BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **218,** 12–16 (1996)
Article No. 0003

# A Region of the Filamentous Phage $\phi$Lf Genome That Can Support Autonomous Replication and Miniphage Production

Nien-Tsung Lin, Fu-Shyan Wen, and Yi-Hsiung Tseng[1]

*Department of Botany and Institute of Molecular Biology, National Chung Hsing University, Taichung 402, Taiwan, Republic of China*

Received November 13, 1995

A 2028-bp fragment from the RF DNA of $\phi$Lf, a filamentous phage of *Xanthomonas campestris* pv. *campestris,* was maintained autonomously as a minireplicon. Upon superinfection of the cells harboring the minireplicon with $\phi$Lf, transducing miniphage particles were released. The minireplicon contained an open reading frame (ORF346) able to encode a polypeptide of MW 39144, which possessed consensus motifs found in the Rep proteins from various sources. These findings suggested ORF346 to be the gene encoding replication initiation protein, gene II (gII) of $\phi$Lf. Upstream to ORF346 were sequences with potential to form hairpin structures and a sequence similar to the integration host factor (IHF) binding site, structures similar to the intergenic region (IR) of the Ff phages. A 15 bp AT-rich core for $\phi$Lf integration was found 37 bp downstream to the IHF binding site. 

Filamentous phage $\phi$Lf, which specifically infects the Gram-negative plant pathogenic bacterium *Xanthomonas campestris* pv. *campestris*, is similar to other filamentous phages (1) in having a circular, single-stranded DNA genome (6.0 kb), using a replicative form (RF) as the replication intermediate, and manifesting a non-lytic life cycle (2). However, it differs from other filamentous phages in possessing the ability of site-specific integration into the host chromosome (3), and infecting through a receptor other than a pilus (unpublished results). Therefore, viewing these differences, we considered $\phi$Lf an interesting material for molecular biological study. As nucleotide sequence analysis is the first step towards understanding of the genome, we have sequenced the *Eco*RV-*Sph*I fragment (1018 bp) from $\phi$Lf which contains four contiguous genes coding for the single-stranded DNA binding protein (gV), the two minor coat proteins (gVII and gIX) and the major coat protein (gVIII), respectively (4). These genes have the same order as that of the Ff phages (f1, fd and M13), V-VII-IX-VIII (1).

It is established that gII, the gene encoding replication initiation protein, is the only phage function involved in the propagation of fd RF, and cloned gII can support replication of a fragment containing the replication origin (*ori*) of fd (5). Since $\phi$Lf is similar to Ff phages, we predicted that $\phi$Lf contains a region analogous to the *ori* and gII of Ff phages. Therefore, efforts were made in our laboratory to seek for such a region in the $\phi$Lf genome. In this study, a 2028-bp fragment of the $\phi$Lf RF DNA was cloned and shown to support autonomous replication and phage production.

## MATERIALS AND METHODS

*Phage, bacterial strains and cultivation conditions.* Filamentous phage $\phi$Lf has been described previously (2). Strain Xc17 was a wild-type *X. c.* pv. *campestris* strain isolated in Taiwan (6). Xc17NT1 was derived from Xc17 by deletion of a 4.4 kb *Hinc*II fragment containing the *attB* for $\phi$Lf integration. Steps involved in making the deletion were as follows. A 6.1 kb *Eco*RV fragment containing the 4.4 kb region was first cloned in pOK12 (7), then this region was replaced by a tetracycline cartridge from pHP45$\omega$-Tc (8), followed by double crossover of the deleted version into the chromosome. Luria

[1] To whom correspondence should be addressed. Fax: 886-4-287-4879.

Abbreviations: RF, replicative form; ORF, open reading frame; IHF, integration host factor; MW, molecular weight; IR, intergenic region; *ori*, origin of replication; m.o.i., multiplicity of infection; ssDNA, single-stranded DNA.



0006-291X/96 $12.00

broth and L agar (9) were used throughout the experiments to cultivate *Escherichia coli* DH5α and JM101, as the cloning hosts, at 37°C and *X. c.* pv. *campestris* at 28°C.

*Phage titer assay.* Double layer (10) was used to determine the phage ϕLf titer using Xc17 as indicator host. Titer of transducing particle was assayed by infecting Xc17NT1 with appropriately diluted suspension containing the transducing particles at m.o.i. less than 0.01, followed by plating the infected cells onto L agar containing kanamycin (50 μg/ml).

*DNA techniques.* Preparation of plasmid, phage ssDNA and phage RF DNA, restriction enzyme digestion, ligation, preparation of α-[$^{32}$P]-labeled probe, DNA hybridization, and agarose gel electrophoresis were carried out as described by Sambrook *et al.* (11). Transformation was performed by electroporation (12). All enzymes were purchased from Promega Corporation (Madison, Wisconsin). α-[$^{32}$P]-dCTP was purchased from Amersham Life Sciences (England, UK).

*Nucleotide sequence analysis.* Segments from the minireplicon OR1 was cloned into M13mp18 and mp19 (13), and the sequences of both strand were determined by the method of Sanger *et al.* (14). Nucleotide sequences were analyzed by using Release 6.01 of PC/GENE (IntelliGenetics). Tfasta (15) was used to search in the Genetics Computer Group (GCG) for homologous sequences.

## RESULTS AND DISCUSSION

### *Cloning of the Fragment of ϕLf RF Capable of Autonomous Replication*

The ϕLf RF DNA (Fig. 1) was partially digested with *Sau*3A1 and ligated with the gene encoding kanamycin resistance (Km cartridge) from pUC4K (16). The ligation mixture was then electroporated into Xc17NT1, using Km$^r$ as the selection marker for transformants maintaining autonomously replicating replicons. The reasons for using the Km cartridge and Xc17NT1 were of two folds. Firstly, the Km cartridge contained no *ori*, therefore, possibility of maintenance depending on an *ori* from a source other than ϕLf could be excluded. Secondly, Xc17NT1 had the region containing *attB* site for ϕLf integration (4.4 kb) deleted, therefore, integration of the recombinant molecules into the chromosome could be avoided. During these experiments, more than 200 Km$^r$ transformants were obtained. Size comparison by agarose gel electrophoresis revealed that the smallest clone, designated pOR1, contained an insert of about 2.0 kb (Fig. 2, lane 4). This *Sau*3A1 partial fragment of the pOR1 insert, extending from kb 0.3 to 2.3 on the ϕLf physical map (Fig. 1), was called minireplicon OR1. All of the other clones were shown to contain inserts overlapping with OR1, by Southern hybridization using the 0.9-kb *Eco*RI-*Hin*dIII fragment inside OR1 as the probe (data not shown).

### *Release of Transducing Miniphage Particles*

pOR1 was stable in Xc17NT1. After ten passages by patch on L agar plate without kanamycin, 99% of the colonies maintained the replicon. Upon superinfection of Xc17NT1(pOR1) with wild-type ϕLf (m.o.i. about 20), miniphage particles were released along with the ϕLf (Fig. 2, lane 1). The DNA from the miniphage particles was sensitive to DNaseI as did the ssDNA from the ϕLf particles, suggesting the miniphage DNA to be single-stranded in nature (data not shown). The miniphage particles exhibited infectivity, able to transduce the host into Km$^r$. Results of infectivity assay indicated that the supernatant from the ϕLf-superinfected Xc17NT1(pOR1) culture contained approximately $8 \times 10^{10}$ PFU of ϕLf and $2.6 \times 10^6$ transducing miniphage particles per ml.

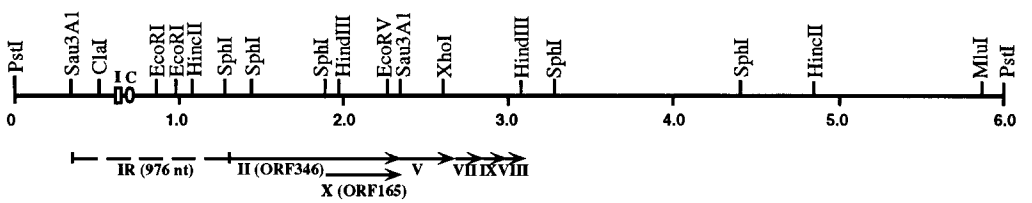


**FIG. 1.** Physical map of ϕLf RF DNA (6.0 kb) linearized at the unique *Pst*I site. Flanked by the *Sau*3A1 sites is the minireplicon OR1 (2028 bp). IR stands for the intergenic region. Letters I and C represent the IHF binding site and the core sequence for site-specific integration of ϕLf, respectively.

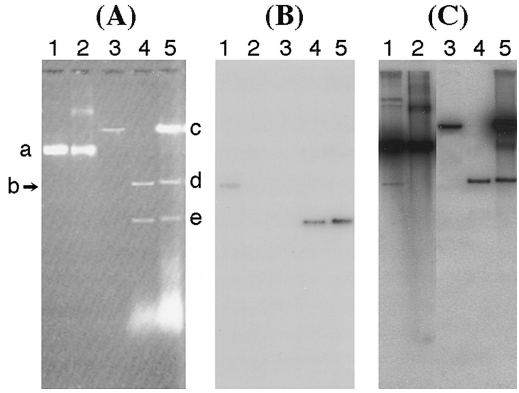


**FIG. 2.** Agarose gel electrophoresis (A) and Southern hybridization of the miniphage DNA using the Km cartridge (B) or the $\phi$Lf RF DNA (C) as a probe. Lanes: 1, ssDNA from the transducing miniphage particles and $\phi$Lf in the culture supernatant of the $\phi$Lf-superinfected Xc17NT1(pOR1); 2, ssDNA from the $\phi$Lf particles; 3 to 5, the *Pst*I-digested RF DNA extracted from Xc17NT1($\phi$Lf), Xc17NT1(pOR1) and $\phi$Lf-superinfected Xc17NT1(pOR1), respectively. Treatment with *Pst*I linearized $\phi$Lf and cleaved at the *Pst*I sites flanking the Km cartridge. Letters: a, $\phi$Lf ssDNA; b, pOR1 ssDNA; c, $\phi$LfRF DNA; d, minireplicon OR1, and e, Km cartridge.

*Nucleotide Sequence of the Minireplicon*

The minireplicon OR1 contained 2028 nucleotides (Fig. 3). The 3′-end of this fragment had an overlap of 340 bp, starting at the *Eco*RV site, with the previously reported sequence of the *Eco*RV-*Sph*I fragment containing genes V, VII, IX and VIII (Fig. 1; 4). Therefore, this overlapping region contained the N-terminus of gene V. After completion of this study, the nucleotide sequence of a total of 2906 bp of the $\phi$Lf RF has been determined.

Computer search of six possible reading frames on the 2028-bp fragment revealed one open reading frame, ORF346. Starting at bp 977 with GTG, ORF346 was able to encode a polypeptide of 39144 dalton. Sequence GAGG, consensus to a ribosome-binding site, was found 6 bp upstream to the GTG (Fig. 3). No sequence similar to a promoter structure was found.

*Comparison of the Sequences*

The deduced amino acid sequence of ORF346 showed 93% identity to that of the ORF344 of the *X. campestris* pv. *citri* filamentous phage cflt, which has been demonstrated to possess integrase activity and shown to have 60% similarity to the amino acid sequence of the f1 gII protein, the replication initiation protein (17). In addition, the deduced protein of ORF346 possessed three regions which show striking similarity to the highly conserved motifs important for the function of replication proteins (Rep) from various sources, which are involved in initiation and termination of rolling circle DNA replication (Fig. 4; 18). Viewing these similarities together with our results showing the function involved in the autonomous replication, we suggest ORF346 to be the replication initiation protein gene, the gII of $\phi$Lf. Compared with ORF346, ORF344 has a deletion of two amino acids, $Gly_{94}$ and $Gln_{95}$, which results in differences in length between the two ORFs.

Gene X (gX) of the Ff phages is contained within gII and has the same amino acid sequence as the C-terminal third of gII (19). A similar case was found in the gII coding region of $\phi$Lf (Fig. 3). This putative gX, starting at bp 1890 with AUG, could encode 165 amino acids and had 48% of the length of gII.

*The Intergenic Region*

In Ff phages, the intergenic region (IR, 508 bp), a region not encoding any protein product, contains two subregions, a core and an enhancer. The core includes the origins for replication

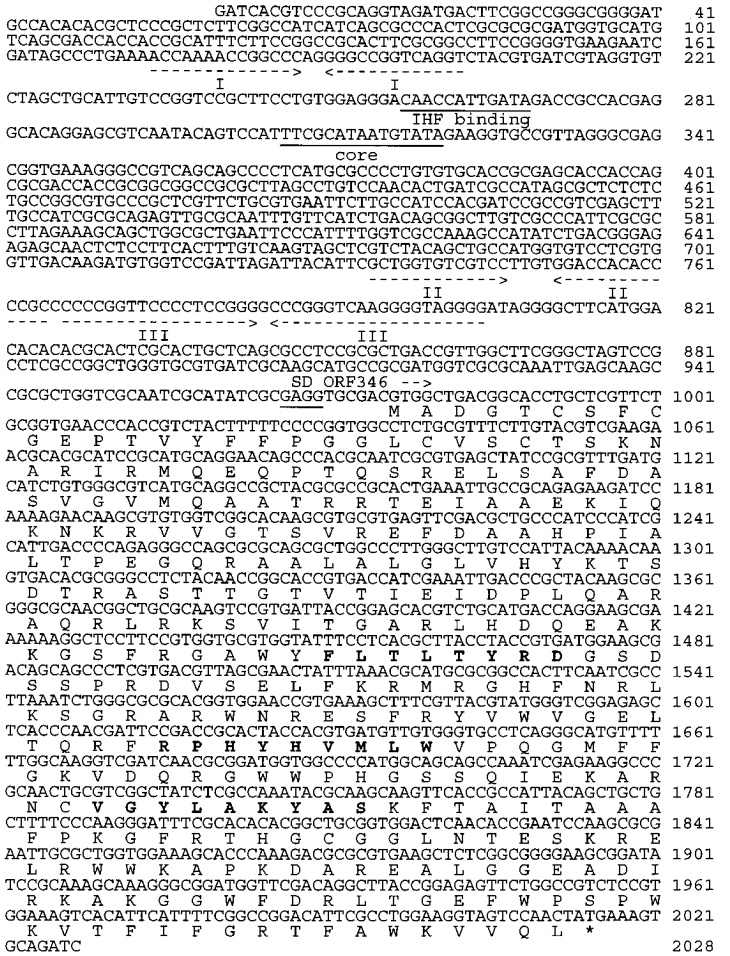


**FIG. 3.** Nucleotide sequence of the minireplicon OR1 (2028 bp) including the intergenic region and ORF346. The coding region is indicated by the amino acids (one-letter code). Underlined are the core, the IHF binding site and the SD sequence. Romantic numerals represent the regions having potential to form hairpin structure. The three regions shown with bold-faced letters represent the conserved motifs found in the Rep proteins from various sources. The nucleotide sequence has been deposited in the Genbank under Accession No. U38235.

initiation of the viral (+) and the complementary (−) strands, and the phage packaging signal, whereas the enhancer is a 150-bp AT-rich sequence including two integration host factor (IHF) binding sites (20). In minireplicon OR1, a stretch of 976 bp upstream to ORF346 contained no identifiable open reading frame (Fig. 3). This 976 bp fragment seemed to contain the IR of $\phi$Lf, since it was sufficient to provide the structures for replication and phage packaging. In the IR of Ff phages, several segments having potential to form hairpin structure are important for normal functions (21). Segments which could form hairpin structure were also found in the 976-bp region (Fig. 3), although the roles they play remain to be studied.

The minireplicon OR1 included most part of the *Pst*I-*Eco*RI fragment (858 bp) which mediates integration of $\phi$Lf into the host chromosome (3). A 15-bp AT-rich sequence within this region (Fig. 3; bp 307), identical to the core for the integration of the *X. c.* pv. *citri* filamentous phage cflt (22), has been shown to serve as the attachment site for $\phi$Lf integration (Lin and Tseng, unpublished results). Thirty-seven bp downstream to the core is a 12 bp sequence similar to the *E. coli* IHF binding site (Fig. 3). It has been established that binding of IHF is required for various functions,

| | 1 | | 2 | | 3 |
|---|---|---|---|---|---|
| Consensus | futltxxx | | xpHuHuuux | | uxxYuxkxxx |
| ϕX174 | FDTLTLAD | 53 | RLHFHAVHF | 70 | VGFYVAKYVN |
| G4 | FDTLTLAD | 53 | RLHFHAVHF | 70 | VGFYVAKYVN |
| SPV4 | FVTLTYSD | 50 | RPHYHICFF | 44 | -ANYTARYTT |
| PHASYL | FLTLTFRD | 37 | RIHYHLLVA | 56 | IGRYVGKYIS |
| ϕLf | FLTLTYRD | 47 | RPHYHVMLW | 29 | VG-YLAKYAS |

**FIG. 4.** Alignment of the conserved sequence motifs in the initiator proteins for rolling circle DNA replication with the ϕLf sequences. Motifs of the Rep proteins from various sources were compiled by Ilyina and Koonin (18), but only several of them are cited here. The letters in upper and lower cases are the amino acid residues conserved in all or at least half of the sequences aligned, respectively. Letter u and x represent a bulky hydrophobic residue (I, L, V, M, F, Y, M) and no consensus in this position, respectively.

including replication and integration of plasmids and phages (23). In Ff phages, two IHF-binding sites near the origins of replication have been shown to enhance DNA replication (20). However, since the IR of ϕLf is involved in both DNA replication and integration, the role played by the IHF-binding site revealed here remains to be elucidated.

*Genome Organization*

Filamentous phages have similar genome organization, i.e., in an order IR-II-X-V-VII-IX-VIII-III-VI-I-IV (1). It has been shown that ϕLf possesses the same gene order V-VII-IX-VII (4). The results obtained in this study showed IR being followed by II, X and V. Thus, within the region already sequenced, ϕLf has the same genome organization as that of other filamentous phages (Fig. 1).

## ACKNOWLEDGMENTS

This work was supported by grant No. NSC83-0203-B005-071 from National Science Council of Republic of China. We thank Dr. C.-N. Sun for reading the manuscript.

## REFERENCES

1. Model, P., and Russel, M. (1988) *in* The Bacteriophages (Calendar, R., , Ed.), Vol. II, pp. 375–455 Plenum Press, New York.
2. Tseng, Y.-H., Lo, M.-C., Lin, K.-C., Pan, C.-C., and Chang, R.-Y. (1990) *J. Gen. Virol.* **71,** 1881–1884.
3. Fu, J.-F., Chang, R.-Y., and Tseng, Y.-H. (1992) *Appl. Microbiol. Biotechnol.* **37,** 225–229.
4. Wen, F.-S., and Tseng, Y.-H. (1994) *J. Gen. Virol.* **75,** 15–22.
5. Mayer, T. F., and Geider, K. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5416–5420.
6. Yang, B.-Y., and Tseng, Y.-H. (1988) *Bot. Bull. Acad. Sinica.* **29,** 93–99.
7. Vieira, J., and Messing, J. (1991) *Gene* **100,** 189–194.
8. Prentki, P., and Krisch, H. M. (1984) *Gene* **29,** 303–313.
9. Miller, J. H. (1972) Experiments in Molecular Genetic Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
10. Eisenstark, A. (1967) *in* Methods in Virology (Maramorosch, K., and Koprowski, H., , Eds.), Vol. I, pp. 449–524. Academic Press, New York.
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
12. Wang, T.-W., and Tseng, Y.-H. (1992) *Lett. Appl. Microbiol.* **14,** 65–68.
13. Yanisch-Peron, C., Vieira, J., and Messing, J. (1985) *Gene* **33,** 103–109.
14. Sanger, F., Nicklen, S., and Coulsin, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
15. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
16. Vieira, J., and Messing, J. (1982) *Gene* **19,** 259–268.
17. Shieh, G.-J., Lin, C.-H., Kuo, J.-L., and Kuo, T.-T. (1995) *Gene* **158,** 73–76.
18. Ilyina, T. V., and Koonin, E. V. (1992) *Nucleic Acids Res.* **20,** 3279–3285.
19. Yen, T. S. B., and Webster, R. E. (1981) *J. Biol. Chem.* **256,** 11259–11265.
20. Greenstein, D., Zinder, N. D., and Horiuchi, K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6262–6266.
21. Zinder, N. D., and Horiuchi, K. (1985) *Microbiol. Rev.* **49,** 101–106.
22. Kuo, T.-T., Chao, Y.-S., Lin, B.-Y., Lin, L.-F., and Feng, T.-Y. (1987) *J. Virol.* **61,** 60–65.
23. Friedman, D. I. (1988) *Cell* **55,** 545–554.