

Site-specific dissection of *E.coli* chromosome by λ terminase

Hirokazu Kotani*, Akira Kawamura, Asako Takahashi, Masako Nakatsuji, Nobutsugu Hiraoka, Kazuo Nakajima and Mituru Takanami¹

Bioproducts Development Center, Takara Shuzo Co. Ltd, Seta 3-4-1, Ohtsu, Shiga 520-21 and

¹Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan

Received April 7, 1992; Revised and Accepted June 9, 1992

DDBJ accession nos D00927 and D00928

ABSTRACT

We have succeeded the targeted cleavage of chromosomes by λ terminase that introduces double-strand cleavages in DNA recognizing the λ *cos* sequence. When chromosomal DNAs of various *Escherichia coli* K-12 strains were subjected to terminase digestion, all were found to contain two common cleavage sites. Therefore, DNAs from λ lysogens in which λ DNA was inserted at different chromosomal sites were specifically cleaved at one more additional site. The two sites, termed *ecos1* and *ecos2*, were mapped at approximately 35.1' and 12.7' of *E.coli* genetic map. The *ecos1* and *ecos2* sites were included in *qin* and *qsr'* regions, respectively. Therefore, the cleavage sites were associated with cryptic prophages. Sequences at the *ecos1* and *ecos2* sites showed 98% homology to the λ *cos* sequence, indicating high fidelity of sequence recognition by the terminase. Since the strategy for integration of a DNA segment into chromosomal DNA through homologous recombination has been established, the dissection method that uses λ terminase should be applicable for gene mapping as well as construction of macro-physical maps of larger genomes.

INTRODUCTION

λ terminase is a hetero-oligomer consisting of the *A* and *Nul* gene products of λ phage (1, 2). The enzyme introduces two nicks, staggered 12 nucleotides apart, at the *cos* site of the λ DNA molecule and packages it into the prohead. The minimal essential region required for this cutting has been narrowed down to position -22 to +38 in the λ DNA coordinate (3, 4). Employing this cutting activity, it seems to be possible to cleave large DNA molecules. Recently, we developed the commercially available λ terminase for cosmid mapping. This enabled us to examine the action of the terminase on the large DNA molecules. We have chosen the *E.coli* chromosome as a model system, for its physical map has been constructed and insertion of λ phage into the genome can be manipulated. The genome of *E.coli* K-12 contains,

in addition to λ phage, at least three λ -related segments, *rac* (5), *qin* (6) and *qsr'* (7). Because two of these segments, *qin* and *qsr'*, contain *cos*-like sequences functionally identical to λ *cos*, it is expected to cleave at these sites by λ terminase *in vitro*. We report here that the *E.coli* DNA in agarose gel blocks is specifically cut λ terminase at *cos* sites.

MATERIALS AND METHODS

Bacterial strains

E.coli W3350, JM109, W3110, and MC1061 were stock strains of our laboratories. *E.coli* HI1005 (MC1000*trpA*_{9605(am)thy}) and HI1098 (HI1005 Δatt^{λ} *nadA* *ilv* *att^{\lambda}* *rho*) were generous gifts from H.Imai (Kyoto University) (8). These strains were used as the substrates for digestion of chromosomal DNAs by λ terminase or as host for DNA manipulations.

Materials

λ terminase, restriction enzymes, DNA ligation kit, Random primer labeling kit, Gene Mapping Membrane of *E.coli* (9), pUC18 DNA and λ DNA were supplied from Takara Shuzo and used according to the manufacturer's instructions. InCert gel for the preparation of agarose gel blocks and Seakem GTG agarose for the pulsed field gel electrophoresis were products of FMC (Philadelphia, PA). [α -³²P] dCTP (>3,000 Ci/mmol) was from Amersham. All other chemicals were of reagent grade.

Digestion of *E.coli* chromosomal DNAs

Agarose gel blocks embedded with *E.coli* DNAs were prepared as described by Schwartz and Cantor (10). These chromosomal DNAs in agarose blocks were digested by λ terminase under the conditions recommended by the supplier (Takara Shuzo). The blocks were equilibrated for 1 hr in 100 μ l of the reaction buffer supplied by the manufacturer. Then 40 units of λ terminase were added and incubation was done 16 hrs at 30°C. The digests were resolved for 40 hrs at 100 mA with linear pulsed time gradient from 40 to 300 sec (Fig. 1a) or for 96 hr at 50 mA with linear pulsed time gradient from 600 to 1,800 sec and then for 12 hr

* To whom correspondence should be addressed at: Ishida cho 910-51, Moriyama, Shiga 524, Japan

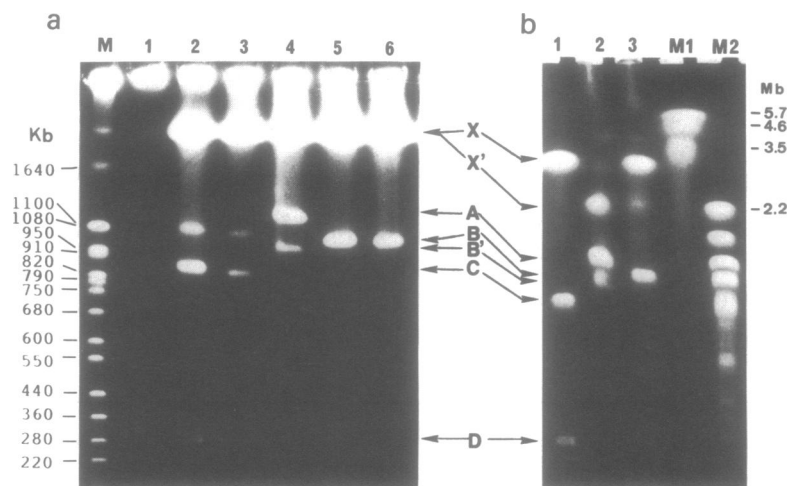


Fig. 1. **a** Dissection of *E. coli* chromosomal DNAs by λ terminase. DNAs from W3350 (λ) (lanes 1,2), HI1005 (λ) (lane 3), HI1098 (λ) (lane 4), MC1061 (lane 5) and JM109 (lane 6) were prepared in agarose gel blocks (1–2 $\mu\text{g}/50 \mu\text{l}$). DNA fragments generated were labeled as in the right side. Lane 1 was undigested control of *E. coli* W3350. In lane M, *Saccharomyces cerevisiae* genome as size makers was run. Their sizes (kb) are given in the left side. **b.** DNAs from W3350 (λ) (lane 1), HI1098 (λ) (lane 2) and JM109 (lane 3) were treated same as **Fig. 1a**, except the conditions of PFGE as described in **Methods and Materials**. In lane M1 and M2, chromosomes of *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* were run as size markers.

300 sec of pulsed time (**Fig. 1b**) in the CHEF-DR system of BioRad, and stained with ethidium bromide.

Cloning and Mapping of *cos1* and *cos2*

All of procedures of DNA manipulations were carried out according to Maniatis *et al* (11). The W3110 DNA was digested with *NspI* and after resolution by gel electrophoresis, Southern hybridization was performed with the λ *cos* probe [the 1 ~ 415 bp region of the λ DNA map (12)]. Then cloning of the two regions into the *SphI* site of pUC18 was carried out by the colonies hybridization, and two clones, pECOS1 and pECOS2, respectively carrying the 2.1 kb and 0.7 kb inserts were isolated. These clones were analyzed by DNA sequencing and hybridization using Gene Mapping Membrane and phage DNAs of the Kohara library of the W3110 chromosome (9).

RESULTS AND DISCUSSION

The genome of *E. coli* K-12 contains two *cos*-like sequences functionally identical to λ *cos*. Murialdo (13) reported that the recombination deficient *E. coli* cells containing the cloned λ terminase genes were lethal because of cleavage of *cos*-like sequences by λ terminase *in vivo*. It is expected to cleave specifically at these sites by λ terminase *in vitro*. As shown in lanes 5 and 6 of **Fig. 1a**, DNAs from the non-lysogens were cleaved into two pieces, indicating that two cryptic *cos* sites in *qin* and *qsr'* were susceptible to the terminase. When other *E. coli* strains, W3110, MC1061, HB101 and JM109, were examined, DNAs of all the strains were found to generate two fragments with almost identical sizes: The smaller one was estimated to be about 1,100 kilo-base pairs (kb) (band B in **Fig. 1a**). The other one (band X in **Fig. 1a**) was too large to estimate the accurate size by co-migrating in the compression zone, but that band was well resolved and estimated to be over 3,000 kb as shown in lane 3 of **Fig. 1b**.

When the DNAs from λ lysogens of W3350 and HI1005 were examined, the smaller one (band B) was split into two pieces

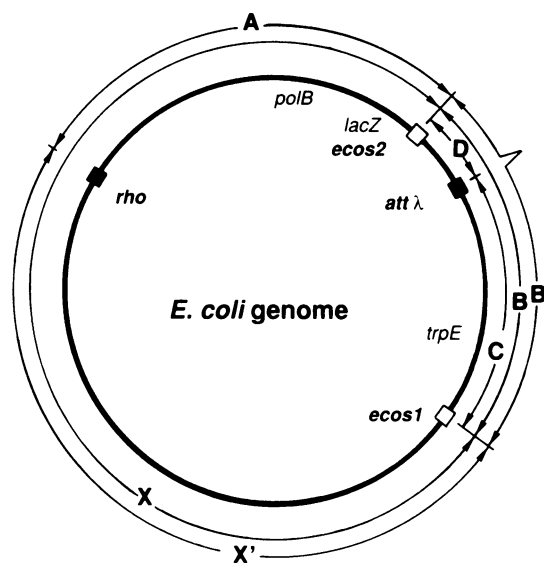


Fig. 2. Schematic cleavage pattern of *E. coli* genome. The cryptic cleavage sites, *ecos1* and *ecos2*, and fragments generated by cleavages were assigned as indicated on the circular map of *E. coli* genome, on which the positions of *att λ* and *rho* sites and used probes are also shown.

about 900 kb and 250 kb (band C and D in **Fig. 1a**, lanes 2 and 3: undigested band B is also seen, and **Fig. 1b**, lane 1). Band D in lane 2 of **Fig. 1a** appears doublets, but we assume that it is an artifact of electrophoresis, for it appears a singlet or doublets depending on runs. We then analysed the DNA from another lysogen, HI1098 (λ). In this strain, the authentic *att λ* region (17') of the *E. coli* genome was deleted and instead, a 1.6 kb fragment carrying *att λ* was inserted near the *rho* gene at 84' (8), to which λ prophage was inserted. By digestion of the HI1098 (λ) DNA, we observed generation of three fragments with different sizes (**Fig. 1a**, lane 4 and **Fig. 1b**, lane 2): a large

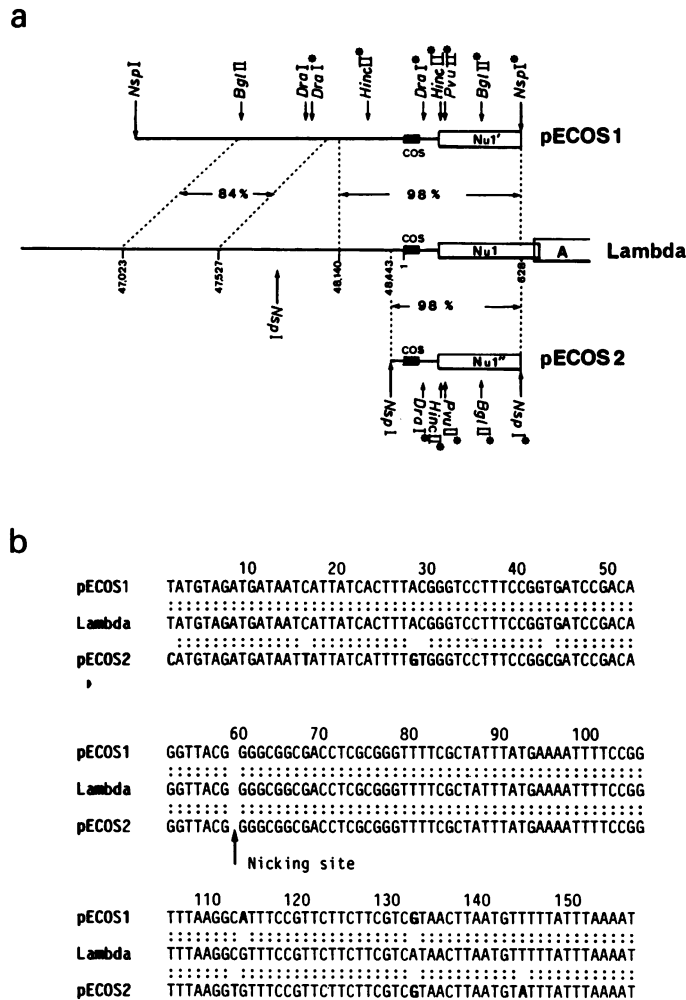


Fig. 3. a The restriction maps of the inserts in pECOS1 and pECOS2 carrying the *ecos1* and *ecos2* sites, respectively. In the middle, the corresponding region of λ genome were indicated, and the *Nul* and *A* genes in boxes. The restriction sites of λ genome that had been conserved in the clones were indicated by asterisks, and the extents of sequence homology among λ genome and clones by percentages. The regions corresponding to the *Nul* gene in the clones were boxed (*Nul*¹ and *Nul*²). b The nucleotide sequences of 160 bp around the λ *cos* region in pECOS1 and pECOS2 were indicated aligned with respect to the λ nicking site.

fragment at the band X' region, about the 1,300 kb fragment (band A) and a fragment (band B') slightly smaller than the 1,100 kb fragment (band B). As shown in lane 2 of Fig. 1b, the size of band X' was estimated to be about 2,300 kb.

To assign the cryptic cleavage sites on the *E. coli* genetic map, DNA fragments resolved were transferred onto nitrocellulose filters, and hybridization tests were carried out with DNA probes carrying *polB* (2') (14), *lacZ* (8') (15) and *trpE* (27') (16). The *trpE* probe predominantly hybridized to band C (in Fig. 1a, lane 2 & 3), band B (in lanes 2, 3, 5 and 6) and band B' (in lane 4). The other two probes hybridized to band A (in lane 4) and band X (in lanes 2, 3, 5, & 6) (data not shown). Taking into account the sizes of fragments, the map positions of two cryptic *cos* sites were coincided with these of *qin* (34') and *qsr*' (12.5') (6, 7). Therefore, we termed these *cos* sites derived from cryptic prophages were *ecos1* in *qin* region and *ecos2* in *qsr*' region (Fig. 2). The shorter size of band B' generated from HI1098 (λ), compared with band B from other strains, is probably due to the

deletion of the *att* ^{λ} region at 17' during the strain construction. The assignment of the two cryptic sites has further been confirmed by cloning experiments, as described below.

To determine the maps and sequences of the two cryptic sites more precisely, the W3110 DNA was digested with *NspI* and after resolution by gel electrophoresis, hybridization was performed with the λ *cos* probe. Two bands of about 2.1 kb and 0.7 kb were detected (data not shown). Then cloning of the two regions into the *SphI* site of pUC18 was carried out, and two clones, pECOS1 and pECOS2, respectively carrying the 2.1 kb and 0.7 kb inserts were isolated. We determined the entire sequences of the inserts and found that the regions encompassing the *cos* region have highly been conserved: the extent of homology was about 98% (Fig. 3a). Fine cleavage maps of pECOS1 and pECOS2 are shown in Fig. 3a, in comparison with the λ genome map. In Fig. 3b, the sequences of 160 bp around the *cos* region are shown, in which the minimal essential and enhancing regions (3, 4) are included. The data imply a high sequence fidelity of cleavages with λ terminase. The sequence data of pECOS1 (Accession No. D00927) also suggest that insertion and rearrangement of the λ -like sequence had occurred in ancient times, as the conserved regions appear in mosaic reported by Kaiser *et al* (6, 7, 13) previously. When the clones were analysed by using the Kohara library (9) of the W3110 chromosome, pECOS1 was hybridized to clone No. 22D11 at 35.1' using Gene Mapping Membrane (clone No. 22D11 was recently added to the library by Dr. Y. Kohara) and pECOS2 to clone No. 23E10 at 12.7' by Southern hybridization (data not shown).

We demonstrated specific cleavages of *E. coli* chromosomes by using with λ lysogens. Although the exact sequences required for full cutting activity are not defined yet, we confirmed that the terminase cleaves the 415 bp *HincII* fragment from positions 48,515 to 199 in the λ coordinate (12) (see Fig. 3a) as efficient as λ genome and cosmids. In the *E. coli* and yeast systems, integration of DNA segments into chromosomal DNAs by homologous recombination has generally been used for genetic analysis. This principal has also been applied for gene targeting in higher eukaryotes (17). Employing λ terminase, therefore, large chromosomal DNAs could be dissected at any sites by inserting a DNA segments tagged with the λ *cos* sequence.

ACKNOWLEDGEMENTS

The authors wish to thank Drs. S. Yasuda, M. Ooki, M. Yanagida and S. Niwa for valuable discussions, Dr. M. Oishi for critical reading of the manuscript, Dr. M. Imai for providing *E. coli* strains, HI1005 and HI1098, and Dr. Y. Kohara for denoting his clone library.

REFERENCES

1. Becker, A., Murialdo, H. & Gold, M. (1977) *Virology*, **78**, 277–290.
2. Summer-Smith, M., Becker, A. & Gold, M. (1981) *Virology*, **111**, 642–646.
3. Feiss, M., Widner, W., Miller, G., Johnson, G. & Christiansen, S. (1983) *Gene*, **24**, 207–218.
4. Miwa, T. & Matsubara, K. (1983) *Gene*, **24**, 199–206.
5. Kaiser, K. & Murray, N. E. (1979) *Mol. Gen. Genet.*, **175**, 159–174.
6. Espion, D., Kaiser, K. & Dambly-Chauduere, C. (1983) *J. Mol. Biol.*, **170**, 611–633.
7. Kaiser, K. (1980) *Mol. Gen. Genet.*, **179**, 547–554.
8. Tsurushita, N., Hirano, M., Shigesada, K. & Imai, M. (1984) *Mol. Gen. Genet.*, **196**, 458–464.
9. Kohara, Y., Akiyama, K. & Isono, K. (1987) *Cell*, **50**, 495–508.

10. Schwartz, D. C. and Cantor, C. R., (1984) *Cell*, **37**, 67–75.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
12. Daniels, D. L., Schroeder, J. L., Szybalski, W., Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F, Peterson, G. & Blattner, F. (1983) In Hendrix, R. W., Roberts, J. W., Stahl, F. W. & Weisberg, R. A. (eds) *Lambda II*. Cold Spring Harbor Press, NY, pp519–676,
13. Murialdo, H. (1988) *Mol. Gen. Genet.*, **213**, 42–49.
14. Iwasaki, H., Nakata, A., Walker, G. & Shinagawa, H. (1990) *J. Bacteriol.*, **172**, 6268–6273.
15. Kalnins, A., Ott, K., Rutter, U. & Muller-Hill, B. (1983) *EMBO J.*, **2**, 593–597.
16. Yanofsky, C., Platt, T., Crawford, I., Nichols, B., Christie, G., Horowitz, H., Van Cleemput, A. & Wu, A. (1981) *Nucleic Acids Res.*, **9**, 6647–6668.
17. Mansour, S. L., Thomas, K. R. & Cappecchi, M. R. (1988) *Nature*, **336**, 348–35.