messing et al., 1981) and Bethesda Research Laboratories (BRL). Bacteriophage S13 w.t. was obtained from I. Tessman. \( [\alpha^{-32P}] \text{-dATP (specific activity 600 Ci/mmol)}, \]
\( [\gamma^{-32P}] \text{-ATP (specific activity 3000 Ci/mmol)}, \]
and polynucleotide kinase were from New England Nuclear Canada. Restriction enzymes were from BRL or Boehringer Mannheim Canada (BMC) and used according to manufacturers' conditions. Dithiothreitol (DTT), DNA polymerase I (Klenow fragment) and dNTPs were obtained from BMC and dideoxynucleoside triphosphates (ddNTPs) from PL Biochemicals. Isopropyl-thio-\( \beta \)-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-\( \beta \)-D-galactoside (X-gal) were obtained from Sigma, BRL, or BMC.

Methods

Cloning of S13 DNA fragments into M13 mp7. General methods for the preparation, isolation, and purification of M13 mp7 RFI DNA were obtained from J. Messing prior to publication (Messing et al., 1981). The vector RFI DNA was linearized by digestion with Eco RI or Hinc II. Bacteriophage S13 RFI DNA was prepared by a modification of the procedure of Kaptein and Spencer (1978) (Lau & Spencer, manuscript in preparation). The S13 DNA was digested with Hae III or Hinf I, the mixture deproteinized with phenol and extracted with ether and the DNA precipitated with 2.5 vol. of ethanol. The DNA was then ligated into the Hinf II or Eco RI linearized vector essentially as described by Winter et al. (1981). The ligation mixture was used to transform competent E. coli JM103 cells which were spread on indicator plates containing IPTG and X-gal. Colorless plaques (clones) were picked (Barnes, 1977) and inoculated into a 1.5-ml JM103 culture. Twenty \( \mu \)l of the culture supernatant, in the presence of sodium dodecyl sulfate (SDS), EDTA, and Bromophenol Blue dye were electrophoresed on a 0.7% agarose gel to screen for size of the DNA inserts. The mobility of the DNA was compared to single-stranded M13 mp7 wild-type DNA.

The synthetic primer and sequencing reactions. Single-stranded DNA preparations from chimeras were isolated from the supernatants of 1-ml cultures (Sanger et al., 1980) except that the DNA was dissolved in 20 \( \mu \)l of 10 mM Tris/HCl, pH 7.4, 0.1 mM EDTA. The preparations were screened, using the ddTTP tracking procedure exactly according to Sanger et al. (1980), but using the newly designed tetradecamer primer (0.1 pmol/reaction). The primer was synthesized with a DNA/RNA synthesizer (Bio Logicals) which uses the phosphite triester method (Alvarado-Urbina et al., 1981) and kindly supplied by Bio Logicals.

Sequencing, using the four ddNTPs, was carried out as follows: for each of the four reactions 5 \( \mu \)l of single-stranded M13 chimeric (template) DNA solution was mixed with 1 \( \mu \)l of tetradecamer primer (0.5 pmol), 1 \( \mu \)l of 10XH buffer (66 mM Tris/HCl, pH 7.4, 66 mM MgCl\(_2\), 500 mM NaCl, 10 mM DTT; Air et al., 1976) and 3 \( \mu \)l of
distilled water. The mixture was heated at 67°C for 10 min and cooled to room temperature. Two µl of the template/primer mixture was then combined with 2 µl of a mixture containing the appropriate dNTP, ddNTP, and [α-32P]-dATP, and then 1 unit of DNA polymerase I (Klenow fragment) in 1 µl was added. The final nucleotide concentrations were:

A reaction: dA 1 µM  
T reaction: dA 1 µM  
G reaction: dA 1 µM  
C reaction: dA 1 µM

dT 26 µM  
G 26 µM  
C 26 µM  
T 26 µM  
G 38 µM  
C 38 µM  
G 38 µM  
C 38 µM  
T 112 µM

dT 38 µM  
G 38 µM  
C 2 µM  
G 72 µM  
C 2 µM  
G 72 µM  
C 2 µM  
T 90 µM

The mixtures were incubated for 15 min at room temperature, and then 1 µl of a solution containing all four dNTPs, each at a concentration of 45 µM, was added (Anderson et al., 1980) and the mixture was incubated for a further 15 min. The reactions were stopped by the addition of 6 µl of 98% deionized formamide containing 0.02% xylene cyanol FF and 10 mM EDTA. Bromophenol Blue was omitted from the dye mix because it often interferes with the interpretation of sequence information on the autoradiogram corresponding to the dye's migration position.

The sequencing reaction mixtures were separated on 0.4-mm sequencing gels of 6%, 8%, or 12% polyacrylamide as described by Sanger and Coulson (1978) and Smith (1980). Electrophoresis was at 1.5 kV for 40-cm gels and 2.5-3 kV for 80-cm gels (BRL Models S0 and S1).

In some experiments the tetradecamer primer was phosphorylated by polynucleotide kinase and [γ-32P]-ATP according to procedure 5B of Maxam and Gilbert (1980). The 5' 32P-labelled primer, approximately 4.0 x 106 d.p.m. in a 2-µl vol., was annealed to the template as described above. The nucleotide mixes were as listed above, except that dATP was supplemented to 0.8 µM for the A reaction and 80 µM for the T, G, and C reactions. The reaction mixes were incubated for 30 min at room temperature without a further 'chase' period.

Following electrophoresis the gels were covered with Saran Wrap (DOW) plastic film and autoradiographed overnight at -70°C, using either Kodak X-Omat AR-5 or Du Pont Cronex 4 X-ray films.

Results

Fig. 1a is an autoradiograph of a sequence determination of the complementary strand of a 72-base-pair (bp) Hae III fragment of SI3 DNA cloned into the Hinc II site of M13 mp7 primed by the synthetic tetradecamer. The M13 mp7 vector sequences flanking the SI3 Hae III 72-bp insert are underlined. When dATP is the radiolabeled nucleotide in the reaction mix, the first base labeled and terminated is the A complementary to the T residue at position 6275 of the M13 mp7 vector sequence (Table 1). Thus the sequence of the cloned fragment is preceded by a 'lead-in' sequence of 16 residues, viz.

5'-ATTCCCCGGATCCGTC-3'  
EcoR I  Bam HI  Hinc II
When the same S13 Hae III fragment chimera was primed by the 26-bp primer from plasmid pSP 14y, the first unambiguously legible base on the sequencing gel (Fig. 1b) is the G complementary to the penultimate C residue (position 6267) of the Bam HI site. Although the start of the sequence of the cloned fragment is still defined, the sequence around the Eco RI recognition site is masked by four bands - two of which are very dark - across all four sequencing tracks, which obliterate sequence information of eight bases (see also Anderson et al., 1980).

The feasibility of cloning an Hinf I DNA fragment into the Eco RI site of M13 mp7 is illustrated in Fig. 2a. The sequence and thus strategic positioning of the tetradecamer primer to the M13 mp7 cloning sites provides sequence information across the Eco RI site. In this case the 'lead-in' sequence is 5'-ATC-3'. Note that this is not 5'-ATTC- (cf. Fig. 1a). Fig. 2b shows the sequence 5'-GATTC at the cloning site on the 5' end of an S13 Hinf I insert resulting from ligating the Hinf I fragment into the Eco RI site. Note that the sequence corresponds to Hinf I recognition sequence and not Eco RI. Fig. 2b also shows an internal Hinf I recognition sequence (GAGTC) which resulted from the ligation of two Hinf I fragments before cloning. If the tetradecanucleotide primer is 5' 32P-labeled it is possible to define completely the recognition sequence 5'-GAATC at the cloning site on the 3' end of the insert (Fig. 3a,b).

Discussion

The synthetic tetradecamer 5'-AAAACGACGGCCAG has several advantages over other currently available primers (Table 1). A denaturation step of the template and primer at 100°C prior to annealing is not required and self-priming does not occur, reducing considerably experimental manipulation and facilitating the screening of clones by the ddTTP-tracking procedure (Sanger et al., 1980; Fields et al., 1981). The heptadecamer (primer 5) has been reported to have an additional binding site on the M13 template resulting in multiple priming sites (Duckworth et al., 1981). Primers 7, 8, and 9 are complementary to the same sequence region and may have this disadvantage. Nevertheless, the pentadecanucleotide, primer 7, has been used to obtain the complete sequence of the cauliflower mosaic virus genome (Gardner et al., 1981). Primers 5, 7, 8, and 9 are complementary to a sequence some twenty bases away from the first cloning site, resulting in superfluous vector sequence information in sequence analyses. Thus the tetradecamer and primers 3 and 4 provide some 20 additional nucleotides of useful sequence information from each sequencing gel. With the tetradecamer primer, a combination of 6% and 8% gels can provide unambiguous sequence data exceeding 350 nucleotides.

The unique sequence of the tetradecamer also provides sequence information through the multipurpose cloning vector restriction sites, especially the first cloning site, in M13 mp7 when [α-32P]-dATP is the radiolabeled nucleotide. When [α-32P]-dCTP is the radiolabel the Eco RI site is not defined, the first base observed being the C, corresponding to position 6271 of the M13 mp7 vector (Table 1). This aspect is important in comparative sequencing and for exact sequence definition of cloned fragments.
Fig. 1. Autoradiograms of sequencing gels of an S13 Hae III 72-bp fragment cloned into the Hinc II site of M13 mp7. (a) The template was primed using the synthetic tetradecamer and the products were separated on an 80-cm 12% sequencing gel. (b) The template was primed using the 26-bp double-stranded primer and analyzed on an 80-cm 6% sequencing gel. The vector sequences flanking the 72-bp Hae III fragment insert are underlined. The dashed lines indicate the relative positions of the two sequences in the two gels. The nucleotide length (n) is given on each autoradiograph. The heavy bands below 40 n in (b) represent the artifact bands due to self-priming discussed in the text. XC and BB indicate the positions of the xylene cyanol FF and Bromophenol Blue dyes.

Fig. 2. Autoradiograms of sequencing gels of S13 Hinf I DNA fragments cloned into the Eco RI site of M13 mp7. The templates were primed with the tetradecamer. (a) Portion of an 8% sequencing gel. The sequence complementary to the viral vector DNA at the 3' end of the insert is marked by a vertical bar. (b) Portion of a 6% sequencing gel. The sequence complementary to the viral vector DNA on the 5' end of the insert is underlined. Note the Hinf I 3'-CTTAG recognition sequence at the 5' cloning site. An internal Hinf I sequence (CAGTC) is observed (see boxed sequence). The nucleotide lengths (n) are indicated.
Fig. 3. Autoradiograms of sequencing gels of an S13 Hinf I DNA fragment cloned into the Eco RI site of M13 mp7. A 5' 32P-labeled tetradecamer primer was used. (a) Separation on a 40-cm 8% sequencing gel. (b) Separation on an 80-cm 8% gel. The Hinf I recognition sequence 5'-GAATC (boxed) is observed at the cloning site. The nucleotide lengths (n) are numbered. XC and BB indicate the positions of the xylene cyanol and Bromophenol Blue dyes.

The feasibility of cloning Hinf I fragments into the Eco RI site of an M13 vector was first observed using guinea-pig satellite DNA (A. Krebs & J. H. Spencer, unpublished observations). This result has been confirmed by sequencing known Hinf I restriction fragments of S13 DNA cloned into the Eco RI sites. Analyses of the two S13 Hinf I clones presented in this paper show the sequence at the cloning site is Hinf I, not Eco RI. This may arise by nucleolytic removal of the terminal A residue of the 5' end of the Eco RI recognition sequence on the vector DNA leaving 5'-ATT- as the overlapping end instead of 5'-AATT-. Alternatively, the terminal A residues of the Eco RI cohesive ends may loop out, and then undergo in vivo repair after transformation. Accordingly, only the Hinf I fragments with the GAATTC recognition sequence can be ligated to the template. This observation increases the versatility of the Eco RI cloning site of the M13 mp7 vector (Messing et al., 1981).

The tetradecamer universal primer is also excellent for the M13 mp8 and mp9 derivatives (unpublished results) and should be useful for rapid sequencing of DNA cloned in plasmids of the pUC series (Messing, 1981) and pUR 222 (Rüther et al., 1981) according to the procedures of Smith et al. (1979) and Wallace et al. (1981).

Acknowledgements

We thank Dr. J. Messing for strains and information prior to publication, Dr. S. Anderson for plasmid pSP 14 and advice, and Dr. R. Deeley for a gift of the pSP 14 RF DNA and helpful discussion. We are grateful to Bio Logicals (Canada) for the generous gift of the tetradecamer primer. We also thank R. Shaver and D. Powers for skillful technical assistance and H. Metz for photography. This work was supported by grant MT 1453 from the Medical Research Council of Canada.

References