Structural and functional features of cis-acting sequences in the basic replicon of plasmid Collb-P9

Katsunori Tanaka, Hiroshi Sakai, Yoichi Honda, Toshiaki Nakamura, Akihisa Higashi¹ and Tohru Komano*

Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, Kyoto 606 and 'Research and Development Center, DAIKEN Corporation, Okayama 702, Japan

Received March 20, 1992; Accepted April 9, 1992

GenBank accession no. M34837

ABSTRACT

We have structurally and functionally analyzed the ciselements essential for Collb-P9 plasmid DNA replication. The putative oriV region encompassed a region of 172 base pairs (bp) located 152 bp downstream of the repZ gene. A typical dnaA box found in this region proved nonessential for the DNA replication of Collb-P9. The ssi signal of Collb-P9 is a homologue of the G-sites of Rl and R100 plasmids. Deletion of the G-site led to 1.5-fold reduction of the copy number, suggesting that although this G-site is not essential, it is important for efficient Collb-P9 DNA replication. In addition, the Collb-P9 replicon is highly and extensively homologous with the P307 (RepFIC) replicon, and highly homologous with the R100 (RepFIIA) replicon around the G-site region. These facts imply a common ancestry from which the plasmids have evolved.

INTRODUCTION

ColIb-P9 plasmid is a large (93.2 kb) conjugative plasmid belonging to the I_1 (I α) incompatibility group and originally isolated from Shigella sonnei strain P9. All the information for autonomous replication, incompatibility, and copy number control of IncI₁ plasmids (ColIb-P9, and the related plasmids R144 and R64) is clustered within a small region of the genome (1, 2). Hama et al. (3) have recently demonstrated that a 1845-bp ColIb-P9 DNA fragment contains sufficient information for autonomous replication and incompatibility, and predicted that a 300-bp region downstream of the repZ gene has features characteristic of other procaryotic origins and most likely functions during initiation of ColIb-P9 DNA replication. Biochemical and genetic analyses showed that three genes, designated repZ, inc, and repY were present in the 1845-bp segment. The repZ gene encodes a 39-kDa protein, which probably functions as an initiator protein for the ColIb-P9 replicon (3, 4). The inc gene, mapped upstream of repZ, phenotypically governs the incompatibility and encodes a small RNA (inc RNA) with ^a size of about ⁷⁰ bases. This inc RNA acts in trans to repress the expression of repZ and functions to maintain a constant copy number of the ColIb-P9 replicon in host cells $(3, 5)$. The $repZ$ leader sequence that encodes a 29-aminoacid polypeptide, designated as $repY$, acts as a positive regulator for the repZ expression. The expression of repZ is coupled with $repY$ translation, which disrupts a secondary structure sequestering the repZ translation initiation signal (4). These structural and functional features resemble those seen in the regulation of repA expression in an IncFII plasmid RI. Although the regulation of repZ expression has been studied in detail, the mode of action of the RepZ protein and the precise initiation site in ColIb-P9 DNA replication are not clear.

We have shown that ColIb-P9 contains an ssi signal close to the replication origin (6). An ssi signal is defined as a specific nucleotide sequence in the ssDNA template, directing the priming of the complementary DNA strand synthesis. Although ^a nucleotide sequence in an origin of DNA replication is sufficient as ^a cis-acting element to support DNA replication of ^a given replicon both in vivo and in vitro, actual initiation sites on the leading or the lagging strands do not necessarily exist inside the origin sequence (7, 8). In RI, DNA synthesis of the leading strand is initiated at a specific site about 400 bp downstream of oriR (8). This sequence, called the G-site, can function in a singlestranded form as a template for the synthesis of a unique primer by host primase.

Since priming of both the continuous leading strand and the discontinuous lagging strand syntheses takes place on the ssDNA templates, studies on ssi signals are particularly significant to understand the molecular mechanism of initiation events. The ssi signal of ColIb-P9 was situated in the 350-nt HaeIII segment of the 1.8-kb SalI(S7) fragment, nearly 400 bp downstream of the replication origin. Introduction of the ssi signal into a mutant of the filamentous phage M13 lacking ori_c results in restoration of phage growth and RFI DNA synthesis. Furthermore, DNA homology studies showed that the nucleotide sequence of the ssi signal of ColIb-P9 was extremely homologous with those of the ssi signals, G-sites, in R1 and R100.

In this report, we describe the minimal replication origin of ColIb-P9 and the function of the dnaA box in this region by the in vivo replication assay. Furthermore, the function of the Gsite in vegetative ColIb-P9 DNA replication is described.

^{*} To whom correspondence should be addressed

MATERIALS AND METHODS

Bacteria, phages, and plasmids

E. coli strain JM109 (9) was used as ^a host bacterium for M13 bacteriophages and plasmids used in this study. BW86 [dnaG3] leu thyA deoB rpsL ColI^t Δ (chlA-uvrB)] (10) was used as a host strain for ColIb-P9. TG1 was used as a host strain for M13mpl8 phage derivatives in in vitro site-directed mutagenesis. Phage M13mpl8 was used as a cloning vector. Plasmids pBR322, and pHSG396, 398, and 399 (11), chloramphenicol resistance derivatives of pUC 19, were used as cloning vectors. The key plasmids containing the modified origins are as follows. Plasmid pKT527 is a pHSG396 derivative carrying a 3.0-kb Bgll-EcoRI ColIb-P9 fragment and the β -lactamase gene from pBR322. Deletion derivatives of plasmid pKT527 were constructed by nibbling the 3.0-kb ColIb-P9 fragment in pKT527 with exonuclease III after cleavage by BamHI and SacI. Plasmid $pHAZ$ (= $pKT527\Delta 1700$) is one of the pKT527 deletion derivatives which has the repZ gene and lacks the sequence downstream from the nucleotide coordinate 1576 (Fig. 3), and used as a vector in in vivo replication assay as shown in Fig. 1. Plasmid pAC8 is a mini-ColIb plasmid, which was constructed by self-ligation of the largest Pvull fragment of pKT527 , carrying the 3.0-kb BgIII-EcoRI ColIb-P9 fragment and the β -lactamase gene,. Other plasmids are described in the text.

Media

L-Broth was used for growth of the bacterial cells. $2 \times \text{YT}$ was used for the preparation of M13 phages. When necessary, chloramphenicol (30 mg/ml), tetracycline (10 mg/ml), or ampicillin (25 mg/ml) was supplemented.

DNA manipulation

Restriction endonucleases and DNA-modifying enzymes were purchased from Takara Shuzo Co., Ltd., Toyobo Co., Bethesda Research Laboratories, Inc., and Boehringer Mannheim GmbH, and $\lceil \alpha^{-32}P \rceil dCTP$ (400 Ci/mmol) and $\lceil \gamma^{-32}P \rceil dATP$ (3,000 Ci/mmol) were purchased from Amersham Corp. Plasmid DNA was extracted by the alkaline denaturation procedure (12). Transformation was done by the method of Chung et al. (13).

Measurement of the relative copy number of the mini-ColIb plasmid derivatives

Relative copy numbers of the mini-ColIb plasmid derivatives were measured by estimation of the single-cell ampicillin resistance in the strain JM109 (14, 15).

Alteration of a dnaA box by site-directed mutagenesis

Nucleotide substitution was introduced into the typical dnaA box by oligonucleotide-directed mutagenesis (16, 17) using oligonucleotide, 5'-GATTTATGTCGTTAACTGAAA-3', which was synthesized in ^a DNA synthesizer (Applied Biosystem ³⁹⁴ DNA/RNA synthesizer). The sequence of the altered dnaA box was confirmed by nucleotide sequencing (18) .

The in vivo replication assay

The in vivo replication assay was done as described below (Fig. 1). The fragment that contains the repZ gene, the β lactamase gene, truncated origin fragment, and no ColEl-type origin was excised as a PvuII fragment from each of the pHAZ derivatives. The fragment was self-ligated and introduced into

Fig. 1. Schematic representation of the in vivo replication assay for definition of the minimal essential sequence for ColIb-P9 DNA replication. The truncated origin fragments were prepared by exonuclease III digestion of an 1.1-kb $E_{CO}RV$ fragment of ColIb-P9 on the pHSG396 vector.

the strain JM109. The in vivo replication activities of the mini-ColIb plasmids were measured based on the numbers of ampicillin resistant transformants.

RESULTS

Definition of the minimal DNA sequence essential for the initiation of ColIb-P9 DNA replication in vivo

To define the minimal DNA sequence as ^a cis-element essential for initiation of ColIb-P9 DNA replication, we did an in vivo replication assay (Fig. 1). Plasmid pHAZ ($pKT527\Delta1700$) containing the repZ gene was used as a vector. As shown in Fig. 1, the truncated origin fragments were inserted into pHAZ. ColIb-P9 origin plasmids, the miniplasmid derivatives, were constructed by self-ligation of the largest PvuIl fragment that had been excised from pHAZ derivatives. The origin functions were estimated based on the in vivo transformation efficiency of JM109 with the miniplasmid derivatives. As shown in Fig. 2, miniplasmids containing WE, A275R, A541R, A304L, A324L, and A423L could form ampicillin-resistant transformants of the strain JM109. However, the transformation efficiency of the miniplasmid containing A423L was much lower han those of the other five described above. With $\Delta 423L$, lack of the intergenic region containing ^a static DNA bend, which is referred to as ^a CIS-sequence in RI (19), may be responsible for the decrease in the transformation efficiency (see Discussion). Existence and sizes of the plasmid DNAs from ampicillin-resistant transformants were confirmed by agarose gel electrophoresis. These results were similar to those obtained by the *in vivo polA*(ts) assay, in which pHAZ derivatives were introduced into ^a strain, MM386,

Fig. 2. The structures and replication abilities of the truncated origin fragments. Each solid line represents the truncated origin fragment. The shadowed area is the putative oriV region defined in this work. End points of the truncated origin fragments are indicated by nucleotide numbers in the coordinates (Fig. 3). The in vivo replication activity was measured by the miniplasmid assay. Each truncated origin fragment was prepared as described in Fig. 1. The asterisk shows the mutation site of the dnaA box. Two vertical arrows indicate the replacing nucleotides. inc; inc gene that phenotypically governs the incompatibility. repY; repY gene that encodes a 29-amino-acid polypeptide functioning as a positive regulator for repZ expression. repZ; repZ gene that encodes an initiator protein, RepZ. dnaA box; DnaA protein binding consensus sequence. G-site; G4-type ssi signal.

Fig. 3. Nucleotide sequence of the putative oriV region of ColIb-P9. Bases are numbered starting from the BgIII site of the physical map shown in Fig. 2. Large boxed region represents the 172-bp oriV sequence. A small box from 1653 to 1661 and a dashed box from 1716 to 1729 represent bp of DnaA protein recognition sequence (dnaA box) and an IHF binding site-like tract, respectively. The coding frame for repZ terminates at TAG sequence (1484-1486). The arrows represent three direct repeats of AANCNGCAAAC. Some possible stem-loop structures are indicated with two arrows pointing to each other. The end points of truncated origin fragments are indicated by L-shaped arrows.

Fig. 4. Physical map and relative copy numbers of mini-ColIb derivatives lacking the G-site. The upper thick solid line represents a 3.0-kb BglII-EcoRI fragment of ColIb-P9. Amp^r indicates β -lactamase gene derived from pBR322. Mini-ColIb deletion derivatives were constructed by exonuclease III digestion of pKT527 following self-ligation of the largest PvuII fragment. The deleted parts of the 3.0-kb BgIII-EcoRI fragment are represented by dashed lines. Relative copy numbers were estimated by single-cell ampicillin resistance levels as described in *Materials and Methods*. + or - represents the presence or absence of the G-site. Restriction sites: B, BglII; E, EcoRI; S, SalI; P, PstI; and V, EcoRV.

Fig. 5. Schematic representation of sequence homologies among ColIb-P9, R100, and P307 replicons. The three replicons are aligned for comparison. HRI, HRII, and HRUI represent regions homologous among plasmid replicons belonging to the RepFIIA family. The asterisks represent highly homologous regions (90-100%) between ColIb-P9 and R100, and between ColIb-P9 and P307.

that produces temperature-sensitive DNA polymerase ^I (data not shown). The minimal essential region may be no shorter than 120 bp because deletion of an additional 47 bp of the 5'-end (positions 1763) lost the replication function (3).

All of the above results showed that the minimal origin of ColIb-P9 was present within a region of 172 bp (positions $1639-1810$), 152 bp downstream of the *repZ* gene (Fig. 3). The sequence of this 172-bp region was similar to the other procaryotic origins (20) in its characteristic structural features including some repetitive sequences to which initiator protein RepZ probably binds, a typical dnaA box, and AT-rich sequences containing an IHF binding site-like tract (Fig. 3).

To test whether this *dnaA* box is actually functional for ColIb-P9 DNA replication or not, we performed ^a mutational analysis. A consensus sequence for dnaA box is ⁵'-(T/C)(T/C)(A/T/C)T- $(A/C)C(A/G)(A/C)T(A/C)-3'$ (21). By site-directed mutagenesis, the dnaA box sequence, 5'-TTATCCACA-3', was changed to 5'-TTAACGACA-3', in which the two absolutely conserved nucleotides T and C were replaced by A and G, respectively. These changes (AC8AM) had no effect on the origin function in the miniplasmid assay (Fig. 2), indicating that this $dnaA$ box was not important as a typical dnaA box for ColIb-P9 DNA replication.

Function of the G-site in the vegetative DNA replication of ColIb-P9

Plasmid ColIb-P9 contains, close to the replication origin, an ssi signal, which is the ColIb-P9 homologue of the G-sites, G4-type ssi signals, of R1 and R100 plasmids (6). A mini-ColIb plasmid pAC8 and its derivatives lacking the G-site were constructed (Fig. 4). To analyze the functional role of the G-site in vivo in the vegetative DNA replication of ColIb-P9, we estimated the relative copy numbers of the mini-plasmids. While a mini-plasmid derivative pAC8A650 that has the G-site was maintained at the same copy number as pAC8, deletion of the G-site led to 1.5-fold reductions in the copy numbers as seen with pAC8A1000 and pAC8Al100 (Fig. 4 and Fig. 5). These results suggested that although the G-site was neither a part of the replication origin nor essential for ColIb-P9 DNA replication, it was important for efficient ColIb-P9 DNA replication. We can not rule out, however, the possibility that the deletions introduced into pAC8 result in alterations in the level of Amp^R expression, leading to perturbation in estimating the plasmid copy numbers.

Sequence homologies with IncFI and IncFII plasmids

An ssi signal of ColIb-P9 is the homologue of the G-sites of RI and R100 (6). This fact suggests some evolutionary relationship among these plasmids. To make the evolutionary relationship clearer, comparison between the ColIb-P9 replicon and the R100 replicon was done by computer-assisted homology research. Plasmid R100 belongs to the IncFII group and its replicon is referred to as the RepFIIA replicon. There was a highly conserved region about 600 bp long including the G-site (Fig. 6). Furthermore, the nucleotide sequence in a region about 2.5 kb long was extensively and highly conserved between the ColIb-P9 replicon and the P307 replicon (Fig. 6). Plasmid P307 belongs to the IncFI group and its replicon is referred to as the RepFIC replicon, which is one of the three replicons described for IncF plasmids. The repZ gene, the minimal replication origin, and the G-site of ColIb-P9 were within this homologous region. These results suggest that the ColIb-P9, the P307, and the R100 replicons have a similar replication mechanism and are evolutionarily related, although these three are grouped into the distinct replicons at the present time.

DISCUSSION

In this study, we have analyzed the structure and function of the cis-element essential for ColIb-P9 DNA replication. The putative $oriV$, a minimal nucleotide sequence as a *cis*-acting element essential for ColIb-P9 DNA replication, has been identified. It was 172 bp long and contained a typical *dnaA* box, AT-rich sequences, repetitive sequences, and an IHF binding site-like tract. We showed that the G-site, which was located downstream of this putative $oriV$ region, was an essential element for efficient ColIb-P9 DNA replication. Furthermore, judging from the finding that the nucleotide sequence near the replication origin was highly conserved among plasmids ColIb-P9, P307, and R100, we suggested that these three plasmids had evolutionarily close relationships.

Although the minimal cis-element essential for ColIb-P9 DNA replication was a 172-bp region downstream of the repZ gene, the replication ability of $\Delta 423L$ was partially defective as evidenced in the in vivo replication assay. This may be because \triangle 423L lacks a part of the intergenic region between repZ and putative $oriV$ which is the ColIb-P9 homologue of the CISsequence of RI. With RI and its related plasmid NR1, the CISsequence contains a static DNA bend and a Rho-dependent type of DNA rearrangement mediated by a 'shufflon' (31).

transcription terminator, and appears to impart the cis-acting behavior to the Rep protein (19, 22). Our preliminary results suggested that the RepZ protein acted *in vivo* on the replication origin preferentially in cis (data not shown). Thus, the intergenic region between $repZ$ and putative $oriV$ may be crucial in mediating the cis-dominant action of the RepZ protein in DNA replication as proposed in the RI replication. The initial stage of ColIb-P9 DNA replication probably involves specific recognition of the $oriV$ sequence by the RepZ protein, but the nature of the interaction and the mechanism of the cis action are not known.

Mutational analysis of the *dnaA* box in the putative $\text{orif } V$ region of ColIb-P9 suggested that it did not play the typical role assigned to a *dnaA* box (Fig. 2). Outside the putative *oriV*, there is a 9-bp sequence (5'-TTATCCAGA-3', 1088 - 1080) which contains one substitution in the *dnaA* box consensus, and overlaps with repZ. Because the 9-bp sequence may compensate for the function of the *dnaA* box in the putative *oriV*, we can not rule out the possibility that DnaA protein is involved in the initiation of plasmid replication. With RI, the evidence for involvement of DnaA protein in the DNA replication includes inconsistencies. Some independent reports (23, 24, 25, 26, 27) suggest that DnaA protein is important in Rl DNA replication, but the degree to which DnaA protein is necessary may depend on some unidentified factors.

In plasmid RI, the G-site is recognized by primase to direct the initiation of the leading strand synthesis and is non-essential but necessary for efficient DNA replication (8). In the plasmids RIOO, F, and R6K, regions of sequence homology with the Gsite of RI have been found and confirmed their ssi functions, indicating the generality of primase recognition sequences (28). We suggested that this G-site is important in efficient replication and is able to direct the initiation of the leading strand synthesis as in the case of RI and R100. In the absence of this G-site, an alternative priming mechanism may function in vivo to rescue ColIb-P9 DNA replication.

Recent studies have indicated that there is extensive homology among basic replicons of the IncF plasmids (29), suggesting a common ancestry. Replicons homologous with RepFIIA replicon are classified as the RepFIIA family (30). The RepFIIA family thus far studied contains three well conserved regions, HRI, HRII, and HRIII, which are separated by two non-conserved regions, NHRI and NHRH. While the NHRs contain the genes encoding plasmid-specific proteins, the HRs contain the regulator element for the expression of the genes. In this study, we found that the ColIb-P9 replicon was highly homologous with the P307 (RepFIC) and partially with the RIOO (RepFIIA) replicon (Fig. 6). These findings suggest that the ColIb-P9 replicon and the P307 replicon, although they have different incompatibility properties based on nonhomologous regions, have a similar replication mechanism and have evolutionarily close relationships to each other. Moreover, considering the report that the P307 (RepFIC) replicon and the R100 (RepFIIA) replicon can be classified as the RepFIIA family (30), the ColIb-P9 replicon has evolutionarily close relationships not only to the P307 replicon but also to the RIOO replicon. Taken together the results presented herein strongly suggest that the ColIb-P9 replicon, an IncI₁ group plasmid, can be classified in the RepFIIA family. Thus, the divergence among the ColIb-P9, the P307 (RepFIC), and the RI00 (RepFIIA) replicons is more likely to have occurred through mechanisms such as homologous recombination, or as a novel

ACKNOWLEDGEMENTS

We thank Dr. K.Mizobuchi for providing the sequence information before publication. This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan.

REFERENCES

- 1. Furuichi, T., Komano, T. and Nisioka, T. (1984) J. Bacteriol, 158, $997 - 1004$.
- 2. Rees, C. E. D., Bradley, D. E. and Wilkins, B. M. (1987) Plasmid, 18, $223 - 236$.
- 3. Hama, C., Takizawa, T., Moriwaki, H., Urasaki, Y. and Mizobuchi, K. (1990) J. Bacteriol., 172, 1983-1991.
- 4. Hama, C., Takizawa, T., Moriwaki, H. and Mizobuchi, K. (1990) J. Biol. Chem., 265, 10666-10673.
- 5. Shiba, K. and Mizobuchi, K. (1990) J. Bacteniol, 172, 1992-1997.
- 6. Tanaka, K., Sakai, T., Honda, Y., Hiasa, H., Sakai, H. and Komano, T. (1991) Plasmid, 25, 125-130.
- 7. Minden, J. S. and Marians, K. J. (1985) J. Biol. Chem., 260, 9316-9325.
- 8. Masai, H. and Arai, K. (1989) J. Biol. Chem., 264, 8082-8090.
- 9. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene, 33, 103-119. 10. Chatfield, L. K., Orr, E., Boulnois, G. J. and Wilkins, B. M. (1982) J. Bacteriol, 152, 1188-1195.
- 11. Takeshita, S., Sato, M., Toba, M., Masahashi, W. and Hashimoto-Gotoh, T. (1987) Gene, 61, 63-74.
- 12. Bimboim, H. C. and Doly, J. (1979) Nucleic Acids Res., 7, 1513- 1523.
- 13. Chung, C. T., Niemela, S. L. and Miller, R. H. (1989) Proc. Natl. Acad. Sci. USA, 86, 2172-2175.
- 14. Honda, Y., Sakai, H., Hiasa, H., Tanaka, K., Komano, T. and Bagdasarian, M. (1991) Proc. Natl. Acad. Sci. USA, 88, 179-183.
- 15. Nordström, K., Molin, S. and Aagaard-Hansen, H. (1980) Plasmid, 4, $332 - 349$.
- 16. Taylor, J. W., Schmidt, W., Cosstide, R., Okuruszek, A. and Eckstein, F. (1985) Nucleic Acids Res., 13, 8749-8764.
- 17. Taylor, J. W., Ott, J. and Eckstein, F. (1985) Nucleic Acids Res., 13, 8765-8785.
- 18. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
- 19. Masai, H. and Arai, K.-i. (1988) Nucleic Acids Res., 16,
- 20. Bramhill, D. and Kornberg, A. (1988) Cell, 54, 915-918.
- 21. Schaefer, C. and Messer, W. (1991) Mol. Gen. Genet., 226, 34-40. 22. Dong, X., Womble, D. D. and Rownd, R. H. (1988) J. Mol. Biol., 202, 495-509.
- 23. Molin, S. and Nordström, K. (1980) J. Bacteriol, 141 , $111 -$.
- 24. Nagata, T., Murakami, Y. and Imai, M. (1988) Mol. Gen. Genet., 213, 163.
- 25. Tang, X. B., Womble, D. D. and Rownd, R. H. (1989) J. Bacteriol, 171, 5290.
- 26. Ortega, S., Lanka, E. and Diaz, R. (1986) Nucleic Acids Res., 14, 4865-.
- 27. Masai, H. and Arai, K.-i. (1987) Proc. Natl. Acad. Sci. USA, 84, 4781-4785.
- 28. Nomura, N., Masai, H., Inuzuka, M., Miyazaki, C., Ohtsubo, E., Itoh, T., Sasamoto, S., Matsui, M., Ishizaki, R. and Arai, K.-i. (1991) Gene, 108, 15-22.
- 29. Bergquist, P. L., Saadi, S. and Maas, W. K. (1986) Plasmid, 15, 19-34.
- 30. Saadi, S., Maas, W. K., Hill, D. F. and Bergquist, P. L. (1987) J. Bacteriol., 169, 1836-1846.
- 31. Komano, T., Kubo, A., Kayamuma, T., Furuichi, T. and Nisioka, T. (1986) J. Bacteriol., 165, 94-100.