Insertion mutant of bacteriophage f1 sensitive to *Eco*RI

(biological linker/directed mutagenesis/f1 cloning vector/f1/pBR322 chimera)

JEF D. BOEKE, GERALD F. VOVIS, AND NORTON D. ZINDER

The Rockefeller University, New York, New York 10021

Contributed by Norton D. Zinder, March 26, 1979

ABSTRACT The nucleotide sequence A-A-T-T was inserted into the intergenic region of the f1 genome at a site cleaved by *Hae* III (cleavage sequence (G-G+C-C). The resultant viable phage mutant (R199) contains a single site sensitive to the restriction endonuclease *Eco*RI (cleavage sequence G+A-A-T-T-C). This phage is sensitive to *Eco*RI restriction and modification *in vivo* and *in vitro*. Its potential for use as a cloning vector has been tested by construction *in vitro* of an f1/pBR322 chimeric phage. The four bases inserted into wild-type f1 to generate the R199 mutant came from a small restriction fragment obtained by digesting plasmid pBR322 with *Eco*RI and *Hind*III. The use of this linker prepared from a biological substrate is an example of a technique for constructing restriction enzyme sites *in vitro*. It is presented as an alternative to the use of synthetic linkers and should be generally applicable.

The filamentous bacteriophage f1 (and the closely related phages M13 and fd) have several properties that make them particularly desirable as vectors for the cloning of foreign DNA. These include the presence of an intergenic region (1, 2), certain sections of which are not required for normal growth of the phage and can hence be used as sites for inserting foreign DNA (3–5). Furthermore, the apparent lack of packaging constraints allows for the cloning of long stretches of foreign DNA (6).

In addition, the single strandedness of the genomic DNA allows for ready purification of large amounts of single-stranded DNA. Single-stranded DNA is required in such procedures as: rapid sequence determination by the dideoxy or plus-minus methods (7, 8); pulse chase restriction mapping (9); directed mutagenesis techniques (10, 11); heteroduplex formation (12); and potentially for determining the transcriptional polarity of cloned genetic material. Finally, the limited host range of the male-specific filamentous phages (13) makes them desirable from a containment point of view, because most potential hosts in nature are female (14).

A desirable feature of any vector for use in recombinant DNA research is the presence of a single site sensitive to a restriction enzyme that generates cohesive termini, such as *Eco*RI (15). For a phage vector it is especially useful if this site is in a nonessential part of the genome. Wild-type f1 possesses no restriction site that fulfills these criteria. We report here the construction in vitro of an f1 mutant that contains a single *Eco*RI site within a nonessential portion of the intergenic region. We chose to introduce an EcoRI site into f1 because a mutant containing such a site is potentially susceptible to EcoRI restriction and modification in vivo and hence distinguishable from wild type, which is not restricted *in vivo* or *in vitro* by EcoRI. To insert an EcoRI site into the f1 genome, we developed a technique in which the linker was isolated from a biological substrate (Fig. 1) and which we present as an alternative to the use of synthetic linkers (16, 17).

MATERIALS AND METHODS

The Escherichia coli strains have been described (18-21). K507 is K38 transformed by the CaCl₂ technique with the plasmid pMB 4 (22), is ampicillin resistant, and carries the RI restriction and modification systems. H560 (23) was obtained from K. Marians. $\left[\alpha^{-32}P\right]TTP$ was purchased from Amersham Radiochemicals. Unlabeled triphosphates were purchased from P-L Biochemicals. T4 ligase and restriction endonucleases Hae II. HincII, and Hae III were obtained from Bethesda Research Laboratories (Rockville, MD). EcoRI and HindIII were purchased from New England BioLabs. DNA polymerase I large fragment (24) was purchased from Boehringer Mannheim. T4 ligase unable to ligate blunt-ended molecules was the generous gift of C. Yehle. Polyacrylamide and agarose gels (25), preparation of f1 replicative form (RF) and single-stranded DNA (26), and isolation of pBR322 RF were as described (27). Isopycnic centrifugation to remove RFI from other species was as described (26). The cohesive termini of pBR322 generated by cleavage with *Eco*RI were filled in by incubation with the large fragment of DNA polymerase I (24) in 0.2 ml of buffer (6.6 mM Tris-HCl, pH 7.6/6.6 mM NaCl/6.6 mM MgCl₂/6.6 mM dithiothreitol) containing 33 μ M each of dATP, dGTP, and dCTP. $|\alpha^{-32}P|TTP$ (60 μ Ci of 350 Ci/mmol; 1 Ci = 3.7 × 10¹⁰ becquerels) was added carrier free to the reaction to monitor incorporation and serve as a marker for the detection of the linker on gels. Enzyme $(3 \mu l)$ (supplied at a nominal concentration of 892 units/ml) was diluted with 15 μ l of 0.1 M phosphate buffer in 50% glycerol. Incubation was at room temperature for 10 min, followed by the addition of 10 μ l of 1 mM TTP and 10 min of additional incubation. The biological linker was recovered by cutting the appropriate portion of the gel into 1-mm squares and soaking the slices in 1.5 ml of buffer (0.6 M sodium acetate/0.1 M Tris-HCl, pH 8.0/2.5 mM EDTA) overnight at 37°C, followed by filtration and ethanol precipitation. The resultant pellet was dissolved in a small volume, reprecipitated, and washed with 70% ethanol before resuspension. Ligation of blunt-ended DNA was performed in a volume of $20 \ \mu$ l in a buffer containing 20 mM Tris-HCl (pH 7.4), 7.5 mM MgCl₂, 1 mM EDTA, and 50 μ g of bovine serum albumin per ml at 4°C fo. 16 hr. Ligation of cohesive termini was carried out in the same buffer but at 19°C. Transfection was by the $CaCl_2$ technique (28). DNA sequence analysis was by the dideoxy method of Sanger *et al.* (8).

RESULTS

Biological Linker Technique. The principle of this method is that by ligation of two blunt-ended pieces of DNA, each containing a portion of the desired restriction endonuclease recognition sequence, it is possible to generate the desired se-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: RF, double-stranded, replicative form DNA—e.g., RFI, superhelical circular form; RFII, circular form with one or more nicks; RFIII, linear, unit-length form (generated by a restriction enzyme *in vitro*); and RFIV, circular form lacking single-strand nicks and superhelicity.



FIG. 1. Outline of the biological linker technique and its use to generate insertion mutants of bacteriophage f1 sensitive to EcoRI. f1 RF was partially digested with *Hae* III to give 40% RFIII, a molecule having blunt ends with 5'-pC residues. This was ligated to a biological linker isolated from pBR322 RF. The linker was prepared by incubating pBR322 RF with EcoRI, then with the Klenow fragment of DNA polymerase I and dATP and $[\alpha^{-32}P]$ TTP, and finally with *Hind*III. The 31-base-pair linker fragment was gel-purified and ligated to the f1 RFIII by using a T4 ligate preparation containing "blunt end" activity. The DNA in the ligation mixture was cleaved with EcoRI. The RFIII-sized material was purified, diluted (to promote intramolecular ligation), and religated to tircularize the products. Asterisks denote residues labeled with ³²P.

quence. This approach was used in the cloning of the cI gene of bacteriophage λ (29). Any DNA fragment having blunt (base paired) ends with 5'-terminal pC residues and 3'-terminal G residues can be used with a biological linker of the type described below to generate terminal *Eco*RI sites on the fragment. The plasmid pBR322 contains a single *Eco*RI site 31 nucleotides from its unique *Hin*dIII site. Thus, it is possible to cut pBR322 with *Eco*RI and then fill in the cohesive termini by incubation with DNA polymerase, dATP, and TTP. After heat inactivation of the polymerase, the RFIII is digested with *Hin*dIII. This digestion generates two fragments, the approximately 4000nucleotide fragment comprising most of the pBR322 genome and a 31-nucleotide fragment that we refer to as the "biological" linker.

One end of the biological linker consists of the cohesive terminus produced by *Hin*dIII digestion, whereas the other consists of a filled-in *Eco*RI terminus (blunt end). Only the latter end is reactive in blunt-end ligation. When excess biological linker is incubated with an appropriate DNA fragment in the presence of T4 ligase having blunt-end activity, any fragment reacting with a molecule of linker on each end will be bounded by *Eco*RI sites. Removal of unreacted linker (30) is unnecessary because neither the linker itself nor its multimers contain any complete *Eco*RI sites to compete for the *Eco*RI enzyme in the subsequent cleavage step necessary to generate cohesive *Eco*RI termini.

Introduction of EcoRI Site into fl Genome. To put an EcoRI site into the intergenic region of f1 by the above technique, we used as our substrate RFIII molecules generated by partial digestion of f1 RFI with the enzyme Hae III, which generates blunt ends with 5'-terminal pC residues and 3'-terminal G residues. The partial digest contained approximately 30% RFI, 30% RFII, and 40% RFIII. Because the f1 genome contains nine Hae III sites, two of which are in the intergenic region, and the other seven in coding regions, it was clear that the insertion of four nucleotides into seven of the nine sites would probably result in lethal frameshift mutations. Of the two sites in the intergenic region, one is known to be nonessential, because the M13 strain mp 1 (3) has an insertion of approximately 800 bases of the *lac* operon inserted in what corresponds to the fl Hae III G/D border. The other fl site [the F/G border (see below)] is near the origin of minus strand replication, and the effects of insertion there are not predictable. After removal of RFI by isopycnic centrifugation, the RFIII was incubated with a ten-fold excess of biological linker (based on the molarity of ends) and T4 "blunt end" ligase (Fig. 2, lane a). The resultant mixture was cleaved with EcoRI and loaded on a 0.8% agarose gel to isolate RFIII from RFII and other contaminants. The RFIII was ligated under dilute conditions by using a preparation of T4 ligase containing no detectable blunt-end activity. The absence of blunt-end activity prevents



FIG. 2. Blunt-end ligation of biological linker to linearized f1 genome and recircularization by intramolecular ligation of genome after EcoRI cleavage. f1 RFIII produced by partial digestion of f1 RFI with *Hae* III was incubated with a 10-fold excess of biological linker. For lane a, 1% of the reaction volume was diluted and loaded on a 2.4-10% polyacrylamide gradient gel, which was then autoradiographed overnight. The lowermost band seen is unreacted linker. The second band from the bottom is linker dimer, which can be formed either by cohesive ligation (of the HindIII ends) or blunt-end ligation. The products in the higher bands are linker trimers, tetramers, etc. In order to create these multimers, one or more blunt-end ligation events is required; hence the extent of multimer formation in a ligation reaction of this type reflects the extent of blunt-end ligation. The uppermost band is RFIII ligated to linker molecules. Ligation products were digested with EcoRI and loaded on a vertical 0.8% agarose slab gel. The band comigrating with marker RFIII was cut out and the RFIII was recovered by electroelution. This DNA was used for intramolecular ligation of the EcoRI cohesive termini. Small aliquots of this reaction were run on a 0.8% agarose gel before (lane b) and after (lane c) this ligation step. Arrows indicate the position of (top to bottom) f1 RFII, RFIII, and RFI OD markers. RFI and RFIV comigrate in this gel system.

the circularization of those RFIII molecules to which no biological linker was added. Circularization of such material would yield infectious molecules that lack an *Eco*Ri site. A small anquot was run on a 0.8% agarose gel to monitor the extent of circularization, which was estimated at approximately 20% (Fig. 2, lanes b and c). After transfection of CaCl₂-treated K414 cells, plaques were picked and tested for their relative efficiency of plating on K38 ($r_{\rm RI}^- m_{\rm RI}^-$) and K507 ($r_{\rm RI}^+ m_{\rm RI}^+$). Several isolates showed classical restriction modification behavior. One such isolate (R199) had a restriction coefficient of 0.2 on K507 compared to K38 and produced typical f1 plaques.

Localization of EcoRI Site. Viability and sensitivity to restriction and modification were presumptive evidence for the proper insertion into the intergenic space. R199 RFI and single-stranded DNA were prepared. To locate precisely the EcoRI site, the RF was subjected to restriction analysis by using endonucleases EcoRI, Hae II, and Hae III (Fig. 3). Compared to wild-type f1, the restriction pattern obtained by digesting R199 with Hae III shows a new band appearing at about 450 base pairs and the loss of fragment G and one of the bands in the DEF cluster. This result was expected because the introduction of an A-A-T-T sequence at either Hae III F/G or D/G should abolish that *Hae* III site and yield in either case a fragment of about 450 base pairs (ref. 25; see Fig. 4). The normal Hae III fragment pattern is restored upon the addition of EcoRI enzyme. The new restriction site was localized at Hae III F/G by a combined Hae II/Hae III restriction digest. Because the Hae III F fragment contains a Hae II site whereas neither G nor D does, the sensitivity of the new 450-base-pair fragment to Hae II indicates that the new fragment is FG. The nucleotide sequence of the mutant region was determined by priming R199 single-stranded DNA with the wild-type fragment HinfK (31), and the sequence G-G-C-C in the wild type was seen to be replaced G-G-A-A-T-T-C-C in R199 at the Hae III F/G border. A second f1 mutant (R209) containing the EcoRI site at the Hae III G/D border was also isolated (data not shown).

Use of R199 as Cloning Vector. To demonstrate the potential usefulness of R199 as a cloning vector, limit *Eco*RI digests of R199 and pBR322, a plasmid conferring tetracycline



FIG. 3. Localization of EcoRI site in R199. Restriction endonuclease digests of wild-type and R199 RF were electrophoresed on a 2.4–7.5% polyacrylamide gel containing 0.5 μ g of ethidium bromide per ml and stained for 10 min in buffer containing 0.5 μ g of ethidium bromide per ml. (Lane a) *Hae* III digest of wild-type RF. Capital letters along left side of figure denote names of f1 *Hae* III fragments. (Lane b) *Hae* III digest of R199 RF. Note the prominent hybrid band (F/G) caused by insertion of the sequence A-A-T-T into the *Hae* III site. (Lane c) Combined *Hae* III/*Eco*RI digest of R199 RF. (Lane d) Combined *Hae* III/digest of wild-type RF. (Lane e) Combined *Hae* III digest of R199 RF.



FIG. 4. Genetic and physical map of f1 genome. The upper line represents the physical map of f1. The letters and vertical lines denote the *Hae* III fragments and sites, respectively. The arrows beneath the line represent the *Hae* II sites. The M13 *Hae* II and *Hae* III cleavage maps are identical except for an additional *Hae* III site in *Hae* III A. The middle line represents the genetic map of f1. Roman numerals indicate f1 complementation groups. IG indicates the intergenic region. The lower line is a measure of the size of the physical (lower, in kilobases) and genetic (upper, in map units) maps. The location of the A-A-T-T insertion in R199 is indicated by the heavy arrow. The *Eco*RI site in R209 maps to the *Hae* III G/D border.

and ampicillin resistance to its host, were ligated by using T4 ligase. The reaction mixture was used to transfect the *polI* mutant H560. In this host, the plasmid pBR322 cannot replicate by itself due to its ColE1 origin of replication (32). All of the ampicillin-resistant transformants isolated produced recombinant phage that exhibited the electrophoretic mobility expected of a chimeric phage containing the two genomes (data not shown). One such isolate, called R208, was studied in detail, and its ability to transduce ampicillin resistance was seen to be sensitive to f1 antiserum. RFI was prepared from R208 and digested with EcoRI. The restriction pattern suggests the presence of an entire copy on pBR322 within R199 (Fig. 5).

DISCUSSION

We have shown how a restriction fragment from the plasmid pBR322 can be used to generate EcoRI restriction sites in vitro. Our technique is presented as an alternative to the synthetic linker techniques recently described (16, 17). In our procedure, two restriction fragments, each of which contains a portion of the desired sequence, are joined together. These sites can then be cleaved by the appropriate restriction enzyme (in our case EcoRI) to leave cohesive termini on those fragments to which the biological linker was ligated. Our method should be generally applicable because other restriction sites could be prepared by following a similar rationale. A small restriction fragment is convenient for use as a biological linker because it allows a simple assay for the success of the blunt-end ligation (Fig. 2, lane a). If the enzymes used in preparation of biological linkers are known to be relatively free of contaminating nucleases, the sequence purity of the linker is guaranteed by the specificity of the enzymes. Such sequence purity may be difficult to attain by using synthetic methods. A substantial proportion of the ends generated by the biological linker technique



FIG. 5. EcoRI cleavage of R208, a pBR322/f1 chimera. RF DNAs were cleaved with restriction endonucleases and run on a 0.8% agarose gel containing 0.5 μ g of ethidium bromide per ml. Lanes: a, undigested pSC101 RF (9.6 kilobases); b, undigested f1 RF (6.5 kilobases); c, undigested R199 RF; d, undigested R208 RF; e, HincII digest of f1 RF (RFIII); f, EcoRI digest of R199 RF; g, EcoRI digest of R208 RF; h, EcoRI digest of pBR322 RF.

are chemically and biologically intact, as shown by the fact that most of the cohesive *Eco*RI termini generated could be ligated to circularize the fragment (Fig. 2, lanes b and c). Currently, the main drawback of the technique is that rather large amounts of pBR322 RF are required for each experiment, because each molecule of pBR322 produces only one molecule of biological linker. We are currently attempting to construct plasmid variants that will carry multiple copies of the biological linker, thereby alleviating this problem.

By use of this technique, the sequence A-A-T-T was inserted at two different sites within the intergenic region of f1. One of these sites, the Hae III F/G border region, has been implicated in the replication of the filamentous phage (33, 34). Based on sequence data, this region has potential secondary structure (35). In fact, the hairpin-like structure containing the *Hae* III F/Gborder has been implicated in the initiation of DNA replication. Therefore, the insertion of foreign DNA at this site might be expected to prevent DNA replication. The isolation of the f1/pBR322 recombinant phage (R208) indicates that replication of such molecules is possible and appears to reduce the likelihood of an absolute requirement for secondary structure in this region for the initiation of DNA replication. Moreover, the use of the poll mutant H560 [in which pBR322 cannot replicate independently (32)] ensures that the bacteriophage origin of replication is used exclusively in chimeric phage production.

The filamentous phage, because their genome consists of a single-stranded DNA molecule, offer a set of unique advantages as cloning vectors. Milligram quantities of both single-stranded and double-stranded DNA can be easily prepared, even when the genome size is increased substantially. In terms of efficient production of cloned material, filamentous phage potentially offer great advantages, because a single wild-type plaque routinely yields greater than 10^9 plaque forming units, and phage production is on the order of 10^{12} per ml. Also, the vector genome size is very small, so that the cloned DNA constitutes a large fraction of the recombinant DNA. The availability of cloned DNA in both single-stranded and double-stranded forms facilitates its rapid structural analysis. By using single-stranded DNA from chimeric phage as template and wild-type or chimeric restriction fragments as primers, pulse-chase restriction mapping and DNA sequence determination by using dideoxy terminators are possible. Also, one should be able to determine the transcriptional polarity of the cloned DNA by hybridization of the RNA in question to genomic DNA from chimeric phages carrying the insert in opposite orientations. Finally, single-stranded DNA is one of the substrates required for directed mutagenesis techniques which employ a mismatched primer (10, 11). Thus, point mutations might be introduced in a specific manner into the foreign genetic material inserted in the phage genome. Limits on the size of the DNA to be cloned have yet to be determined.

We thank Dr. P. Modrich for supplying us with the plasmid pMB 4, P. McGraw for her excellent technical assistance, K. Horiuchi and P. Model for critically reading the manuscript, and C. Yehle for generous gifts of enzyme. This work was supported in part by funds from the National Science Foundation and the National Cancer Institute.

 Vovis, G. F., Horiuchi, K. & Zinder, N. D. (1975) J. Virol. 16, 674–684.

- Enea, V., Horiuchi, K., Turgeon, B. G. & Zinder, N. D. (1977) J. Mol. Biol. 111, 395–414.
- Messing, J., Gronenborn, B., Müller-Hill, B. & Hofschneider, P. H. (1977) Proc. Natl. Acad. Sci. USA 74, 3642–3646.
- Herrmann, R., Neugebauer, K., Schaller, H. & Zentgraf, H. (1978) in *The Single-Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 473–476.
- 5. Ray, D. S. & Kook, K. (1978) Gene 4, 109-119.
- Vovis, G. F., Ohsumi, M. & Zinder, N. D. (1977) in Molecular Approaches to Eukaryotic Gene Systems, eds. Wilcox, G., Abelson, J. & Fox, C. F. (Academic, New York), pp. 55-61.
- Sanger, F. & Coulson, A. R. (1975) J. Mol. Biol. 94, 441-448.
 Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 9. Seeburg, P. H. & Schaller, H. (1975) J. Mol. Biol. 92, 261-277.
- Hutchison, C. A., III, Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P. & Smith, M. (1978) J. Biol. Chem. 253, 6551–6560.
- 11. Razin, A., Hirose, T., Itakura, K. & Riggs, A. D. (1978) Proc. Natl. Acad. Sci. USA 75, 4268-4270.
- 12. Vovis, G. F., Horiuchi, K., Hartman, N. & Zinder, N. D. (1973) Nature (London) New Biol. 246, 13-16.
- Zinder, N. D., Valentine, R. C., Roger, M. & Stoeckenius, W. (1963) Virology 20, 638-640.
- 14. Lederberg, J. (1951) Science 114, 68-69.
- Mertz, J. E. & Davis, R. W. (1972) Proc. Natl. Acad. Sci. USA 69, 3370–3374.
- Bahl, C. P., Marians, K. J., Wu, R., Stawinsky, J. & Narang, S. A. (1977) Gene 1, 81–92.
- Scheller, R. H., Dickerson, R. E., Boyer, H. W., Riggs, A. D. & Itakura, K. (1977) Science 196, 177–180.
- Garen, A. & Siddiqi, O. (1962) Proc. Natl. Acad. Sci. USA 48, 1121-1127.
- 19. Lyons, L. B. & Zinder, N. D. (1972) Virology 49, 45-60.
- 20. Cole, R. S. (1971) J. Bacteriol. 106, 143-149.
- 21. Ohsumi, M., Vovis, G. F. & Zinder, N. D. (1978) Virology 89, 438-449.
- Betlach, M., Hershfield, V., Chow, L., Brown, W., Goodman, H. M. & Boyer, H. W. (1976) Fed. Proc. Fed. Am. Soc. Exp. Biol. 35, 2037–2043.
- Vosberg, H. P. & Hoffmann-Berling, H. (1971) J. Mol. Biol. 58, 739–753.
- Klenow, H. & Henningsen, I. (1970) Proc. Natl. Acad. Sci. USA 65, 168–175.
- 25. Horiuchi, K., Vovis, G. F., Enea, V. & Zinder, N. D. (1975) J. Mol. Biol. 95, 147-165.
- 26. Model, P. & Zinder, N. D. (1974) J. Mol. Biol. 83, 231-251.
- 27. Clewell, D. B. (1972) J. Bacteriol. 110, 667-676.
- 28. Mandel, M. & Higa, A. (1970) J. Mol. Biol. 53, 159-162.
- Backman, K., Ptashne, M. & Gilbert, W. (1976) Proc. Natl. Acad. Sci. USA 73, 4174–4178.
- Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) Cell 15, 687-701.
- Horiuchi, K., Vovis, G. F. & Model, P. (1978) in *The Single-Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 113–137.
- 32. Kingsbury, D. T. & Helinski, D. R. (1970) Biochem. Biophys. Res. Commun. 41, 1538–1544.
- Horiuchi, K. & Zinder, N. D. (1976) Proc. Natl. Acad. Sci. USA 73, 2341–2345.
- Geider, K., Beck, E. & Schaller, H. (1978) Proc. Natl. Acad. Sci. USA 75, 645-649.
- Schaller, H., Beck, E. & Takanami, M. (1978) in *The Single-Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 139–163.