Molecular mechanisms of early steps of ColE2 plasmid replication initiation

2007 March

Han Man

Contents

Title	1
Contents	2-3
Chapter I. Introduction	4-9
References	10-13
Chapter II. The ColE2 Rep protein binds to the origin DNA as a monomer	14-28
Abstract	15
Introduction	16
Materials and methods	17-19
Results and discussion	20-22
Figure legends	23
References	24-26
Figures	27-28
Chapter III. Characterization of Specific Interaction between the Initiator Protein (Rep) andOrigin of Plasmid ColE2-P9	29-70
Abstract	30
Introduction	31-33
Materials and methods	34-38
Results	39-45
Discussion	46-49
References	50-53
Figure legends	54-57
Tables and Figures	58-70

Chapter IV. Identification of the region important for dimerization of Rep protein 71-78
Abstract 72
Materials and methods 73
Results and discussion 74
References 75
Figure legends 76
Figures 77-78

Chapter V. Mechanism of origin melting of ColE2-P9 plasmid 79-107
Abstract 80
Introduction 81-82
Materials and methods 83-85
Results and discussion 86-92
Figure legends 93-95
References 96-99
Figures 100-107

Chapter VI. Acknowledgments 108-109

Chapter VII. Publications 110-111

I Introduction

DNA replication is a universal process for transmission of genetic information in all organisms. In the classic replicon model (18), regulated DNA replication requires a cis-acting DNA sequence (the replicator) and a trans-acting factor (the initiator) that binds to the replicator, which is now called the origin (of replication), thus resembling the proposed made shortly before for the regulatory circuits in gene expression. For the last 40 years, the replicon has consistently received multiple experimental support (28). DNA replication can be divided into three stages: initiation (8, 14, 28), where initiation proteins bind to the origin, locally unwind the duplex, and introduce other replication proteins like the helicase, primase and DNA polymerase into the open complex to form the replisome; elongation, where DNA synthesis take place; and termination, where replication forks collide and and concatenated DNA molecules are separated into two daughter molecules.

In bacteria the three phases are relatively well understood. Among various bacterias, DNA replication of *Escherichia coli* is the best characterized (19). ATP-bound DnaA (initiator) binds as a monomer specifically to each of the four asymmetric 9-bp consensus sequences (5'-TTA/TTNCACA) or its closely related sequences, the DnaA boxes in the origin (OriC). These ATP-bound DnaA proteins oligomerize with additional ATP-bound DnaA proteins to form a large nucleoprotein structure containing 20-40 DnaA proteins, which induces local unwinding of the AT-rich region adjacent to OriC. The DnaA proteins act as a replisome organizer for the subsequent loading of proteins required for the replication process (32). A complex of the helicase loader, the DnaC protein, together with helicase, the DnaB proteins, is associated with the DnaA-origin complex (7, 30). Following ATP hydrolysis by DnaC, DnaB is assembled around the DNA at the origin (30).Then primase, the DnaG proteins, is associated with the DnaB protein to synthesis the primer, which is utilized by DNA polymerase III to initiate bidirectional DNA synthesis.

The study of DNA replication initiation in eukaryotic replicons initially followed path slower than in bacteria. Eukaryal linear chromosomes, on the contrary to the circular bacterial one, present multiple origins of replication to allow copying the complete genome timely along the S-phase of the cell cycle. The existence of many replication origins introduces further complexity in the mechanisms that control eukaryotic DNA replication (5, 13, 29). A complex of six proteins called ORC (origin recognition complex, orc 1-6), specifically binding to replicators in an ATP-dependent

way (6, 25) was first characterized in *S. cerevisiae*.

In *S. cerevisiae* all six ORC subunits are essential for viability and remain associated with ARS through all the cell cycle (2, 11, 44). Yeast ORC seems to include a single copy of each subunit (4) but subcomplexws defective in Orc1 (31) and Orc6 (36) have been isolated in higher eukaryotes, together with others including extra copies of some subunits (42), being proposed to control DNA replication and cell cycle progression. Thus those two subunits would be associated in a looser way with an Orc2-5 core (10, 43). Local melting of the origin occurs either by ORC alone or in conjunction with additional factors. Either before or during its association with Cdc6, local melting occurs. ATP-bound Cdc6 associates with ORC and the MCM helicase at the origin. ATP hydrolysis by Cdc6 causes conformational changes that release the helicase. Cdt1p is likely to be important for loading of MCM helicase. (33). Initiating the replication from pre-RC (pre replication complex) requires the activity of Cyclin-Dependent Kinases (CDKs). After the helicase is loaded onto the DNA, RPA coats the exposed ssDNA. Then the rest of the replication machinery like PCNA clamp and DNA pol α is recruited to initiate DNA synthesis (1).

Early studies suggested that archaeal replication is a simpler version of the eukaryotic process, as many of the proteins participating in archaeal DNA replication appear to be more similar to those found in eukarya than those in bacteria (12, 35). It became apparent, however, that most of the archaeal replication proteins are eukaryotic-like, some are more similar to those found in bacteria and others are archaeal-specific factors (polD). Initiation of replication in archaea begins when Cdc6/Orc1 homologue (the likely archaeal OBP, might exert all the functions described for eukaryal ORC and Odc6 OBPs) recognizes and binds to the origin. The binding causes a distortion and localized melting in the DNA. ATP-bound Cdc6/Orc1 associates with the Cdc6/Orc1-origin complex and with the MCM helicase. Following ATP hydrolysis the Cdc6/Orc1 protein releases the helicase MCM. After the helicase is loaded onto the DNA, SSB (Ssb-like protein) coats the exposed ssDNA. This SSB-ssDNA complex is the substrate onto which primase, the DNA polymerase, and the rest of the replication machinery is recruited to initiate DNA synthesis (20).

Plasmids are double-stranded circular or linear extrachromosomal DNA molecules. Although plasmids replicate autonomously, they coexist stably with a host by limiting their replication to one round per bacterial generation. Circular plasmids replicate

primarily by either asymmetric rolling circle (RC) replication or theta replication. In RC replication of plasmids, a plasmid-encoded initiator protein (Rep) intiates replication by making a strand-specific nick at the plasmid double-stranded origin (DSO), which then serves as a primer for replication (26, 41). RC replication is the most common mode of replication for small (<10 kb), promiscuous, multicopy Gram-positive bacterial plasmids, and is also used by some Gram-negative bacterial plasmids (9, 21). Theta replication is initiated by RNA primers at the origin (21, 34). The overall structure of origins for replicons that replicate by the theta-type mechanism and that are independent on DNA Pol I is highly similar to the *E.coli* OriC origin. Typically, these plasmid origins consist of a set of short repeated DNA sequences (iterons), specifically recognized by the initiator Rep protein; One or two binding sites for E.coli DnaA initiator protein (*dnaA* boxes); an A+T rich sequence adjacent to the iterons which often contains repeats similar to the 13-mer repeats of the E.coli OriC; and commonly GATC sequences which are the sites of methylation by host Dam methylase. Initiation of replication follows in general the model proposed for replication at OriC (28). The initiator Rep protein specifically recognizes and binds to the iterons of the origin and various degrees of assistance of the DnaA protein, alters the DNA conformation and initiates a localized melting of the strands at the adjacent A+T rich region, and successive loading of the replisomal helicase, primase, and the DNA Pol III HE (8, 22). In iteron replicons, F plasmid is well characterized. The crystal structure of the RepE initiator protein of plasmid F was solved as the first three-dimensional structure of a prokaryotic initiator protein (27). The RepE protein consists of topologically similar N- and C- terminal domains related to each other by internal pseudo 2-fold symmetry. By using HTH DNA binding motifs, the two domains bind to two consecutive major grooves on one face of the DNA helix of each of the 19-bp directly repeated sequences (the iterons) in the origin. The amino acid sequence homology between the RepE protein and some other plasmid initiator proteins, such as those of pPS10 and pSC101 suggested that those proteins might bind to their iterons similarly (27). Although the P1 RepA protein can not be aligned due to its poor sequence homology with the RepE protein, the molecular modeling suggested that it is also structurally similar to the class of plasmid initiator proteins to which the RepE protein belongs (37). In Pol I-dependent plasmids like ColE1, ColE2, pAMβ1, replisome assembly occurs after an initial synthesis of the leading strand carried out by Pol I. This synthesis opens the duplex and

activates a PriA-dependent primosome assembly site which attracts the helicase, primase, and Pol III HE.

The plasmid ColE2-P9 (ColE2) is a circular duplex DNA molecule of about 7 kb (15). It is present at about 10 to 15 copies per host chromosome (3, 16). The initiator protein (Rep; 35 kDa) of plasmid ColE2 is the only plasmid-specified *trans*-acting factor required for the initiation of plasmid replication (23, 24, 46). Initiation of plasmid replication also requires host DNA polymerase I (17, 38), but not RNA polymerase and DnaG primase (17, 39). The minimal *cis*-acting region, the origin (Ori; Fig. 1A), required for ColE2 DNA replication consists of 31 bp (45). The ColE2 Rep protein binds specifically to Ori DNA, as revealed by a filter binding assay (23) and by an electrophoresis mobility shift assay (EMSA) (45). In an *in vitro* ColE2 DNA replication system with crude extracts of *Escherichia coli*, ColE2 DNA replication starts at a unique site in Ori and proceeds unidirectionally (17, 39). The ColE2 Rep protein is unique among other initiator proteins in that it is also a plasmid-specific primase. It synthesizes a 3-nucleotide primer RNA which has a unique structure of 5' ppApGpA (40). Host DNA polymerase I specifically uses the primer RNA to start DNA synthesis, and then form a D-loop structure, into which various replication proteins of *E. coli* like DnaB helicase, DnaG primase and DNA polymerase III holoenzyme, are introduced to continue replication of ColE2 DNA.

The replication initiator protein (Rep) of plasmid ColE2-P9 (ColE2) is multifunctional. We are interested in how Rep binds to the origin (Ori) to perform various functions. Detailed analysis of the binding property of ColE2 Rep and Ori is an important prerequisite to investigate the following steps after Rep binding to Ori, such as the formation of the open complex and primer synthesis. In this study, I used the electrophoresis mobility shift assay (EMSA) and dimethyl sulfate (DMS) protection assay to characterize the interaction of Rep and Ori. Three regions in Rep involved in Ori binding and corresponding three binding sites in Ori are identified. A model for the mode of the ColE2 Rep-Ori binding is proposed (chapter III). The ColE2 Rep proteins are present mostly in a dimeric form with some multimers larger than dimers in solution, while the form of Rep binding to Ori is not known. The results in this study demonstrate that, Rep binds to Ori as a monomer I propose here the efficient dimerization of Rep might be involved in keeping the copy number of the ColE2 plasmid at the normal level by limiting the amount of the active monomeric form of Rep in the cell (chapter II). I

further identified the region important for dimerization of Rep. It is located within the area containing regions II and III important for Rep-Ori interaction. This raises a possibility that dimerization of Rep affected the interaction of regions II and III of Rep with Ori and this might explain why Rep binds to Ori as a monomer but not as a dimer (chapter IV). Rep is also a plasmid-specific primase. Synthesis of the primer by the ColE2 Rep protein requires opening of the DNA in the origin to expose the template strand. I examined how the interaction of ColE2 Rep and Ori leads to opening of the origin DNA. Bases on my research, I propose a model for the molecular mechanism of early steps of ColE2 plasmid replication initiation (chapter V). My research should provide insights into the mechanism of initiation of ColE2 DNA replication.

.

References

1. **Aparicio, O. M., A. M. Stout, and S. P. Bell.** 1999. Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. Proc Natl Acad Sci U S A **96:**9130-5.
2. **Aparicio, O. M., D. M. Weinstein, and S. P. Bell.** 1997. Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell **91:**59-69.
3. **Bazaral, M., and D. R. Helinski.** 1968. Circular DNA forms of colicinogenic factors E1, E2 and E3 from Escherichia coli. J Mol Biol **36:**185-94.
4. **Bell, S. P.** 2002. The origin recognition complex: from simple origins to complex functions. Genes Dev **16:**659-72.
5. **Bell, S. P., and A. Dutta.** 2002. DNA replication in eukaryotic cells. Annu Rev Biochem **71:**333-74.
6. **Bell, S. P., and B. Stillman.** 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature **357:**128-34.
7. **Davey, M. J., L. Fang, P. McInerney, R. E. Georgescu, and M. O'Donnell.** 2002. The DnaC helicase loader is a dual ATP/ADP switch protein. Embo J **21:**3148-59.
8. **del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas.** 1998. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev **62:**434-64.
9. **del Solar, G., M. Moscoso, and M. Espinosa.** 1993. Rolling circle-replicating plasmids from gram-positive and gram-negative bacteria: a wall falls. Mol Microbiol **8:**789-96.
10. **Dhar, S. K., L. Delmolino, and A. Dutta.** 2001. Architecture of the human origin recognition complex. J Biol Chem **276:**29067-71.
11. **Diffley, J. F., J. H. Cocker, S. J. Dowell, and A. Rowley.** 1994. Two steps in the assembly of complexes at yeast replication origins in vivo. Cell **78:**303-16.
12. **Edgell, D. R., and W. F. Doolittle.** 1997. Archaea and the origin(s) of DNA replication proteins. Cell **89:**995-8.
13. **Gilbert, D. M.** 2001. Making sense of eukaryotic DNA replication origins. Science **294:**96-100.
14. **Helinski, D. R., A.E. Toukdarian, and R.P. Novick.** 1996. Replication control

and other stable maintenance mechanism of plasmids, p. 2295-2324. *In* F.C. Neidhardt, Curtiss, R.,III, J. Ingraham, E.C.C.Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (eds), Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington D.C.

15. **Herschman, H. R., and D. R. Helinski.** 1967. Comparative study of the events associated with colicin induction. J Bacteriol **94:**691-9.
16. **Horii, T., and T. Itoh.** 1988. Replication of ColE2 and ColE3 plasmids: the regions sufficient for autonomous replication. Mol Gen Genet **212:**225-31.
17. **Itoh, T., and T. Horii.** 1989. Replication of ColE2 and ColE3 plasmids: in vitro replication dependent on plasmid-coded proteins. Mol Gen Genet **219:**249-55.
18. **Jacob, F., S. Brenner, and F. Couzin,.** 1963. On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol **28:**329-348.
19. **Kaguni, J. M.** 1997. Escherichia coli DnaA protein: the replication initiator. Mol Cells **7:**145-57.
20. **Kelman, L. M., and Z. Kelman.** 2003. Archaea: an archetype for replication initiation studies? Mol Microbiol **48:**605-15.
21. **Khan, S. A.** 1996. in Genetic Engineering (vol. 18) (Setlow, J.K., ed.), pp. 183-201, Plenum Press.
22. **Khan, S. A., and D. K. Chattoraj.** 1998. Initiation of DNA replication in phages and plasmids-a workshop summary. Plasmid **40:**1-11.
23. **Kido, M., H. Yasueda, and T. Itoh.** 1991. Identification of a plasmid-coded protein required for initiation of ColE2 DNA replication. Nucleic Acids Res **19:**2875-80.
24. **Kingsbury, D. T., and D. R. Helinski.** 1970. DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem Biophys Res Commun **41:**1538-44.
25. **Klemm, R. D., R. J. Austin, and S. P. Bell.** 1997. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. Cell **88:**493-502.
26. **Koepsel, R. R., R. W. Murray, W. D. Rosenblum, and S. A. Khan.** 1985. The replication initiator protein of plasmid pT181 has sequence-specific

endonuclease and topoisomerase-like activities. Proc Natl Acad Sci U S A **82:**6845-9.

27. **Komori, H., F. Matsunaga, Y. Higuchi, M. Ishiai, C. Wada, and K. Miki.** 1999. Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 A resolution. Embo J **18:**4597-607.

28. **Kornberg, A., and T. Baker.** 1992. DNA replicarion. 2nd edn. W.H.Freeman, New York, NY.

29. **Lei, M., and B. K. Tye.** 2001. Initiating DNA synthesis: from recruiting to activating the MCM complex. J Cell Sci **114:**1447-54.

30. **Marszalek, J., and J. M. Kaguni.** 1994. DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli. J Biol Chem **269:**4883-90.

31. **Mendez, J., X. H. Zou-Yang, S. Y. Kim, M. Hidaka, W. P. Tansey, and B. Stillman.** 2002. Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. Mol Cell **9:**481-91.

32. **Messer, W. a. C. W.** 1996. Initiation of chromosome replication, p. 1579-1601. *In* F.C. Neidhardt, Curtiss, R.,III, J. Ingraham, E.C.C.Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (eds), Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington D.C.

33. **Nishitani, H., Z. Lygerou, T. Nishimoto, and P. Nurse.** 2000. The Cdt1 protein is required to license DNA for replication in fission yeast. Nature **404:**625-8.

34. **Novick, R. P.** 1989. Staphylococcal plasmids and their replication. Annu Rev Microbiol **43:**537-65.

35. **Olsen, G. J., and C. R. Woese.** 1997. Archaeal genomics: an overview. Cell **89:**991-4.

36. **Prasanth, S. G., K. V. Prasanth, and B. Stillman.** 2002. Orc6 involved in DNA replication, chromosome segregation, and cytokinesis. Science **297:**1026-31.

37. **Sharma, S., B. K. Sathyanarayana, J. G. Bird, J. R. Hoskins, B. Lee, and S. Wickner.** 2004. Plasmid P1 RepA is homologous to the F plasmid RepE class of initiators. J Biol Chem **279:**6027-34.

38. **Tacon, W., and D. Sherratt.** 1976. ColE plasmid replication in DNA polymerase I-deficient strains of Escherichia coli. Mol Gen Genet **147:**331-5.

39. **Takechi, S., and T. Itoh.** 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. Nucleic Acids Res **23:**4196-201.

40. **Takechi, S., H. Matsui, and T. Itoh.** 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. Embo J **14:**5141-7.

41. **Thomas, C. D., D. F. Balson, and W. V. Shaw.** 1990. In vitro studies of the initiation of staphylococcal plasmid replication. Specificity of RepD for its origin (oriD) and characterization of the Rep-ori tyrosyl ester intermediate. J Biol Chem **265:**5519-30.

42. **Thome, K. C., S. K. Dhar, D. G. Quintana, L. Delmolino, A. Shahsafaei, and A. Dutta.** 2000. Subsets of human origin recognition complex (ORC) subunits are expressed in non-proliferating cells and associate with non-ORC proteins. J Biol Chem **275:**35233-41.

43. **Vashee, S., P. Simancek, M. D. Challberg, and T. J. Kelly.** 2001. Assembly of the human origin recognition complex. J Biol Chem **276:**26666-73.

44. **Wyrick, J. J., J. G. Aparicio, T. Chen, J. D. Barnett, E. G. Jennings, R. A. Young, S. P. Bell, and O. M. Aparicio.** 2001. Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science **294:**2357-60.

45. **Yagura, M., S. Y. Nishio, H. Kurozumi, C. F. Wang, and T. Itoh.** 2006. Anatomy of the replication origin of plasmid ColE2-P9. J Bacteriol **188:**999-1010.

46. **Yasueda, H., T. Horii, and T. Itoh.** 1989. Structural and functional organization of ColE2 and ColE3 replicons. Mol Gen Genet **215:**209-16.

II The ColE2-P9 Rep Protein Binds to The Origin DNA As a Monomer

Abstract

The Rep proteins of some plasmid replicons have two functions. Dimers bind to the operator sequences acting as auto-repressors, whereas monomers bind to the iterons to initiate replication of DNA. The ColE2 Rep proteins are present mostly in a dimeric form with some multimers larger than dimers in solution, while the form of Rep binding to Ori is not known. We used an EMSA-based method to determine the molecular weight of Rep in the Rep-Ori complex. The result suggested that Rep binds to Ori as a monomer. In addition, the result of EMSA using the Rep protein fused with the maltose binding protein and the His_6-tag also supported this conclusion. We proposed that dimerization of Rep might probably be involved in keeping the copy number of the ColE2 plasmid at the normal low level by limiting the amount of active monomeric forms of Rep in the host cell.

Introduction

The plasmid ColE2-P9 (ColE2) is a circular duplex DNA molecule of about 7 kb [1]. It is present at about 10 to 15 copies per host chromosome [2, 3]. The initiator protein (Rep; 35 K) of plasmid ColE2 is the only plasmid-specified *trans*-acting factor required for the initiation of plasmid replication [4-6]. The initiation of plasmid replication also requires host DNA polymerase I [7, 8], but not RNA polymerase and DnaG primase [8, 9]. The minimal *cis*-acting region, the origin (Ori; Fig. 1A), required for ColE2 DNA replication consists of 31 bp [10]. The ColE2 Rep protein binds specifically to Ori DNA, as revealed by a filter binding assay [5] and by an electrophoresis mobility shift assay (EMSA) [10]. In an *in vitro* ColE2 DNA replication system with crude extracts of *Escherichia coli* cells, ColE2 DNA replication starts at a unique site in Ori and proceeds unidirectionally [8] [9]. The ColE2 Rep protein is unique among other initiator proteins in that it is also a plasmid-specific primase. It synthesizes a 3-nucleotide primer RNA which has a unique structure of 5' ppApGpA [11]. Host DNA polymerase I specifically uses the primer RNA to start DNA synthesis, and then form a D-loop structure, into which various replication proteins of *E. coli* like DnaB helicase, DnaG primase and DNA polymerase III holoenzyme, are introduced to continue replication of ColE2 DNA. The ColE2 Rep protein has a potential HTH DNA binding motif based on amino acid sequence homology to some known *E. coli* DNA binding proteins [12]. The ColE2 Rep proteins are mainly present as dimers in solution .However, it has not been examined yet whether Rep binds to Ori as a monomer, dimer or oligomer, which is essential for proposing a model of the Rep-Ori interaction.

In this paper, we used an EMSA-based method to determine the molecular weight of the ColE2 Rep protein in the Rep-Ori complex. The result suggested that Rep binds to Ori as a monomer. In addition, the result of EMSA using Rep protein fused with the maltose binding protein and the His_6-tag also supported the notion that Rep binds to Ori as a monomer.

Materials and Methods

Strains and plasmids. The *E. coli* strains used were AG1*recA1* [12] and BL21(DE3) (Novagen). The plasmids used have been described elsewhere [13], except for those described below.

Construction of plasmids. To construct pMAL243cRI [14] carrying the *malE*-fused ColE2 *rep* gene, the 0.95-kb *EcoR*I-*Hind*III fragment of pBE243RI-II [10] carrying the ColE2 Rep coding region was inserted between the *EcoR*I and *Hind*III sites (within the MCS) of pMAL-cRI (NEB). To construct pETmal243, the *Pvu*II-*Hind*III fragment of pMAL243cRI with the *malE*-fused ColE2 *rep* gene was cloned between the *Xba*I (filled with T4 DNA polymerase) and *Hind*III sites of pET-21a(+) (Novagen). pETE2+T7+site RIICm [15] was constructed by inserting the *BamH*I (filled by T4 DNA polymerase) fragment with chloramphenicol resistance gene into *Sal*I site of pET21a-E2Rep+site[10]. To construct pACE2H+T7, the *Bst*EII-*Sca*I fragment of pETE2+T7+site RIICm containing chloramphenicol resistance gene was cloned between the *Bst*EII-*Sca*I sites of pACYC177 [13], followed by selfligation after digesting with *Pst*I and filled with T4 DNA polymerase, then the *EcoR*I-*Bgl*II fragment of this plasmid containing the chloramphenicol resistance gene was replaced with the *EcoR*I-*Bgl*II fragment of pET21a-E2Rep+site with the ColE2 *rep* gene.

Determination of the molecular weight of Rep in the Rep-Ori complex. A series of native polyacrylamide gels (20×20 cm^2) covering a range of 4.5-10% acrylamide concentration (29:1 acrylamide:bis-acrylamide) were used to perform EMSA. Prerun was done for 15 min in 0.5×TBE buffer. A 38-bp DNA fragment containing the origin was obtained by annealing of 34mer Ori1 (5'-TTAATGAGACCAGATAAGCCTTATCAGATAACAG-3') and Ori2 (5'-TTCGCTGTTATCTGATAAGGCTTATCTGGTCTCA-3') oligonucleotides and 3'-end labeling with [α-^{32}P]dATP and Klenow fragment at room temperature for 30 min. Binding reactions were performed with 5 nM 38-bp Ori fragments and 100 nM Rep protein in 20 μl of the reaction mixture containing the binding buffer (200 mM Tris-HCl pH7.5, 10 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 10% ethylene glycol), 0.04 μg/μl BSA and 0.02 μg/μl salmon DNA. After 15-min incubation at 25°C, 1 μl loading buffer (10 mM Tris-HCl pH7.5, 0.025% bromophenol blue) was added to 5 μl sample and

loaded onto the prerun gel. One μg of each protein standard (non-denatured molecular weight markers, Sigma) was mixed with the same loading buffer and run alongside the binding reaction mixture. Electrophoresis was run at 4°C at 420V (constant voltage) until the running dye reached 1 cm from the bottom of the gel. Then the gel was treated as described [16], except that the lane containing the binding reaction mixture was used to expose the imaging plate (Fuji) and analyzed by using BAS 1500 (Fuji) and the image reader v 1.7J software. The relative mobility (Rf) of the Rep-Ori complex and the protein standards were determined by dividing their migration distances from the origin of the separating gel to the center of the protein bands by the migration distances of the bromphenol blue tracking dye from the origin of the same gel. Some protein standards contained more than one band due to the presence of charge isomers, and the Rf of the major isomer was used as described (Sigma Technical Bulletin. NO. MKR-137). The data were analyzed according to Ferguson [17]. First, a plot of the logarithm of the relative mobilities against the percent gel concentrations was constructed for each protein standard and the Rep-Ori complex. The slope of such a curve is the retardation coefficient (Kr). From these curves, individual slopes (Kr) were determined for each protein and the negative slope was plotted against the molecular weight of each protein standard. The resulting linear plot was used to determine the molecular weight of the Rep-Ori complex from its negative slope. All the curves were made by using Excel software (linear approximation).

Enzymes and chemicals. Enzymes used were from commercial sources. [α-^{32}P] dCTP (3,000 Ci/mmol) and [α-^{32}P] dATP (3,000 Ci/mmol) was obtained from Amersham Biosiences.

Purification of Rep proteins. The proteins were purified from BL21 (DE3) cells essentially as described [10], except for the step of expression. Cells were grown for 3 hr at 37°C in 2 ml terrific broth [9] supplemented with 20 μg/ml kanamycin for the cells expressing the His-Rep proteins alone, 50 μg/ml ampicillin for those expressing the MBP-Rep proteins alone and with 50 μg/ml ampicillin and 20 μg/ml kanamicin for the cells coexpressing the His-Rep and MBP-Rep proteins. The culture was diluted to 1 to 2 $\times 10^7$ cells/ml with 50 ml fresh medium, grown at 25°C to 1×10^8 cells/ml and treated with 0.4 mM IPTG for additional 4hr.

EMSA using the fusion Rep proteins. The origin DNA probe used in EMSA was

generated by digesting pBlue22+wtori [10] with *BssH*II and 3'-end labeling with [α-^{32}P]dCTP and Klenow fragment at room temperature for 30 min. The 230-bp labeled *BssH*II fragment containing the ColE2 origin were purified by PCR purification Kit (QIAGEN). Five nM labeled origin DNA fragment was incubated for 15 min at 25°C with the crude extracts containing His-Rep, MBP-Rep or His-Rep and MBP-Rep in 27 μl reaction mixtures containing the binding buffer (200 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 10% ethylene glycol), 40 μg/ml BSA and 20 μg/ml salmon DNA. Aliquots were loaded onto 8% polyacrylamide-0.5×TBE gels (20×20 cm^2). Electrophoresis was run at 4°C at 420 V (constant voltage for 40 min. DNA was visualized as described above.

Results and Discussion

ColE2 Rep binds to Ori as a monomer

The ColE2 Rep proteins are present mostly in a dimeric form in solution [14]. We sought to determine whether Rep binds as a monomer or as a dimer to its specific site in Ori. We used an EMSA-based method [16], which is an adaptation of that of Ferguson [17] to determine the molecular weight of the Rep-Ori complex (Fig. 2; for details see the experimental procedures). A series of native polyacrylamide gels covering a range of acrylamide concentration of 4.5-10% were used. In non-denaturing gels, mobility is influenced by molecular weight, conformation and charge. Ferguson's analysis allows the molecular weight of a native protein to be determined indirectly by eletrophoresing it, along with a number of standard proteins, on a series of non-denaturing gels of different polyacrylamide concentrations. As the ionic conditions are identical in each experiment, the decrease in mobility observed as the polyacrylamide concentration is increased is due to the sieving effect of the gel and hence is related to the size and shape of the native protein but not its charge. Unknown molecular weights can be determined based on the changes in mobility.

First, a plot of the logarithm of the relative mobilities against the percent gel concentration is constructed for each protein standard or the Rep-Ori complex (Fig. 1A), and the negative slopes were determined. Then, the negative slopes were plotted against the molecular weights of protein standards (Fig. 1B). By using the resulting linear plot molecular weight of the Rep-Ori complex was determined to be 66.1 K ($\pm$0.4 K) from its negative slope. Subtracting the contribution of the molecular weight of the origin DNA oligonucleotide (38 bp; 24.4 K) provides an estimate of the molecular weight of the total Rep protein component in a single Rep-Ori complex, 41.7.K ($\pm$0.4 K). The molecular weight of the His_6-tagged Rep protein used in the experiment is 35 K, and so we concluded that Rep binds to Ori as a monomer.

Analysis of specific Rep-Ori binding by EMSA using the fusion Rep proteins

To further confirm the conclusion obtained above, the Rep protein fused with either

a His_6-tag (His-Rep) or a maltose binding protein (MBP-Rep) at the C or N terminal end, respectively, was used to perform EMSA. The complex of His-Rep (about 35 K) and the Ori fragment migrated much faster than that made with MBP-Rep (about 76 K) by virtue of its lower molecular weight (Fig. 2, lanes 2 and 3). When various amounts of the proteins were used, both His-Rep and MBP-Rep formed single species of the protein-DNA complexes with the Ori fragment (data not shown). We extracted total proteins from *E. coli* cells producing His-Rep and MBP-Rep protein, simultaneously. After incubation of the crude extract with the Ori fragment, only two complexes corresponding to those formed with His-Rep and with MBP-Rep alone were detected (Fig. 2, lane 4). No band of an intermediate mobility was seen. We then used the NTA-Ni resin to remove the MBP-Rep proteins from the same cell extract. Only His-Rep and heterodimers of His-Rep and MBP-Rep proteins were purified, because MBP-Rep protein alone did not specifically bind to the NTA-Ni resin (data not shown). The eluted Rep proteins were incubated with the Ori fragment and a similar result (Fig. 2, lane 5) was obtained to that using the crude extract (Fig. 2, lane 4). The presence of the complex corresponding to that formed with MBP-Rep alone and the absence of the complex with an intermediate mobility corresponding to the complex containing the heterodimer of His-Rep and MBP-Rep indicated that the heterodimer of MBP-Rep and His-Rep was indeed formed but that it did not form a complex with the Ori fragment and that MBP-Rep dissociated from the heterodimer formed a complex with the Ori fragment. We concluded that both MBP-Rep and His-Rep proteins bind to Ori as a monomer and not as a dimer.

A possible biological significance of the dimer of Rep.

The Rep proteins of some replicons, such as pPS10 [18], pSC101 [19], F [20] and R6K [21] have an important second function. They recognize invertedly repeated sequences (operators) which overlap with the promotor of their own coding genes, acting as auto-repressors. For pPS10 [18, 22], pSC101 [23] and F [24] replicons, dimers of their Rep proteins bind to the operator sequences, whereas monomers bind to the iterons in the origins to initiate replication of DNA. For R6K replicon, not only the monomer but also the dimer of π protein bind to the iterons [25-28]. The ColE2 Rep

proteins are mainly present as dimers in solution [14], while they bind to Ori as monomers as decribed above. The ColE2 Rep protein does not seem to function as a repressor of trascription of its own coding gene, as only very faint affinity, if any, of Rep and the DNA fragment containing the promoter region of the *rep* gene was detected even in the presence of excess Rep proteins (data not shown). The copy number of the ColE2 plasmid is regulated by plasmid-coded antisence RNA against the *rep* mRNA, which maintained a constant level of expression of the ColE2 Rep protein in the host cells at a post-transcriptional step [29-31]. But this mechanism alone might not be able to respond rapidly to sudden variation in the concentration of Rep. We propose here the efficient dimerization of Rep might be involved in keeping the copy number of the ColE2 plasmid at the normal level by limiting the amount of the active monomeric form of Rep in the cell.

Figure legends.

Fig. 1. Determination of the molecular weight of Rep-Ori complex. (A) Calibration curves. Logarithm of relative mobility (R_f) plotted against percentage acrylamide. Showing the relationship between mobility of each species and the gel concentration: α-lactalbumin (✦); Carbonic Anhydrase (■); Albumin, egg (●); BSA, monomer (▲); BSA, dimer (✕); BSA, trimer (✱); Rep-Ori complex (+). The negative slope of each protein standard and Rep-Ori complex were determined. (B) The negative slope of each line in (A) plotted against the molecular weight of standards: 1, α-lactalbumin, 14.2 K; 2, Carbonic Anhydrase, 29 K; 3, Albumin, egg, 45 K; 4, BSA, monomer, 66 K; 5, BSA, dimer, 132 K; 6, BSA, trimer, 198 K. The resulting line of plot was used to determine the molecular weight of Rep-Ori complex (●) from its negative slope.

Fig. 2. Specific Rep-Ori binding with MBP-Rep and His-Rep. Crude cell extracts or purified proteins were incubated with the origin DNA fragments. Lane 1, no protein; lanes 2~4, crude extracts of *E.coli* cells with pACE2H+T7 producing the His-Rep proteins (lane 2), with pETmal243 producing the MBP-Rep proteins (lane 3), and with both plasmids producing the MBP-Rep and His-Rep proteins simultaneously (lane 4); lane 5, purified proteins (by NTA-Ni resins) using the same crude extract used for lane 4.

References

[1] H. R. Herschman, and D. R. Helinski, Comparative study of the events associated with colicin induction, J Bacteriol 94 (1967) 691-699.

[2] M. Bazaral, and D. R. Helinski, Circular DNA forms of colicinogenic factors E1, E2 and E3 from Escherichia coli, J Mol Biol 36 (1968) 185-194.

[3] T. Horii, and T. Itoh, Replication of ColE2 and ColE3 plasmids: the regions sufficient for autonomous replication, Mol Gen Genet 212 (1988) 225-231.

[4] H. Yasueda, T. Horii, and T. Itoh, Structural and functional organization of ColE2 and ColE3 replicons, Mol Gen Genet 215 (1989) 209-216.

[5] M. Kido, H. Yasueda, and T. Itoh, Identification of a plasmid-coded protein required for initiation of ColE2 DNA replication, Nucleic Acids Res 19 (1991) 2875-2880.

[6] D. T. Kingsbury, and D. R. Helinski, DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1, Biochem Biophys Res Commun 41 (1970) 1538-1544.

[7] W. Tacon, and D. Sherratt, ColE plasmid replication in DNA polymerase I-deficient strains of Escherichia coli, Mol Gen Genet 147 (1976) 331-335.

[8] T. Itoh, and T. Horii, Replication of ColE2 and ColE3 plasmids: in vitro replication dependent on plasmid-coded proteins, Mol Gen Genet 219 (1989) 249-255.

[9] S. Takechi, and T. Itoh, Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism, Nucleic Acids Res 23 (1995) 4196-4201.

[10] M. Yagura, S. Y. Nishio, H. Kurozumi, C. F. Wang, and T. Itoh, Anatomy of the replication origin of plasmid ColE2-P9, J Bacteriol 188 (2006) 999-1010.

[11] S. Takechi, H. Matsui, and T. Itoh, Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication, Embo J 14 (1995) 5141-5147.

[12] S. Hiraga, T. Sugiyama, and T. Itoh, Comparative analysis of the replicon regions of eleven ColE2-related plasmids, J Bacteriol 176 (1994) 7233-7243.

[13] A. C. Chang, and S. N. Cohen, Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid, J Bacteriol 134 (1978) 1141-1156.

[14] K. Matsumoto, and T. Itoh. unpublished data.

[15] M. Yagura, and T.Itoh. unpublished data.

[16] K. Orchard, and G. E. May, An EMSA-based method for determining the molecular weight of a protein--DNA complex, Nucleic Acids Res 21 (1993) 3335-3336.
[17] K. A. Ferguson, Starch-Gel Electrophoresis--Application to the Classification of Pituitary Proteins and Polypeptides, Metabolism 13 (1964) SUPPL:985-1002.
[18] D. Garcia de Viedma, R. Giraldo, M. J. Ruiz-Echevarria, R. Lurz, and R. Diaz-Orejas, Transcription of repA, the gene of the initiation protein of the Pseudomonas plasmid pPS10, is autoregulated by interactions of the RepA protein at a symmetrical operator, J Mol Biol 247 (1995) 211-223.
[19] C. Vocke, and D. Bastia, The replication initiator protein of plasmid pSC101 is a transcriptional repressor of its own cistron, Proc Natl Acad Sci U S A 82 (1985) 2252-2256.
[20] L. A. Rokeach, L. Sogaard-Andersen, and S. Molin, Two functions of the E protein are key elements in the plasmid F replication control system, J Bacteriol 164 (1985) 1262-1270.
[21] M. Filutowicz, G. Davis, A. Greener, and D. R. Helinski, Autorepressor properties of the pi-initiation protein encoded by plasmid R6K, Nucleic Acids Res 13 (1985) 103-114.
[22] D. Garcia de Viedma, A. Serrano-Lopez, and R. Diaz-Orejas, Specific binding of the replication protein of plasmid pPS10 to direct and inverted repeats is mediated by an HTH motif, Nucleic Acids Res 23 (1995) 5048-5054.
[23] D. Manen, L. C. Upegui-Gonzalez, and L. Caro, Monomers and dimers of the RepA protein in plasmid pSC101 replication: domains in RepA, Proc Natl Acad Sci U S A 89 (1992) 8923-8927.
[24] M. Ishiai, C. Wada, Y. Kawasaki, and T. Yura, Replication initiator protein RepE of mini-F plasmid: functional differentiation between monomers (initiator) and dimers (autogenous repressor), Proc Natl Acad Sci U S A 91 (1994) 3839-3843.
[25] J. Wu, M. Sektas, D. Chen, and M. Filutowicz, Two forms of replication initiator protein: positive and negative controls, Proc Natl Acad Sci U S A 94 (1997) 13967-13972.
[26] M. Urh, University of wisconsin-madison, ed. 1996.
[27] M. Urh, J. Wu, K. Forest, R. B. Inman, and M. Filutowicz, Assemblies of replication initiator protein on symmetric and asymmetric DNA sequences depend on

multiple protein oligomerization surfaces, J Mol Biol 283 (1998) 619-631.
[28] S. Kunnimalaiyaan, R. Kruger, W. Ross, S. A. Rakowski, and M. Filutowicz, Binding modes of the initiator and inhibitor forms of the replication protein pi to the gamma ori iteron of plasmid R6K, J Biol Chem 279 (2004) 41058-41066.
[29] T. Sugiyama, and T. Itoh, Control of ColE2 DNA replication: in vitro binding of the antisense RNA to the Rep mRNA, Nucleic Acids Res 21 (1993) 5972-5977.
[30] S. Takechi, H. Yasueda, and T. Itoh, Control of ColE2 plasmid replication: regulation of Rep expression by a plasmid-coded antisense RNA, Mol Gen Genet 244 (1994) 49-56.
[31] H. Yasueda, S. Takechi, T. Sugiyama, and T. Itoh, Control of ColE2 plasmid replication: negative regulation of the expression of the plasmid-specified initiator protein, Rep, at a posttranscriptional step, Mol Gen Genet 244 (1994) 41-48.
[32] M. Han, K. Aoki, M. Yagura, and T. Itoh. unpublished data.

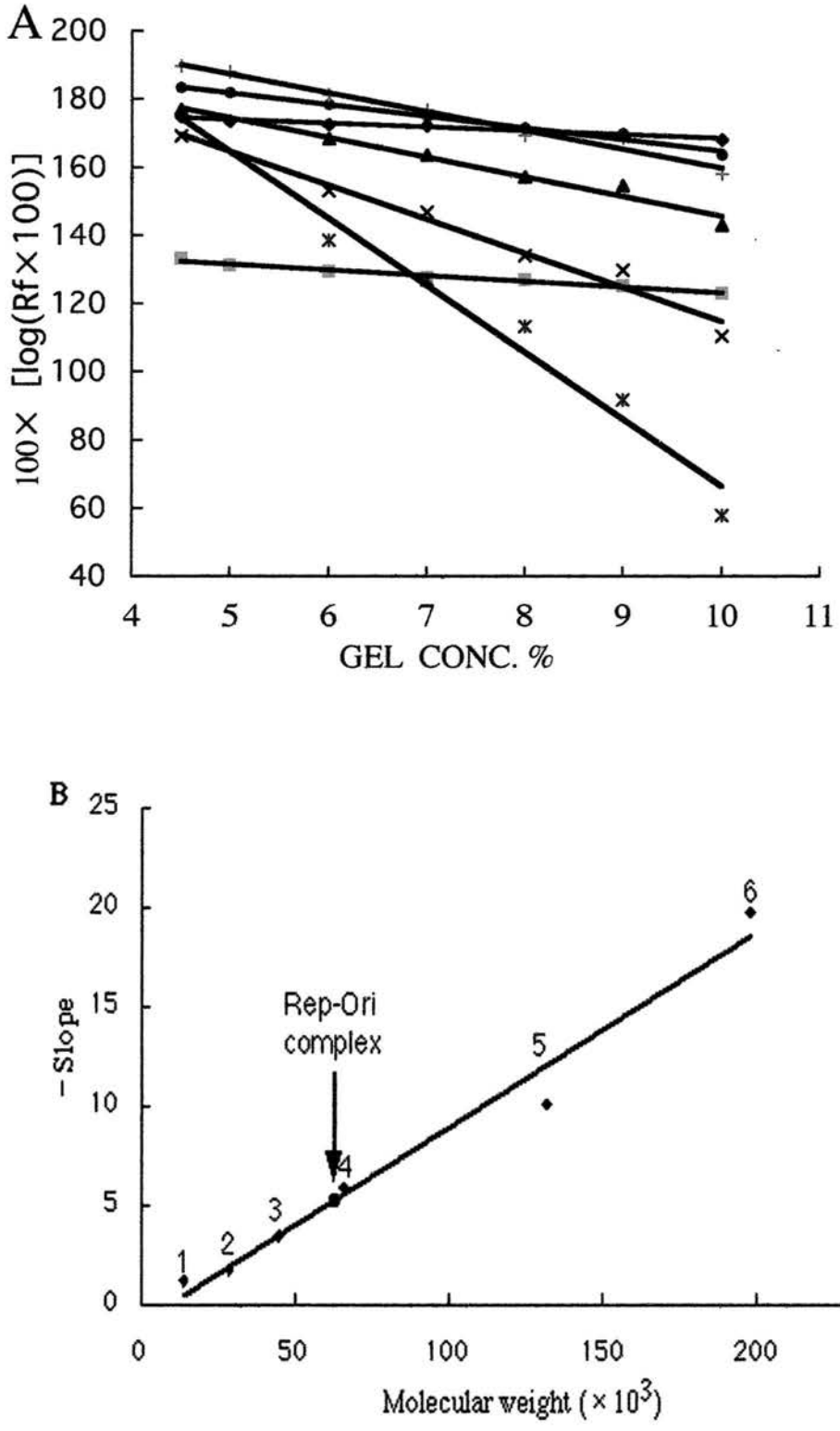

A
200
180
160
140
120
100
80
60
40
100× [log(Rf×100)]
4
5
6
7
8
9
10
11
GEL CONC. %
B
25
20
15
10
5
0
- Slope
Rep-Ori
complex
1
2
3
4
5
6
0
50
100
150
200
250
Molecular weight (× 10³)

Fig. 1

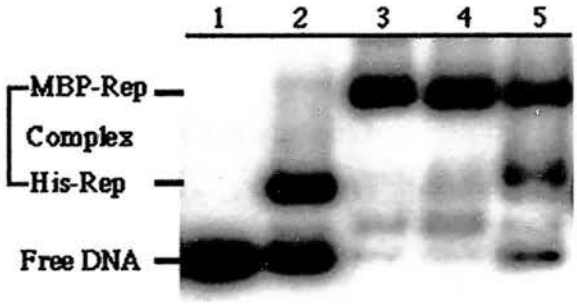


Fig. 2

III. Characterization of Specific Interaction between the Initiator Protein (Rep) andOrigin of Plasmid ColE2-P9

ABSTRACT

The replication initiator protein (Rep) of plasmid ColE2-P9 (ColE2) is multifunctional. We are interested in how Rep binds to the origin (Ori) to perform various functions. We used the wild type and variants of Rep to study Rep-Ori interaction by both *in vitro* and *in vivo* approaches, including biochemical analyses of protein-DNA interaction and *in vivo* replication assay. We identified three regions (I, II and III) of Rep located in the C terminal half and corresponding three binding sites (I, II and III) in Ori important for Rep-Ori interaction. We showed that region I containing a putative HTH motif is necessary and sufficient for the specific Ori recognition, interacting with site I of the origin DNA from the major groove. Region II interacts with site II of the origin DNA from the adjacent minor groove in the left half of Ori and region III interacts with site III next to the template sequence for the primer synthesis, which is one and a half turn apart from site I on the opposite surface of the origin DNA. A putative linker region located between the two DNA binding domains (regions II and III) was identified, which might provide Rep an extended conformation suitable for binding to the two separate sites in Ori. Based on the results presented in this paper, we propose a model for Rep-Ori interaction, in which Rep binds to Ori as a monomer.

INTRODUCTION

DNA replication is a universal process for transmission of genetic information in all organisms. In the classic replicon model (15), initiation of DNA replication requires a cis-acting DNA sequence (the replicator) and a trans-acting factor (the initiator) that binds to the replicator, which is now called the origin (of replication). The origin-bound initiators melt the duplex DNA strands in a region adjacent to the origin to form an open complex (5, 10, 21). Then, the initiators introduce other replication proteins like the helicase, primase and DNA polymerase into the open complex to form the replisome. Among various initiators, the DnaA protein of *Escherichia coli* is the best characterized (16). Its ATP-bound form binds as a monomer specifically to each of the four asymmetric 9-bp consensus sequences (5'-TTA/TTNCACA) or its closely related sequences, the DnaA boxes in the origin (OriC). These ATP-bound DnaA proteins oligomerize with additional ATP-bound DnaA proteins to form a large nucleoprotein structure containing 20-40 DnaA proteins, which induces local unwinding of the AT-rich region adjacent to OriC. The DnaA proteins act as a replisome organizer for the subsequent loading of proteins required for the replication process (26).

In most plasmid replicons, initiation of plasmid replication requires plasmid-encoded initiator proteins, called Rep (5). The crystal structure of the RepE initiator protein of plasmid F was solved as the first three-dimensional structure of a prokaryotic initiator protein (20). The RepE protein consists of topologically similar N- and C- terminal domains related to each other by internal pseudo 2-fold symmetry. By using HTH DNA binding motifs, the two domains bind to two consecutive major grooves on one face of the DNA helix of each of the 19-bp directly repeated sequences (the iterons) in the origin. The amino acid sequence homology between the RepE protein and some other plasmid initiator proteins, such as those of pPS10 and pSC101 suggested that those proteins might bind to their iterons similarly (20). Although the P1 RepA protein can not be aligned due to its poor sequence homology with the RepE protein, the molecular modeling suggested that it is also structurally similar to the class of plasmid initiator proteins to which the RepE protein belongs (31). The replication initiation of these iteron plasmids involves binding of the plasmid encoded initiator proteins to the iterons and form an open complex at their cognate origins to recruit the replisome (5, 17).

The plasmid ColE2-P9 (ColE2) is a circular duplex DNA molecule of about 7 kb (11). It is present at about 10 to 15 copies per host chromosome (1, 13). The initiator protein (Rep; 35 kDa) of plasmid ColE2 is the only plasmid-specified *trans*-acting factor required for the initiation of plasmid replication (18, 19, 40). Initiation of plasmid replication also requires host DNA polymerase I (14, 34), but not RNA polymerase and DnaG primase (14, 35). The minimal *cis*-acting region, the origin (Ori; Fig. 1A), required for ColE2 DNA replication consists of 31 bp (38). The ColE2 Rep protein binds specifically to Ori DNA, as revealed by a filter binding assay (18) and by an electrophoresis mobility shift assay (EMSA) (38). In an *in vitro* ColE2 DNA replication system with crude extracts of *Escherichia coli*, ColE2 DNA replication starts at a unique site in Ori and proceeds unidirectionally (14, 35). The ColE2 Rep protein is unique among other initiator proteins in that it is also a plasmid-specific primase. It synthesizes a 3-nucleotide primer RNA which has a unique structure of 5' ppApGpA (36). Host DNA polymerase I specifically uses the primer RNA to start DNA synthesis, and then form a D-loop structure, into which various replication proteins of *E. coli* like DnaB helicase, DnaG primase and DNA polymerase III holoenzyme, are introduced to continue replication of ColE2 DNA.

The ColE2 Rep protein has a potential HTH DNA binding motif based on amino acid sequence homology to some known *E. coli* DNA binding proteins (12). Comparative studies on the replication of plasmid ColE2 and its close relative, plasmid ColE3-CA38, have shown that interaction of the Rep proteins with the origins of these two plasmids is plasmid-specific, although both of their Rep proteins and origins show a high degree of similarity. By using chimeric *rep* genes and chimeric origins, the two regions (A and B) in the C-terminal regions of the Rep proteins and the two sites (α and β) in the origins important for the determination of plasmid-specificity were identified (Fig. 1A and C; 32). The region B in the C-terminal region of Rep corresponding to the second recognition helix of a potential HTH DNA binding motif (12) was proposed to be a part of the sequence-specific DNA binding domain, and region A might be a linker connecting two domains in Rep involved in DNA binding (32).

Mutation analyses (38) have revealed that Ori may be divided into three subregions (I, II and III). Subregion I is important for stable binding of Rep and contains site β. Subregion II is

important for binding of Rep and for initiation of DNA replication and contains site α.. Subregion III is important for DNA replication but apparently not for binding of Rep and contains the site where the primer is synthesized. It has also been suggested that there are three Rep binding elements in Ori, one (site a) in subregion I and two (sites b and c) in subregion II.

Detailed analysis of the binding property of ColE2 Rep and Ori is an important prerequisite to investigate the following steps after Rep binding to Ori, such as the formation of the open complex and primer synthesis, providing insights into the mechanism of initiation of ColE2 DNA replication. In this paper we used the electrophoresis mobility shift assay (EMSA) and dimethyl sulfate (DMS) protection assay to characterize the interaction of Rep and Ori. We identified three regions in Rep involved in Ori binding and corresponding three binding sites in Ori. We propose a model for the mode of the ColE2 Rep-Ori binding.

MATERIALS AND METHODS

Strains and plasmids. The *E. coli* strains used were AG1*recA1* (12) and BL21(DE3) (Novagen). The plasmids used have been described elsewhere (3, 13, 27, 32, 38, 40), except for those described below.

Construction of plasmids expressing Rep fragments. The plasmid, pET21a-E2Rep +site (38) is a derivative of pET21a(+) (Novagen) carrying the ColE2 *rep* gene, in which two new EndoR cleavage sites (*Sac*II and *Sal*I) were generated without any changes in the amino acid residues. Plasmid pACYC177 (3) was digested with *Dra*I, followed by selfligation to construct pACYC177dAp, in which the β-lactamase gene was inactivated. To construct pHME2Rep+Km, the entire pACYC177dAp cleaved with *BstB*I was inserted at the *BstB*I site of pET21a-E2Rep+site located in the region encoding the C-terminal one third of the Rep protein. To construct pHMC132, the 4.12-kb *EcoR*I (filled with T4 DNA polymerase)-*Bgl*II fragment of pHME2Rep+Km containing pACYC177dAp and portions of the ColE2 *rep* gene was replaced with the PCR fragments (amplified with oligonucleotides KM3F and KM3R) digested with *Bgl*II enzyme. The construction of pHMC94, pHMC62 and pHMC37 was as described for pHMC132 except for the oligonucleotides used for amplification of the PCR fragments (d203 and KM3R for pHMC94, d236 and KM3R for pHMC62, and d260 and KM3R for pHMC37). Plasmid pHMC274 was constructed by digesting pETa21-E2Rep+site with *EcoR*I and *Spl*I, followed by filling with T4 DNA polymerase and selfligation. To construct pHMC169, the *BamH*I (filled with T4 DNA polymerase)-*Bgl*II fragment of pHME2Rep+Km containing pACYC177dAp and portions of the ColE2 rep gene was replaced with the RCP fragment (amplified with oligonucleotides KM2F and KM2R) digested with *Mbo*I (filled with T4 DNA polymerase) and *Bsp119*I. To construct pETE2RepC117 (23) pET21a-E2Rep+site was digested with *EcoR*I and *Asu*II, followed by filling with Klenow fragment and selfligation. To construct pHMC45, the *EcoR*I (filled with T4 DNA polymerase)-*Bgl*II fragment of pHME2Rep+Km containing pACYC177dAp and portions of the ColE2 *rep* gene was replaced with the PCR fragment (amplified with oligonucleotides KM3F and KM3R) digested with *Xho*I (filled with T4 DNA polymerase) and *Bgl*II.

Construction of plasmids expressing mutant Rep proteins with single amino acid substitutions. To introduce single amino acid substitutions in the C terminal half of the Rep protein, the PCR in the presence of $MnCl_2$ was performed in a reaction mixture (20 μl) containing 0.02 μg of the template DNA (pETa21-E2Rep+site), 10 pmol each of oligonucleotides KM3F and KM3R, 0.5 units Gene Taq DNA polymerase (Wako), 1 μl of 10×Gene Taq buffer, 0.2 mM each of 4dNTPs and 0.35 mM $MnCl_2$. Amplification was done through 30 cycles of denaturation (94°C 1 min), annealing (55°C 40 sec), and polymerization (72°C 30 sec). The PCR fragments digested with *Sal*I and *Bgl*II were inserted between the *Sal*I and *Bgl*II sites of pHME2Rep+Km and the nucleotide sequences of the PCR-amplified portions of the resultant plasmids were determined to identify single amino acid substitutions in the Rep proteins encoded by the mutant plasmids. In some cases *in vivo* replication activity (see below) of the resultant plasmids was measured before sequencing to screen the mutant Rep proteins defective in replication. The plasmid carrying the mutant *rep* gene encoding RepΔC20 protein was among them. The mutagenic oligonucleotide KA1 (Table 1) and oligonucleotide KM3F were used to construct plasmids producing the mutant Rep proteins RepT284W and RepR287Q.

Construction of other plasmids. The 1.5-kb *BamH*I-*Bgl*II fragment of pEC22X43 (13), containing the ColE2 *rep* gene with an *Xba*I linker inserted in the RNAI-coding region, and the 2.1-kb *BamH*I-*Bgl*II fragment of pTI20 (40), containing the ColE1 replicon and the β-lactamase gene, were ligated to construct pTI243+. The 2.4-kb *Eco*RI-*Dra*I fragment of pTI243+, containing the ColE1 replicon and the ColE2 *rep* gene, and the 1.4-kb *Eco*RI-*Sma*I segment of pKC7 (27) containing the kanamycin resistance gene, were ligated to construct pTIK243+. The 252-bp *Eco*RI-*Eco*RV fragment of pTI51dr8 (38) and the 258-bp *Eco*RI-*Eco*RV fragment of pTI51dr4 (38) were cloned between the *Eco*RI and *Eco*RV sites of pBluescript KS+, followed by digestion with *Nhe*I and *Xho*I and by selfligation to construct pHMdr8 and pHMdr4, respectively. The 419-bp *Eco*RI-*Hind*III fragment of pTI51dr56 (38) was cloned between the *Eco*RI-*Hind*III sites of pBluescript KS+, then digested with *Nhe*I and *Xho*I, followed by selfligation to construct pHM dr56.

Media, enzymes, and chemicals. The media used have been described elsewhere (13).

Chemicals, enzymes and antibiotics used were from commercial sources. [α-^{32}P] dCTP (3,000 Ci/mmol) was obtained from Amersham Biosiences. Oligonucleotides used for mutagenesis were obtained from Pharmacia and are listed in Table 1.

***In vivo* replication assay.** Derivatives of pETa21-E2Rep+site producing the mutant Rep proteins with single amino acid substitutions or pETa21-E2Rep+site (as a positive control) were introduced into *E. coli* AG1 cells carrying pTIK243+ and pEC22s (32) with the chloramphenicol resistant gene and the ColE2 origin. Cells were plated onto a LB agar plate containing ampicillin (50 μg/ml). After incubation for 14 hr at 37°C, 20 colonies selected randomly were transferred to a new LB agar plate containing ampicillin and incubated for 14 hr at 37°C. Cells grown from each colony were transferred to three new LB agar plates containing ampicillin (50 μg/ml) and kanamycin (20 μg/ml), ampicillin (50 μg/ml) and chloramphenicol (20 μg/ml), and ampicillin only, and incubated for 14 hr at 37°C. During the incubation only with ampicillin, pTIK243+ (containing a ColE1 replicon) were excluded from cells by pET21a+E2Rep+site (containing a pBR322 replicon) or its derivatives due to unidirectional incompatibility (9, 37). Growth or no growth of cells on the plates containing ampicillin and chloramphenicol indicates whether the mutant Rep proteins produced by pET21a+E2Rep+site derivatives retain the replication activity or not.

Purification of Rep protein. The ColE2 Rep protein, its fragments and mutant Rep proteins with a His_6-tag at the C-terminus were purified from BL21 (DE3) cells carrying derivatives of pET21a(+) producing these peptides essentially as described (38), except for the step of expression of the peptides. Cells were grown for 3 hr at 37°C in a 2 ml of terrific broth (35) supplemented with 50 μg/ml ampicillin. The culture was diluted to 1 to 2×10^7 cells/ml with 50 ml fresh medium, grown at 25°C to 1×10^8 cells/ml and treated with 0.4 mM IPTG for another 4hr. The protein concentration was determined by the method of BCA (33). Bovine serum albumin (BSA) was used as a standard. The amount of the overexpressed wild type Rep protein in the whole cell lysate was about 20% of the total *E. coli* proteins , and the soluble Rep protein in the supernatant of high speed centrifugation was about 10 to 20% of the total Rep protein. Most of the Rep fragments showed better solubility than the intact Rep and more than 50% of the total amounts were in the supernatant fractions, except for

RepC274 or RepC94 whose solubility was similar to or lower (about 10%) than that of the wild type Rep protein, respectively. The solubility of the overexpressed mutant Rep proteins were approximately similar to that of the wild type Rep protein.

EMSA. The origin DNA probe used in the electrophoresis mobility shift assay was generated by digesting pBlue22+wtori (38) with *BssH*II and 3'-end labeling with [α-^{32}P]dCTP and Klenow fragment at room temperature for 30 min. The 230-bp labeled *BssH*II fragment containing the ColE2 origin and another 83-bp *BssH*II fragment without the ColE2 origin (as a negative control) were purified by PCR purification Kit (QIAGEN). Five nM labeled origin DNA fragment was incubated for 15 min at room temperature with the wild type or mutant Rep proteins (up to 500 nM) in 27 µl reaction mixtures containing the binding buffer (200 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 10% ethylene glycol), 40 µg/ml BSA and 20 µg/ml salmon DNA. Aliquots were loaded onto 8% polyacrylamide-0.5×TBE gels (20×20 cm^2). Electrophoresis was run at 4°C at 420 V (constant voltage). DNA was visualized by using Fuji BAS1500 phosphorimager and image reader v 1.7J software.

Western blot analysis. BL21 (DE3) cells carrying derivatives of pET21a+E2Rep+site producing mutant Rep proteins with a single amino acid substitution were grown to the stationary phase in the absence of IPTG and washed in buffer (1 mM Na_2HPO_4, 1 mM NaH_2PO_4, 1 M NaCl, 20 mM imidazole and 0.5% tween20). Proteins in the whole cell extracts were separated on 15% SDS-PAGE and detected by staining with Coomassie brilliant blue. Samples for Western blotting were separated in parallel on another 15% SDS-PAGE and proteins were transferred to a polyvinyliden difluoride sheet (PVDF; Atto). The sheet was treated with the anti-T7 tag antibody (Novagen) at a 1:8000 dilution for 1h and then with NBT/BCIP (Roche) at room temperature as recommended by the supplier (Novagen manual).

Dimethyl sulfate footprinting on supercoiled plasmid. The wild type or mutant Rep proteins (up to 400 nM) were allowed to bind to 0.2 pmol supercoiled DNA of pBlue22+wtori in the binding buffer (54 µl) used in EMSA for 10 min at room temperature. Modification reactions were started by precisely adding 1.1 µl of pure DMS (Sigma) at a final concentration of 200 mM and incubated at room temperature. Reactions were stopped exactly after 3 min by adding 110-µl stop solution (1 M β-mercaptoethanol, 20 mM EDTA) and DMS

was thoroughly eliminated by ethanol precipitation.

Primer extension. Modified plasmid DNA molecules treated with DMS were redissolved in 15 μl of distilled Water (miliQ). Primer extension was performed in a 10-μl reaction mixture containing 0.1 pmol plasmid DNA, 0.5 units of EX Taq (TAKARA), 1 μl of 10×EX Taq buffer, 0.25 mM each of 4dNTPs, 1 pmol each of FITC-labeled M13 reverse primer (5'-f GGAAACAGCTATGACCA-3') or M13-20 primer (5'-f GTAAAACGACGGCCAGT-3'). Amplification was done through 36 cycles of denaturation (95°C 1 min), annealing (50°C 1 min) and polymerization (72°C 1 min). The DNA molecules were purified by ethanol precipitation, disolved in 2 μl of loading dye solution and heated at 90°C for 2 min. Samples were loaded on a 8% polyacrylamide-urea sequencing gel and run in 1.2×TBE buffer (108 mM Tris-HCl, 2.4 mM EDTA-2Na, 108 mM boric acid, pH 8.3) for 8 hr (DSQ 1000, DNA sequencer; Shimadzu). The nucleotides modified by DMS were precisely localized by comparison with dideoxy-sequencing ladders obtained on unmodified pBlue22+wtori plasmid using the same primer used for primer extension and run side by side.

RESULTS

Analysis of the specific Rep-Ori binding by EMSA. In order to identify the regions in Rep important for Ori recognition and binding, we constructed a series of partial Rep fragments lacking either the N terminal portions or the C terminal portions and the mutant Rep proteins with single amino acid substitutions (Fig. 1B and C). The selection of end points in partial Rep fragments were made with reference to the result of the secondary structure prediction (Psi-Pred) of Rep (Fig. 1C).

We then compared the Ori binding activity of the intact Rep protein and its partial fragments by EMSA (Fig. 2A). The wild type Rep protein bound specifically to Ori sequence, as reported previously (38). All of the Rep fragments including RepC117 (23) also bound specifically to Ori. This indicated that they folded properly to form the sequence-specific DNA binding domains (DBDs). Note that detailed comparison of the binding affinity of the intact Rep protein and the partial fragments is difficult, as we do not know the ratios of the active forms of these proteins in the soluble fractions. RepC37 consisting of the C terminal 37 amino acids of the Rep protein with a putative HTH DNA binding motif (positions 261 to 297, Fig. 1C) bound to Ori specifically (Fig. 2A). This suggested that the putative HTH motif alone is probably sufficient for the sequence specific binding to Ori.

To show further the importance of the putative HTH motif in the specific recognition of Ori, we used a deletion mutant Rep protein named RepΔC20 lacking the C terminal 20-amino-acid region of Rep containing the second recognition helix of the putative HTH motif (Fig. 1C). It failed to form a complex with Ori DNA (Fig. 2B), indicating that it can not bind to Ori specifically. This result suggested that the second helix of the putative HTH motif is critical to the specific Ori DNA recognition. We further used four mutant Rep proteins with single-amino-acid substitutions in the second helix of the putative HTH motif at positions 284, 285, 286 and 287 (Fig. 1C) and examined their Ori binding activity by EMSA (Fig. 2C and D). All of them were defective, showing the importance of these amino acids. All these results also suggested that the second helix of the putative HTH motif is critical to the specific Ori DNA recognition and these four amino acid residues may be involved directly in the specific

recognition. All these mutant Rep proteins and RepΔC20 lost the *in vivo* activity to initiate DNA replication at Ori (measured by the method described in Materials and Methods).

Substitutions or deletion of amino acids at the C terminal region might have reduced or abolished the stability of the mutant Rep proteins. To exclude the possibility, the expression of the T7-tagged wild type and mutant Rep proteins in uninduced *E. coli* cells was detected by using the anti-T7 tag antibody (Fig. 3B), as the Rep proteins were invisible by dye-staining (Fig. 3A). The amounts of the mutant Rep proteins in uninduced *E. coli* extracts were approximately similar to that of the wild type Rep. This suggested that all the mutant Rep proteins were as stable as the wild type Rep protein and that the defects of the mutant Rep proteins in initiation of *in vivo* DNA replication were not due to the absence of these proteins *in vivo*.

DMS protection assay to identify the Rep contact sites in Ori and additional domains in Rep involved in the Ori binding. As it is difficult to examine detailed properties of the Rep-Ori interaction and to identify other DNA binding domains in Rep besides the putative HTH motif in the C terminal region by using EMSA, we performed dimethyl sulfate (DMS) protection assay for further analyses. DMS methylates the N7 position of guanine from the major groove and the N3 position of adenine from the minor groove on dsDNA substrates (24). Adenine residues are less reactive to DMS than guanine residues (and therefore the intensities of signals at A residues in the DMS footprints are usually lower than those at G residues). On ssDNA, it modifies the N1 position of adenine and, to a lesser extent, the N3 position of cytosine (25). We treated supercoiled plasmid DNA molecules containing Ori with DMS in the presence or absence of Rep and various Rep fragments and modified residues in the Ori sequence were mapped by primer extension (Fig. 4).

Binding of the intact ColE2 Rep protein to Ori resulted in the protection of some bases at or close to the three sites (I, II and III in Fig. 4A). In site I, G at position 6 (6G) on the top strand and 10G and 11G on the bottom strand were protected. In addition, 5T, 7A and 8G on the top strand and 9T and 12T on the bottom strand seemed to be protected. In site II, 14A, 16A and 17A on the top strand were significantly protected. In site III, 24A and 25G on the bottom strand were protected. These results indicated that Rep interacts with those bases in

sites I and III of Ori from the major grooves and with those in site II from the minor groove based on the specificity of DMS modification. Sites I and II are located adjacent to each other roughly on one surface of dsDNA in the left half of Ori and site III on the opposite surface next to the primer sequence.

Stimulation of DMS modification was detected at 12A and 18G on the top strand and 19G on the bottom strand. These bases are located at the boundary between sites I and II and just outside of site II, suggesting some distortion of dsDNA upon binding of Rep. In site III, modification at 25C on the top strand was greatly stimulated and modification at neighboring positions (23A, 24T, 26A, 27G, 28A and 29T) on the top strand was also stimulated. Stimulation at 28A was very weak in the result shown (Fig. 4A). These bases in site III are all on the strand complementary to the template strand for the primer synthesis, suggesting possible distortion or opening of dsDNA in site III of Ori upon binding of Rep. Such a conformational change of dsDNA might be required for the primer synthesis by Rep. Protection at 6G on the top strand and at 10G, 11G and 25G on the bottom strand and stimulation at 27G are consistent with the results obtained by using the Rep protein without the His_6-tag and linear Ori DNA (22).

Very similar results (Fig. 4B) were obtained with RepC132 (Fig. 1) as compared with the intact Rep protein. These results suggested that the C-terminal region of Rep with 132 residues is sufficient for interaction with all the three sites in Ori and that the N-terminal region is unnecessary for the specific Ori DNA binding. Rep C132 is also capable of inducing a conformational change of dsDNA at and around site III. RepC37 containing the putative HTH motif alone was capable of apparently very weak binding to Ori specifically as shown by EMSA (Fig. 2A). In fact, we were able to detect a very faint protection in a part of site I in the presence of far excess amounts of RepC37 over Ori in the DMS protection assay (data not shown), suggesting that the interaction between the putative HTH motif alone with Ori was very unstable. On the other hand, when RepC45 containing an additional 8 amino acids absent in RepC37 was used in the DMS protection assay, the bases in site I were clearly protected (Fig. 4C). The additional amino acids in RepC45 seemed to stabilize the binding of the putative HTH motif to Ori. Based upon the results of EMSA and those of DMS protection

assays, we propose that the region from positions 261 to 297 in Rep (region I) contains a DBD, which recognizes and binds to site I of Ori.

To identify the region in Rep involved in binding to site II, we used RepC62, which contains additional 16 amino acids at the N terminus of RepC45 (Fig. 1). The additional region contains region A of Rep involved in determining the specificity of the Rep-Ori binding in plasmids ColE2-P9 and ColE3-CA38 (32). RepC62 showed a clear protection at the bases in site II (Fig. 4D). This suggested that the stable binding of Rep to site II of Ori requires the amino acid residues of positions 236 to 252. We propose that the second binding domain is located to the right of position 236 and down to around position 260 in Rep (region II) making a contact with the Ori DNA from the minor groove. A closer look at the results of RepC62 revealed weak but significant stimulation at 24A and 25G in site III on the bottom strand by binding of RepC62 (Fig. 4D). This suggested that the binding of RepC62 to the sites I and II induced destabilization of dsDNA in site III approximately one turn apart.

The results presented above also allowed us to narrow the region in Rep containing the DNA binding domain interacting with site III of Ori, which is to the right of position 166 (tentative region III). To further delimit the region containing the third domain, we used two additional deletion Rep proteins with the deletion end points to the right of position 166, namely RepC94 and RepC117 (Fig. 1). The result with RepC94 was essentially similar to that with RepC62, except that an additional residue (20G) on the bottom strand between sites II and III was protected (Fig. 4E). This might result from interaction of the additional amino acids in RepC94 with the Ori DNA. RepC117 was capable of binding to site III in addition to the sites I and II and also of inducing a conformational change of dsDNA in and around site III (Fig. 4F), although the interaction with site III was weaker than RepC132 and the wild type Rep. This suggested that the region binding to site III of Ori is present to the right of position 180, and the presence of the region from position 166 to 180 stabilizes the interaction of Rep to the site III of Ori. A notable amino acid sequence GLGRN is present in this region (Fig. 1C), in which the first G and sequence GRN are strictly conserved in the Rep proteins of the ColE2-related plasmids (12). We constructed mutant Rep proteins with amino acid substitutions at these positions (G172E, G174W, R175Q, N176D) by random mutagenesis

using Mn^{2+} PCR. All these mutant Rep proteins were defective in initiating *in vivo* plasmid ColE2 DNA replication (data not shown), indicating the importance of the conserved residues (G-GRN) for the activity of Rep. The binding affinity of RepN176D to the Ori DNA was apparently similar to that of the intact Rep as examined by EMSA (Fig. 5A). The result of DMS protection assay using RepN176D (Fig. 5B) was similar to that using the wild type Rep or RepC132 (see Fig. 4A1 and B1), except that the protection of the bases on the bottom strand at site III by RepN176D was hardly detectable. Note that the modification of bases on the bottom strand at site III of Ori was stimulated, when site III was not bound by Rep, as observed with RepC94 and RepC62 (Fig. 4D and E). RepN176D was also fully capable of inducing a conformational change of dsDNA at and around site III as revealed by stimulation of modification of bases by DMS on the top strand at site III. These results supported the notion that the presence of the region (positions 166 to 180) containing position 176 stabilizes the interaction of Rep with site III of Ori. Weaker protection on the bottom strand and stronger stimulation of some bases on the top strand at site III by RepN176D than by RepC117 (compare Fig. 5B and Fig. 4F1) might result from the presence of some region in the remaining N terminal half of Rep absent from RepC117 which further induces conformational change of dsDNA at and around site III of Ori.

A potential linker region connecting the binding domains of Rep. In order to analyze the structures and functions of various regions of Rep, we constructed mutant Rep proteins with single amino acid substitutions all through from the N terminus to the C terminus by random mutagenesis and tested their replication activity (details to be described elsewhere). We obtained many mutant Rep proteins defective in replication activity, which were located in the regions important for the origin binding (Fig. 1C). These amino acids seem to be involved directly in the origin binding activity of Rep or in formation of the proper structures of these regions. On the other hand, we noticed a region (around positions 191 to 211), where almost all the mutant Rep proteins obtained (14 mutants at 13 positions) retained the full replication activity, although in some of them the properties of amino acids were significantly changed by mutations (Fig. 1C). These results suggested that the original amino acid sequence in this region is not important for the Rep activity. We propose that this region might function as a

linker connecting the two regions (domains) of Rep involved in the Ori binding. Furthermore, RepC94 with a deletion end in the putative linker region protected the residue (20G) on the bottom strand between sites II and III (Fig. 4E), suggesting a possibility that the putative linker region crosses over the major groove between sites II and III. The presence of a putative linker region (around positions 190 to 211) further delimits region III containing the third DBD between positions 166 and 190.

Interaction of region III of Rep with site III of Ori. To further analyze the interaction of region III of Rep with site III of Ori, we used a deletion mutant origin (Oridr8; Fig. 1A and Table 2) lacking site III in EMSA (Fig. 6A). RepC132 bound to Oridr8 much less efficiently than to the intact origin (Fig. 6A, lane 3 and 2). RepC117, however, bound to Oridr8 and the intact origin roughly with a similar affinity, which might be due to unstable binding of RepC117 to site III of Ori. The DMS protection assay revealed that RepC117 interacted with Oridr8 much less efficiently than with the intact origin (data not shown). In contrast to the protection signals at three sites of the intact origin (see Fig. 4F), only site I of Oridr8 was protected weakly. Both RepC94 and RepC62 bound to the intact origin and Oridr8 with a similar affinity (Fig. 6A). Note that they protected only sites I and II of Ori in DMS protection assay (Fig. 4D and E). These results suggested a possibility that region III of Rep form an independent functional binding domain and interacts with site III of Ori. The results with RepC94 and RepC62 suggested that when region III of Rep is missing, the interaction of the deletion Rep proteins with the sites I and II of Ori is not affected by the presence or absence of site III. The additional amino acid sequence in RepC117 might interfere with the interaction of regions I and II of Rep with sites I and II of Ori, when site III of Ori is missing.

The sequence to the right of the site III of the minimal region of Ori. The deletion mutants of Ori lacking the sequence to the right of site III (Oridr56 and Oridr4) failed to initiate DNA replication (38). Nevertheless they retained similar and slightly weaker IncB activity as compared with the intact origin, suggesting that these two deletion origins are capable of binding to Rep as efficiently as the intact origin. We do not know, however, whether they interact with Rep in the same mode as that of the intact origin, and some difference in the binding mode could affect the initiation of DNA replication at the origin. To

address this problem, we used RepC132 having a similar binding activity to that of the intact Rep protein (Fig. 4A and B) but showing higher solubility (data not shown) and supercoiled plasmid DNA containing Oridr56 and Oridr4 in DMS protection assays (Fig. 5C and D). Both of them gave similar results to that of the intact origin DNA (see Fig. 4B1). Namely three sites (I, II and III) were protected and a conformational change of dsDNA at and around site III was detected. These results suggested that the sequence to the right of site III of Ori does not affect the binding mode of Rep and Ori and we propose that the sequence to right of site III is essential for the initiation of replication after the step of binding of Rep to Ori. The nucleotides 26T and 27G on the bottom strand next to site III of Oridr4 were also strongly modified by DMS, due to the presence of G instead of C at position 27. This guanine residue along with the neighboring thymine residue was protected by the binding of RepC132, supporting the notion that Rep interacts with site III of the origin DNA from the major groove.

DISCUSSION

We identified three regions in the C terminal region of Rep important for the Rep-Ori interaction, two of them (regions I and II) interact with the origin DNA from the adjacent major (site I) and minor (site II) grooves in the left half of Ori and the remaining region III interacts with site III next to the template sequence for the primer synthesis from the major groove, which is one and a half turn apart from site I on the opposite surface of the origin DNA (Fig. 7A and B). We propose that the interaction of region III of Rep with site III of Ori positions the primase domain in the N terminal region of Rep properly on the primer template. Sites I, II and III identified by the DMS protection assay are included within the three Rep protein binding elements (sites a, b and c, respectively) proposed by *in vivo* analyses of mutant origins with deletions and single-base-pair substitutions (38). We also identified a potential linker region located between regions II and III of Rep by the mutation analysis, which is proposed to cross over the major groove between sites II and III of Ori. Region I of Rep containing a potential HTH motif based on amino acid sequence homology to known *E. coli* DNA binding proteins is located at the C-terminus of Rep. We showed here that the region is essential for the specific binding of Rep to Ori. How regions II and III of Rep bind to Ori is yet unclear.

We found that ColE2 Rep protein has sequence homology and secondary structure similarity with the γδ resolvase (39), based on the alignment analysis (Fig 8A). Only the C terminal regions of these proteins were shown here. Although the entire sequence homology was very low, the binding regions I and II of the ColE2 Rep protein showed homology with the residues within the HTH motif and the extended arm region of the γδ resolvase important for the binding to the site I of the res DNA. The crystal structure of the γδ resolvase dimer complexed with the 34-bp site I DNA has been determined (39). Each γδ resolvase monomer has an N-terminal catalytic domain (positions 1-120), a 3-helix-bundle C-terminal DNA-binding domain (positions 148-183) that binds specifically to the recognition sequence from the major groove, and an extended arm region (positions 121-147) that connects the two domains and interacts with DNA from the adjacent minor groove. The arm region contains an extended turn and a helix. It is worth mentioning that the left half region of the res DNA site I

has a sequence similarity to the left half region of the plasmid ColE2 origin (Fig. 8B). An intriguing possibility is that the C-terminal DNA binding domains of the ColE2 Rep protein (region I containing a putative H-T-H motif and region II containing a turn and an α helix based on the secondary structure prediction) might adopt DNA-binding modes related to those of the monomer of the γδ resolvase (a 3-helix-bundle DBD and an extended arm region). A GRK sequence in the turn region (extended turn region) of the γδ resolvase is also found in the putative turn region of ColE2 Rep (Fig. 8A). G and R are directly involved in binding to the res DNA site I from the minor groove. Similarly the GRK sequence in ColE2 Rep might interact with the ColE2 origin DNA from the minor groove. A closer look at the result of the DMS protection assay with RepC45 which contains the GRK sequence at the N-terminal end revealed that 16A in site II appeared to be protected though very weakly (Fig. 4C).

The presence of the region from position 166 to 180 within region III stabilizes interaction of Rep with site III of Ori. Interestingly, the GLGRN sequence in this region important for the Rep activity as shown above, which is conserved in the Rep proteins of the ColE2-related plasmids (12), shows a homology to the RLGRD sequence in the catalytic domain of γδ resolvase (Fig. 8A). The second arginine in the RLGRD of the monomer of the γδ resolvase is involved in catalysis and is directly hydrogen bonded to the DNA backbone (39). Similarly the arginine in the GLGRN sequence of ColE2 Rep might interact with the DNA backbone to stabilize the Rep-Ori binding.

RepC117 lacking the region from position 166 to 180 bound to the three sites of Ori like the intact Rep protein, although the protection of site III of Ori by RepC117 was weaker than that by the intact Rep protein in DMS protection assay (Fig. 4A and 4F). The origin melting activity by RepC117 was very low compared with that by the intact Rep protein (8). This raised a possibility that the region (positions 166 to 180) is important for the *in vivo* replication of the plasmid ColE2, not merely in stabilizing the interaction of Rep with dsDNA of Ori as a part of the third binding domain. We speculate that this region might stabilize the open complex structure by binding to single stranded DNA of the melted region, which is important for progress of the primer synthesis by Rep.

The effective lengths of binding regions on DNA where DNA binding proteins interact

depend upon the structures, sizes and numbers of the DNA binding domains, and also upon whether the proteins interact as monomers, dimers or oligomers (6). In most plasmid replicons like the iteron-regulated plasmids (F, pSC101, pPS10, miniP1, etc.), the Rep proteins consist of about 300 amino acids and the Rep-binding sites in the origins are about 20 bp long (2, 5). The Rep54 monomer of plasmid F has extensive polar interactions between the two recognition helices of two DNA-binding domains and two adjacent major grooves on one surface of the iteron DNA (20). The Rep monomers of pPS10, pSC101 and miniP1 were suggested to have similar structures and might bind to their iterons similarly (31). The RepA monomer of pPS10 was suggested to have an extended form suitable for iteron binding by the biochemical analysis (4, 7). The Rep binding site in the ColE2 origin is about 20 bp long, based on the DMS protection assay, which is relatively long for the 132-amino acid C-terminal region of ColE2 Rep, if we assume a compact globular form. Therefore, a rather extended conformation of the C-terminal region of ColE2 Rep appears to be likely. We have recently found that Rep binds to Ori as a monomer (8). The nucleotide sequences of sites I, IIand III in the ColE2 origin are asymmetric. This would require an asymmetric conformation of the DBDs of Rep, and a monomer with three separate DBDs would be the simplest form of Rep suitable for this requirement.

Linker regions covalently connecting two DNA-binding domains in single proteins are found in the prokaryotic transcriptional activators of the AraC family. In MarA which belongs to the AraC family, two helix domains are linked by an α-helix that appears to provide the flexibility necessary for the two domains to interact with two successive major grooves of the DNA (28). Such linker regions in DNA binding proteins are more commonly found in eukaryotic systems. In the mammalian transcription factors such as Oct-1 and POU homeodomain proteins, for example, two DBDs are connected by a linker region, which serves as a flexible tether between the two domains and permits different relative arrangements of the two domains suitable for the binding sites on the DNA (30). The potential linker region of ColE2 Rep is predicted to form an α helix (Fig. 1B) and has sequence homology with the helix 2 of the Iβ subdomain of the MuA end binding domain (Iβγ) (Fig. 9). The Iβ subdomain can be divided into two structural elements, and the helix2 serves to bridge

these two elements (29). The putative linker region of ColE2 Rep might form a similar structure and link two DNA-binding domains on both sides and allow Rep to bind to Ori properly.

The ColE2 Rep protein is a multifunctional initiator, although it is a small protein. The N-terminal half of Rep is unnecessary for the origin DNA recognition as shown in this paper, whereas it is indispensable for replication of the plasmid ColE2. Rep is a primase, which can synthesize a 3-nucleotide primer RNA specific for the ColE2 origin which is utilized by the host DNA polymerase I to initiate replication. Therefore the N-terminal region must contain the primase domain. We are interested in how Rep binds to Ori to perform its multiple functions. The findings in this paper should help us to understand the early steps of initiation of the plasmid ColE2 replication in more details by future analyses.

REFERENCES

1. **Bazaral, M., and D. R. Helinski.** 1968. Circular DNA forms of colicinogenic factors E1, E2 and E3 from Escherichia coli. J Mol Biol **36:**185-94.
2. **Betteridge, T., J. Yang, A. J. Pittard, and J. Praszkier.** 2003. Interaction of the initiator protein of an IncB plasmid with its origin of DNA replication. J Bacteriol **185:**2210-8.
3. **Chang, A. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J Bacteriol **134:**1141-56.
4. **Corbett, R. J., and R. S. Roche.** 1984. Use of high-speed size-exclusion chromatography for the study of protein folding and stability. Biochemistry **23:**1888-94.
5. **del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas.** 1998. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev **62:**434-64.
6. **Garvie, C. W., and C. Wolberger.** 2001. Recognition of specific DNA sequences. Mol Cell **8:**937-46.
7. **Giraldo, R., J. M. Andreu, and R. Diaz-Orejas.** 1998. Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. Embo J **17:**4511-26.
8. **Han, M., K. Aoki, M. Yagura, and T. Itoh. unpublished data.**
9. **Hashimoto-Gotoh, T., and K. N. Timmis.** 1981. Incompatibility properties of Col E1 and pMB1 derivative plasmids: random replication of multicopy replicons. Cell **23:**229-38.
10. **Helinski, D. R., A.E. Toukdarian, and R.P. Novick.** 1996. Replication control and other stable maintenance mechanism of plasmids, p. 2295-2324. *In* F.C. Neidhardt, Curtiss, R.,III, J. Ingraham, E.C.C.Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (eds), Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington D.C.

11. **Herschman, H. R., and D. R. Helinski.** 1967. Comparative study of the events associated with colicin induction. J Bacteriol **94:**691-9.

12. **Hiraga, S., T. Sugiyama, and T. Itoh.** 1994. Comparative analysis of the replicon regions of eleven ColE2-related plasmids. J Bacteriol **176:**7233-43.

13. **Horii, T., and T. Itoh.** 1988. Replication of ColE2 and ColE3 plasmids: the regions sufficient for autonomous replication. Mol Gen Genet **212:**225-31.

14. **Itoh, T., and T. Horii.** 1989. Replication of ColE2 and ColE3 plasmids: in vitro replication dependent on plasmid-coded proteins. Mol Gen Genet **219:**249-55.

15. **Jacob, F., S. Brenner, and F. Couzin,.** 1963. On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp Quant Biol **28:**329-348.

16. **Kaguni, J. M.** 1997. Escherichia coli DnaA protein: the replication initiator. Mol Cells **7:**145-57.

17. **Khan, S. A., and D. K. Chattoraj.** 1998. Initiation of DNA replication in phages and plasmids-a workshop summary. Plasmid **40:**1-11.

18. **Kido, M., H. Yasueda, and T. Itoh.** 1991. Identification of a plasmid-coded protein required for initiation of ColE2 DNA replication. Nucleic Acids Res **19:**2875-80.

19. **Kingsbury, D. T., and D. R. Helinski.** 1970. DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem Biophys Res Commun **41:**1538-44.

20. **Komori, H., F. Matsunaga, Y. Higuchi, M. Ishiai, C. Wada, and K. Miki.** 1999. Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 A resolution. Embo J **18:**4597-607.

21. **Kornberg, A., and T. Baker.** 1992. DNA replicarion. 2nd edn. W.H.Freeman, New York, NY.

22. **Matsui, H., and T. Itoh. unpublished data.**

23. **Matsumoto, K., and T. Itoh. unpublished data.**

24. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc Natl Acad Sci U S A **74:**560-4.

25. **Melnikova, A. F., R. Beabealashvilli, and A. D. Mirzabekov.** 1978. A study of unwinding of DNA and shielding of the DNA grooves by RNA polymerase by using methylation with dimethylsulphate. Eur J Biochem **84:**301-9.

26. **Messer, W., and C. Weigel.** 1996. Initiation of chromosome replication, p. 1579-1601. *In* F.C. Neidhardt, Curtiss, R.,III, J. Ingraham, E.C.C.Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (eds), Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington D.C.

27. **Rao, R. N., and S. G. Rogers.** 1979. Plasmid pKC7: a vector containing ten restriction endonuclease sites suitable for cloning DNA segments. Gene **7:**79-82.

28. **Rhee, S., R. G. Martin, J. L. Rosner, and D. R. Davies.** 1998. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc Natl Acad Sci U S A **95:**10413-8.

29. **Schumacher, S., R. T. Clubb, M. Cai, K. Mizuuchi, G. M. Clore, and A. M. Gronenborn.** 1997. Solution structure of the Mu end DNA-binding ibeta subdomain of phage Mu transposase: modular DNA recognition by two tethered domains. Embo J **16:**7532-41.

30. **Scully, K. M., E. M. Jacobson, K. Jepsen, V. Lunyak, H. Viadiu, C. Carriere, D. W. Rose, F. Hooshmand, A. K. Aggarwal, and M. G. Rosenfeld.** 2000. Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. Science **290:**1127-31.

31. **Sharma, S., B. K. Sathyanarayana, J. G. Bird, J. R. Hoskins, B. Lee, and S. Wickner.** 2004. Plasmid P1 RepA is homologous to the F plasmid RepE class of initiators. J Biol Chem **279:**6027-34.

32. **Shinohara, M., and T. Itoh.** 1996. Specificity determinants in interaction of the initiator (Rep) proteins with the origins in the plasmids ColE2-P9 and ColE3-CA38 identified by chimera analysis. J Mol Biol **257:**290-300.

33. **Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk.** 1985. Measurement of protein using bicinchoninic acid. Anal Biochem **150:**76-85.

34. **Tacon, W., and D. Sherratt.** 1976. ColE plasmid replication in DNA polymerase I-deficient strains of Escherichia coli. Mol Gen Genet **147:**331-5.

35. **Takechi, S., and T. Itoh.** 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. Nucleic Acids Res **23:**4196-201.

36. **Takechi, S., H. Matsui, and T. Itoh.** 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. Embo J **14:**5141-7.

37. **Tomizawa, J., and T. Itoh.** 1981. Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. Proc Natl Acad Sci U S A **78:**6096-100.

38. **Yagura, M., S. Y. Nishio, H. Kurozumi, C. F. Wang, and T. Itoh.** 2006. Anatomy of the replication origin of plasmid ColE2-P9. J Bacteriol **188:**999-1010.

39. **Yang, W., and T. A. Steitz.** 1995. Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. Cell **82:**193-207.

40. **Yasueda, H., T. Horii, and T. Itoh.** 1989. Structural and functional organization of ColE2 and ColE3 replicons. Mol Gen Genet **215:**209-16.

Figure legends

FIG. 1. The origin and Rep protein of plasmid ColE2. (A) The nucleotide sequence of the origin of the plasmid ColE2-P9. Positions are numbered from left to right. Sites α and β (indicated by arrows below the sequences) are the specificity-determinants of the Rep-Ori interaction between ColE2 and ColE3 (34, 40). Subregions I, II and III of the 31-bp minimal region identified and sites a, b and c proposed as the Rep protein binding elements by mutation analyses (40) are indicated below the sequence. Sites I, II and III of the ColE2 origin bound by Rep identified by DMS protection assay in this study are boxed. The positions of the ends of deletions from the right (dr8, dr4 and dr56; Table2) are indicated by hooked arrows. (B) Schematic representations of the wild type and partial fragments of Rep (solid bars) used. Rep fragments with deletions from the N terminus were named with the numbers of the remaining residues in the C terminal regions. RepΔC20 is a stop codon mutant Rep protein lacking 20 residues in the C terminal region. The position numbers of the residues at the N and C termini are indicated. (C) The C terminal 147-amino acid region of the ColE2 Rep protein important for binding to the origin DNA. The putative HTH motif showing sequence homology with *E. coli* transcription factors (14) and regions A and B involved in determining the specificity of the Rep-Ori interaction in the plasmids ColE2-P9 and ColE3-CA38 (34) are indicated by double-headed arrows below and above the sequence, respectively. The positions of the deletion ends of the Rep fragments are indicated by hooked arrows. The single-amino-acid substitutions in mutant Rep proteins which retained or lost the in vivo activity to initiate DNA replication are indicated above or bellow the sequence, respectively. The predicted α helix regions (Psi-Pred) are indicated by solid bars bellow the sequence. Regions I, II and III of the Rep protein important for binding to the origin DNA identified in this study are boxed and a putative linker region is indicated by a bracket above the sequence.

FIG. 2. Specific Ori DNA binding of the ColE2 Rep protein and its partial fragments and mutant Rep proteins analyzed by EMSA. Various concentrations of peptides indicated at the top of each panel were incubated with end-labeled DNA fragments with (Ori DNA) or without (Ori-free DNA) the origin sequence and the DNA-protein complexes and free DNA

fragments were analyzed by electrophoresis on a 8% polyacrylamide gel. Partial Rep fragments with deletions from the N terminus (A; the positions of DNA-protein complexes are indicated by black circles), Rep Δ C20 (B; crude extracts containing these proteins were used), mutant Rep protein at position 285 (C) and mutant Rep proteins at positions 284, 286 and 287 (D) were compared with the wild type Rep protein. The additional shifted bands with faster mobilities were occasionally observed depending on the preparations of the Rep proteins used. They were more often observed with the proteins of less purity.

FIG. 3. The amounts of the ColE2 Rep protein and the mutant Rep proteins with a T7 tag in uninduced BL21 (DE3) cells carrying plasmids with the cloned genes encoding these proteins. (A) The total *E. coli* proteins in BL21 (DE3) cells were analyzed by 15% SDS-PAGE. (B) A half amount of each sample used in (A) was loaded on the SDS-PAGE and the western blot analysis was carried out with anti-T7 tag antibody.

FIG. 4. DMS footprints of ColE2 Rep and its partial fragments bound to Ori DNA. Supercoiled plasmid DNA carrying Ori was incubated with or without the intact ColE2 Rep protein or its fragments and treated with DMS as described in Materials and Methods. The modified sites on the top and bottom strands were mapped by primer extension using FITC-labeled primers on a sequencing gel (left panels). Control sequencing ladders (A, G, T, C) were obtained by using the same primers. The 37-bp origin region is indicated with an arrow in each gel. Nucleotides protected by Rep against DMS modification are indicated by triangles and those showing increased DMS sensitivity upon Rep binding are indicated by circles. A1: no Rep (lane 1), 100nM Rep (lane 2) and 200nM Rep (lane 3); B1: no RepC132 (lane 1), 100nM RepC132 (lane 2); C1: no RepC45 (lane 1), 200nMRepC45 (lane 2) and 400nM Rep45 (lane 3); D1: no RepC61 (lane 1), 100nM RepC61 (lane 2) and 200nM RepC61 (lane 3); E1: no RepC94 (lane 1), 100nM RepC94 (lane 2) and 200nM RepC94 (lane 3); F1: no RepC117 (lane 1), 100nM RepC117 (lane 2) and 200nM RepC117 (lane 3). Scans of lanes with naked DNA and Rep-bound DNA are shown by gray and black lines, respectively (right panels). Quantification of the experiments shown in the sequencing gels

was performed by using NIH Image 1.62 software. A schematization of the 37-bp origin region boxed is presented and the three sites I, II and III bound by the wild type Rep are indicated by the gray background. The symbols are as described for the left panels. A2: lanes 1 and 3 of A1; B2: lanes 1 and 2 of B1; C2: lanes 1 and 3 of C1; D2: lanes 1 and 3 of D1; E2: lanes 1 and 3 of E1; F2: lanes 1 and 3 of F1.

FIG. 5. Specific Rep-Ori DNA binding of mutant Rep protein and mutant origins. (A) EMSA using the wild type and RepN176D was performed as in Fig. 3. (B) DMS protection assay with (+; 100nM) or without (–) RepN176D was performed as in Fig. 5. (C and D) DMS protection assay using supercoiled plasmid DNA carrying Oridr56 (C) and Oridr4 (D) and with (+) or without (-) RepC132, was performed as in FIG. 5.

FIG. 6. Specific binding of various Rep fragments to the wild type and Oridr8 analyzed by EMSA. (A) Various Rep fragments were incubated with the end-labeled wild type Ori and Oridr8 fragments. The concentration of each Rep fragment was adjusted at the level that was unable to saturate the origin DNA. The protein-DNA complexes (indicated by dots) and free DNA fragments were analyzed as in FIG. 3. (B) Schematic representations of various Rep fragments bound to the wild type Ori and Oridr8. Sites I, II and III of Ori are indicated below the boxes representing the binding sites of the origin region (shown by the lines). Regions I, II and III of Rep are indicated in the ovals representing the domains of Rep.

FIG. 7. Summary and schematic representation of the results of the DMS protection assay in Fig. 5 (A) and a model of Rep-Ori interaction (B). Rep-binding sites I, II and III (brackets) of Ori are shown above a B-form DNA double helix structure and regions I, II and III of Rep interacting with sites I, II and III of Ori are indicated by squares.

FIG. 8. Comparisons of the colE2 Rep protein and the γδ resolvase and of their cognate binding sites. (A) Alignment of the amino acid sequences of the ColE2 Rep and the γδ resolvase. Identical residues are indicated between the two sequences. The predicted and

known secondary structures are shown above and bellow the amino acid sequences of the ColE2 Rep protein and the γδ resolvase, respectively: H, α helix; E, β strand; C, coil. The 3-helix-bundle DNA binding domain (containing a HTH motif) and the extended arm region are marked with double-headed arrows bellow the secondary structure of the γδ resolvase, and the helixE and HTH motif are shadowed. The three regions (I, II and III) of the ColE2 Rep protein important for the origin DNA binding deduced from the DMS footprint analysis and the putative linker region deduced from a genetic analysis are indicated above the secondary structure of the ColE2 Rep protein with double-headed arrows and a bracket, respectively and the putative HTH motif is shadowed. Among the residues in the HTH motif of the γδ resolvase involved in binding to site I of the res DNA, those involved directly in base-specific interaction from the major and minor grooves are indicated by outlined letters, while the other residues involved are underlined. (B) Alignment of the nucleotide sequences of the ColE2 origin DNA and the site I of the res DNA bound by the γδ resolvase. The regions with high homology are highlighted in gray and the identical residues are indicated by vertical lines. The primer sequence in the ColE2 origin is indicated by the arrow above the sequence. Twelve-bp inverted repeats in the site I of the res DNA are indicated by arrows bellow the sequence. The regions recognized and bound from the major and minor grooves by the C-terminal binding domains of the γδ resolvase are shown by brackets bellow the sequence, while the Rep binding sites I and II of the ColE2 origin identified in this study are shown by brackets above the sequence.

FIG. 9. Alignment of the putative linker region of the ColE2 Rep protein and the helix 2 of the Iβ subdomain of the MuA transposase. The putative linker region of the ColE2 Rep protein deduced from the mutation analyses in this paper is indicated by a bracket above the sequence. The predicted α helix region of the ColE2 Rep protein and the known α helix region (helix2) of the Iβ subdomain are indicated by solid bars.

TABLE 1. Oligonucleotides used for mutagenesis

Name	Sequence (5'-3')
KM2F	AAATCCGCGGAACGGGCAC
KM2R	GGACCTTTTCGAACAGATGG
KM3F	TAGCGTCGACAAAAATTACG
KM3R	GTGGTGGTGCTCAGATCT
d203	TGCCGTGATCCAGCGTGTC
d236	TACACACAGGAAATTCTCA
d260	TAAGCGCGCAGCAGTTCCT
KA1	TTGCGAGATCTGGATCACATTTTAA- TTTTYGGTAGTACSWCGCTCGAC

Table 2. Mutant Origins from the right

Name	Sequence
	1 10 20 30 40
ColE2 Ori 5'-	AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCCCT
dr56	AAAATGAGACCAGATAAGCCTTATCAGATAACAtCGatga
dr4	AAAATGAGACCAGATAAGCCTTATCAcAagcttatcgatg
dr8	AAAATGAGACCAGATAAGCCcaAgCttATcgatGataagc

Residues in deletion mutants different from those in the intact ColE2 Ori are shown by lowercase letters.

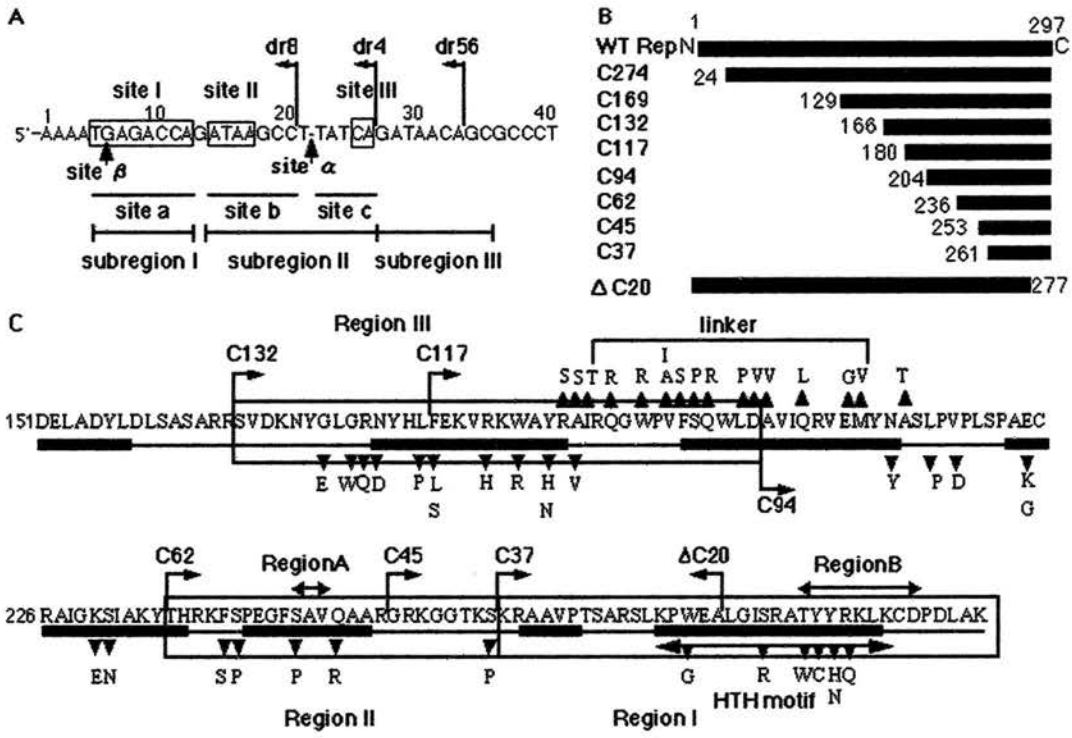

A
dr8
dr4
dr56
site I
site II
site III
1
10
20
30
40
5'-AAAATGAGACCAGATAAGCCT-TATCAGATAACAGCGCCCT
site β
site α
site a
site b
site c
subregion I
subregion II
subregion III
B
1
297
WT RepN
C
C274
24
C169
129
C132
166
C117
180
C94
204
C62
236
C45
253
C37
261
ΔC20
277
C
Region III
linker
C132
C117
I
SSTR R ASPR PVV L GV T
151DELADYLDLSASARFSVDKNYGLGRNYHLFEKVRKWAYRAIRQGWPVFSQWLDAVIQRVEMYNASLPVPLSPAEC
E WQD PL H R H V
S
N
C94
Y P D K
G
C62
RegionA
C45
C37
ΔC20
RegionB
226 RAIGKSIAKYTHRKFSPEGFSAVQAARGRKGGTKSKRAAVPTSARSLKPWEALGISRATYYRKLKCDPDLAK
EN
SP P R
P
G R WCHQ
N
HTH motif
Region II
Region I

Fig. 1

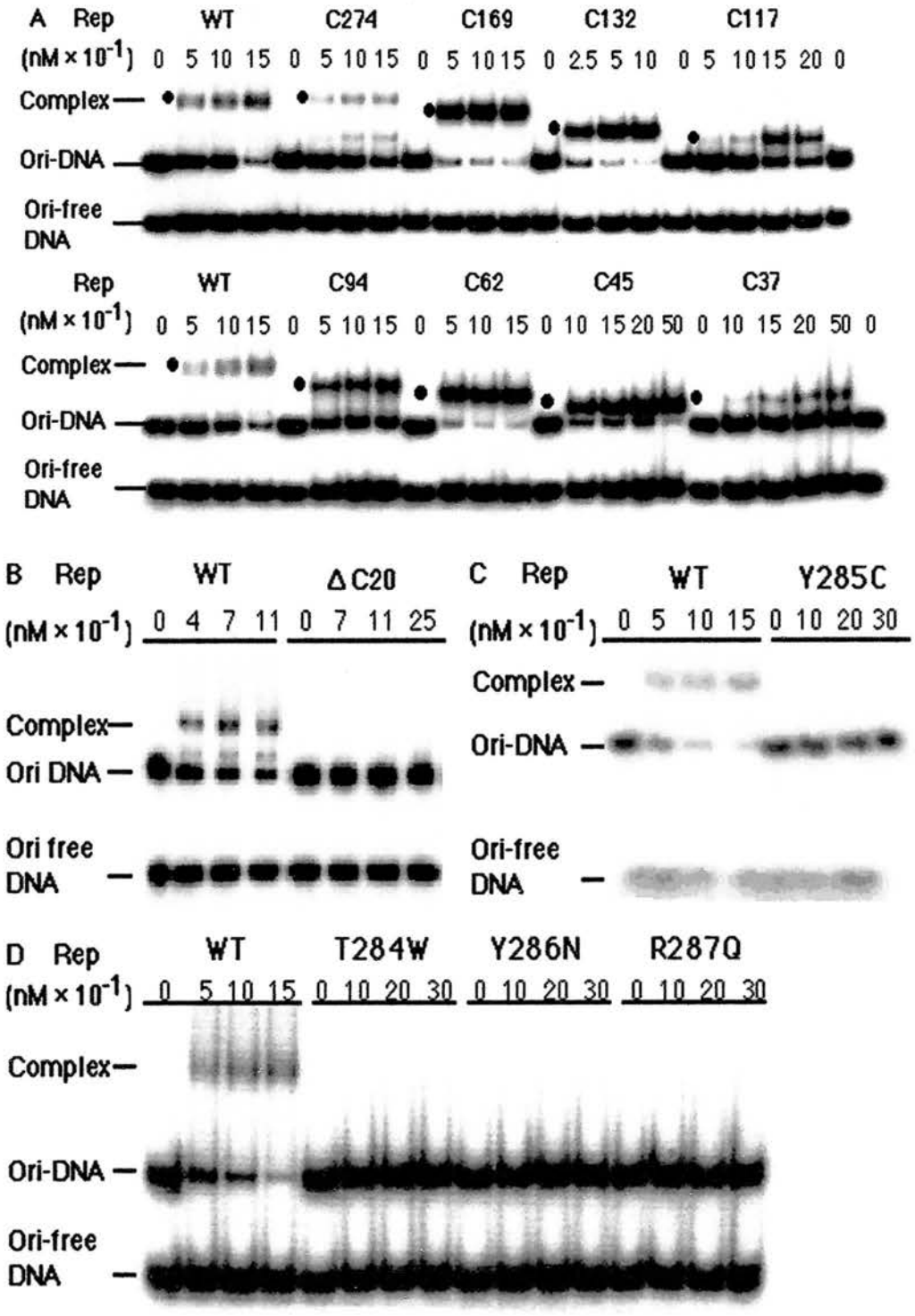

A Rep
WT
C274
C169
C132
C117
(nM × 10⁻¹) 0 5 10 15 0 5 10 15 0 5 10 15 0 2.5 5 10 0 5 10 15 20 0
Complex
Ori-DNA
Ori-free DNA
Rep
WT
C94
C62
C45
C37
(nM × 10⁻¹) 0 5 10 15 0 5 10 15 0 5 10 15 0 10 15 20 50 0 10 15 20 50 0
Complex
Ori-DNA
Ori-free DNA
B Rep
WT
ΔC20
(nM × 10⁻¹) 0 4 7 11 0 7 11 25
Complex
Ori DNA
Ori free DNA
C Rep
WT
Y285C
(nM × 10⁻¹) 0 5 10 15 0 10 20 30
Complex
Ori-DNA
Ori-free DNA
D Rep
WT
T284W
Y286N
R287Q
(nM × 10⁻¹) 0 5 10 15 0 10 20 30 0 10 20 30 0 10 20 30
Complex
Ori-DNA
Ori-free DNA

Fig. 2

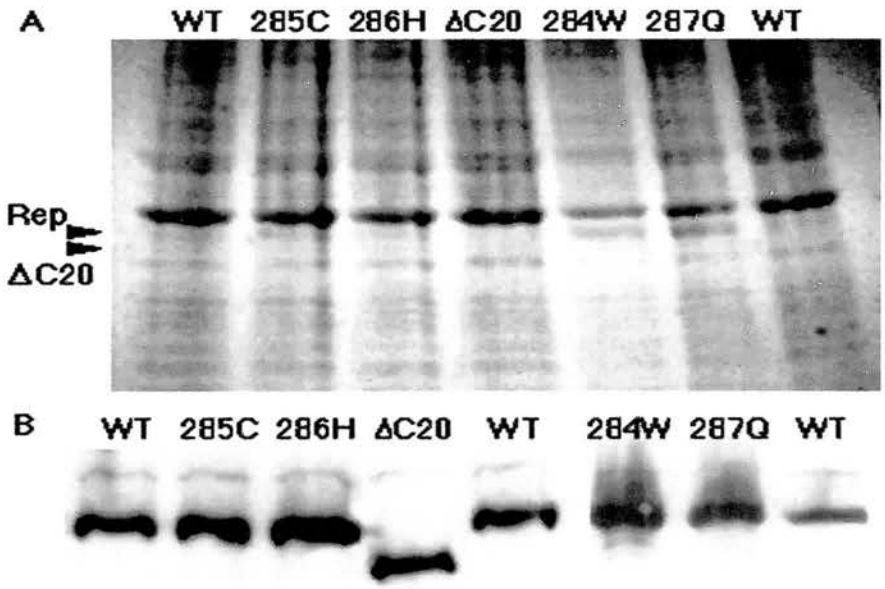
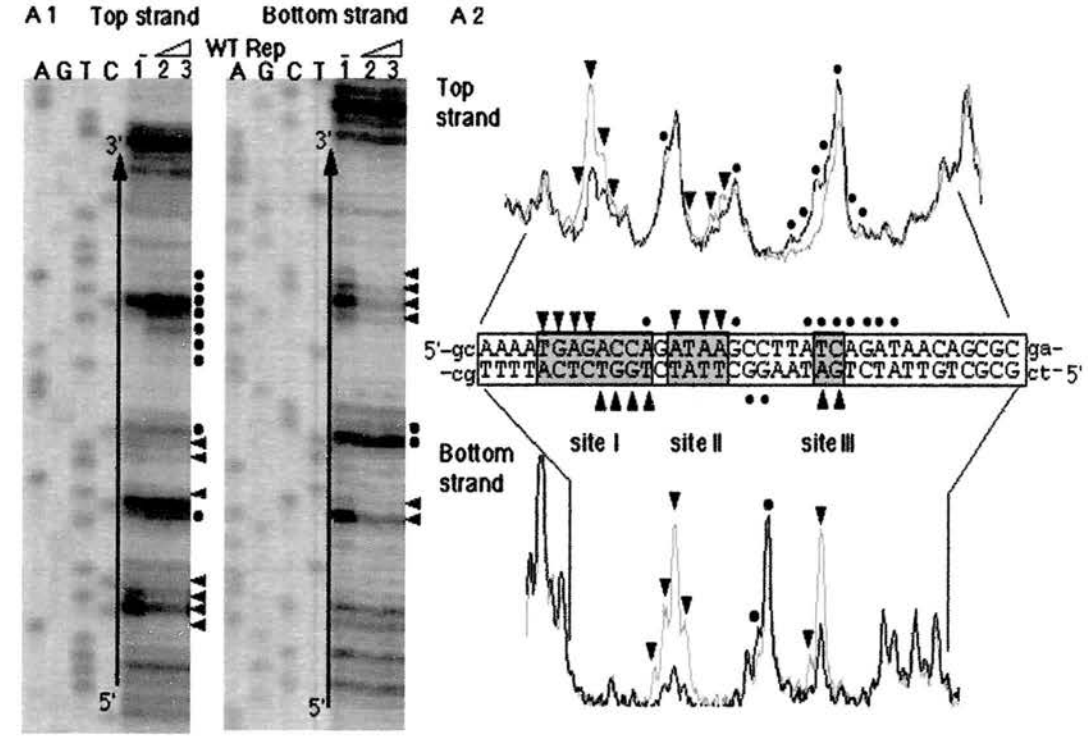
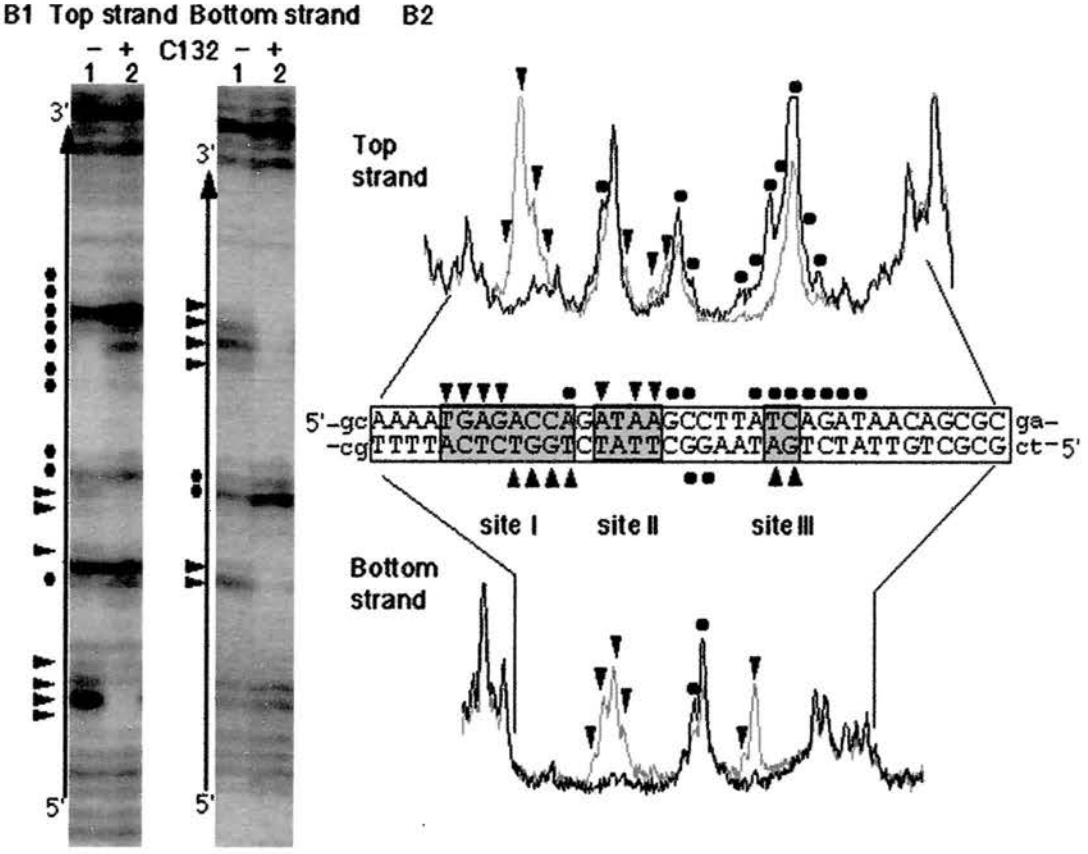
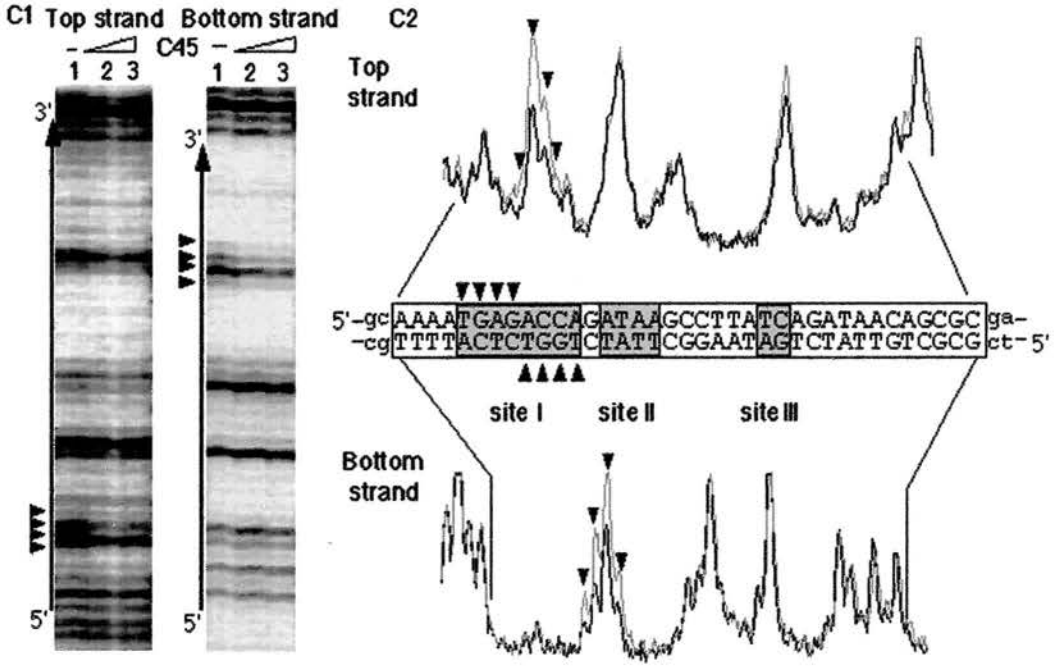
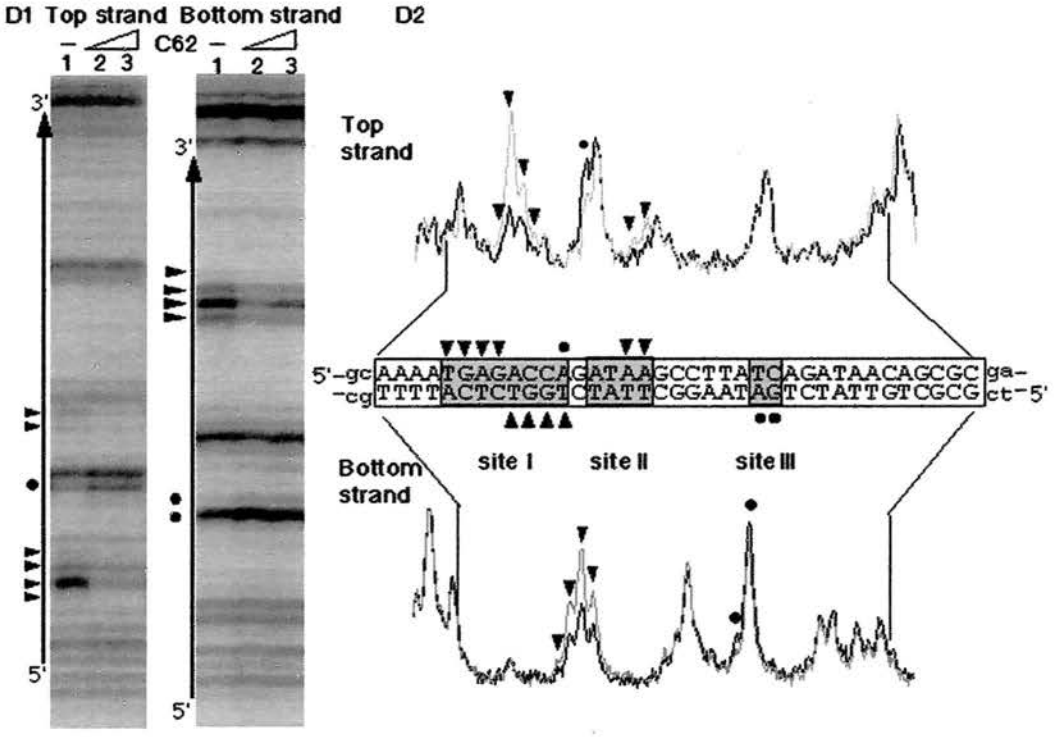

A
WT 285C 286H ΔC20 284W 287Q WT
Rep
ΔC20
B
WT 285C 286H ΔC20 WT 284W 287Q WT

Fig. 3


A 1
Top strand
Bottom strand
A 2
WT Rep
A G T C 1 2 3
A G C T 1 2 3
3'
5'
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand


B1 Top strand Bottom strand
B2
C132
– + – +
1 2 1 2
3'
5'
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand


C1 Top strand Bottom strand
C45
1 2 3
1 2 3
3'
5'
C2
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand


D1 Top strand Bottom strand
C62
1 2 3
1 2 3
3'
5'
D2
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand

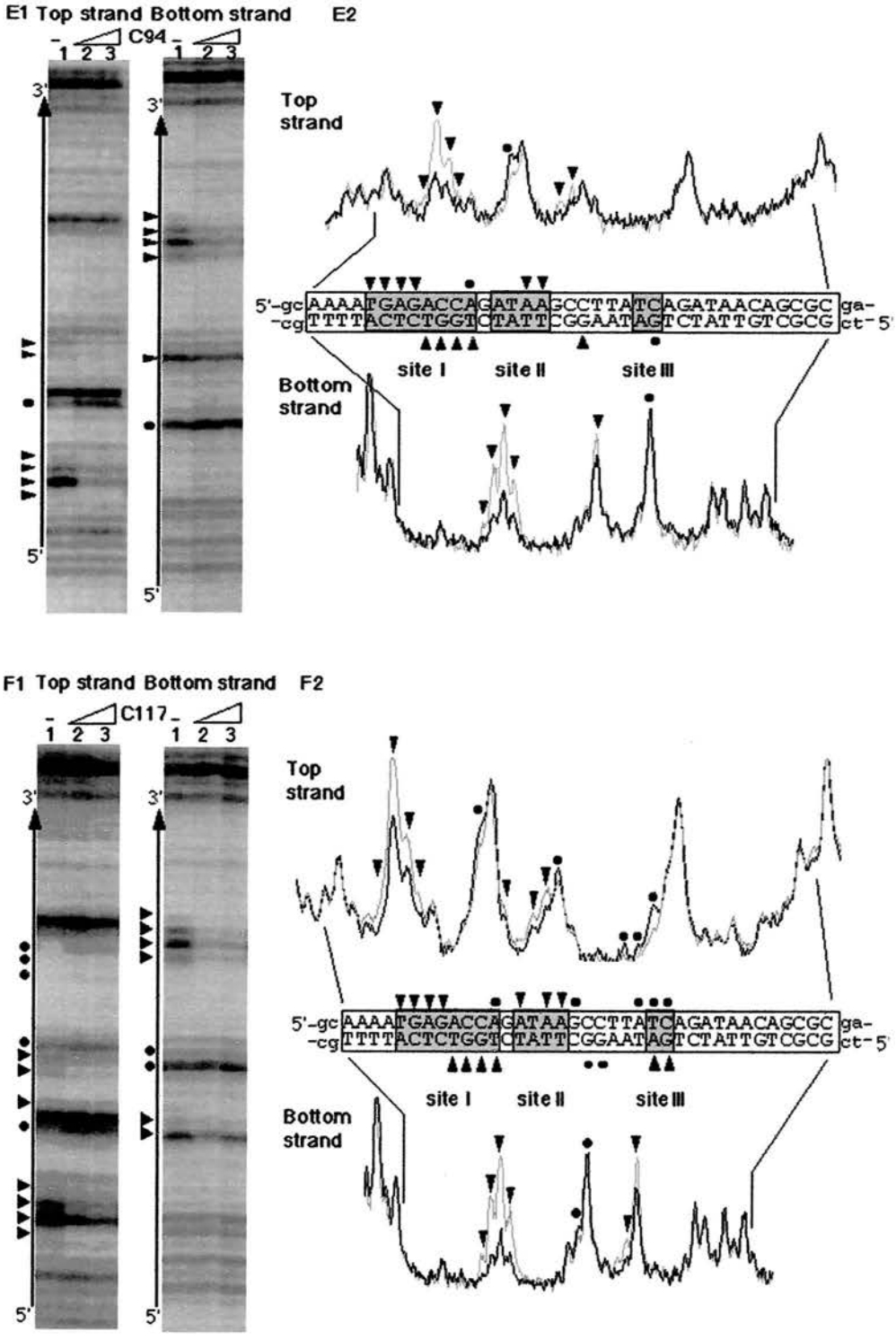

E1 Top strand Bottom strand
E2
C94
1 2 3
1 2 3
3'
5'
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand
F1 Top strand Bottom strand
F2
C117
1 2 3
1 2 3
3'
5'
Top strand
5'-gcAAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCga-
-cgTTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCGct-5'
site I
site II
site III
Bottom strand

Fig. 4

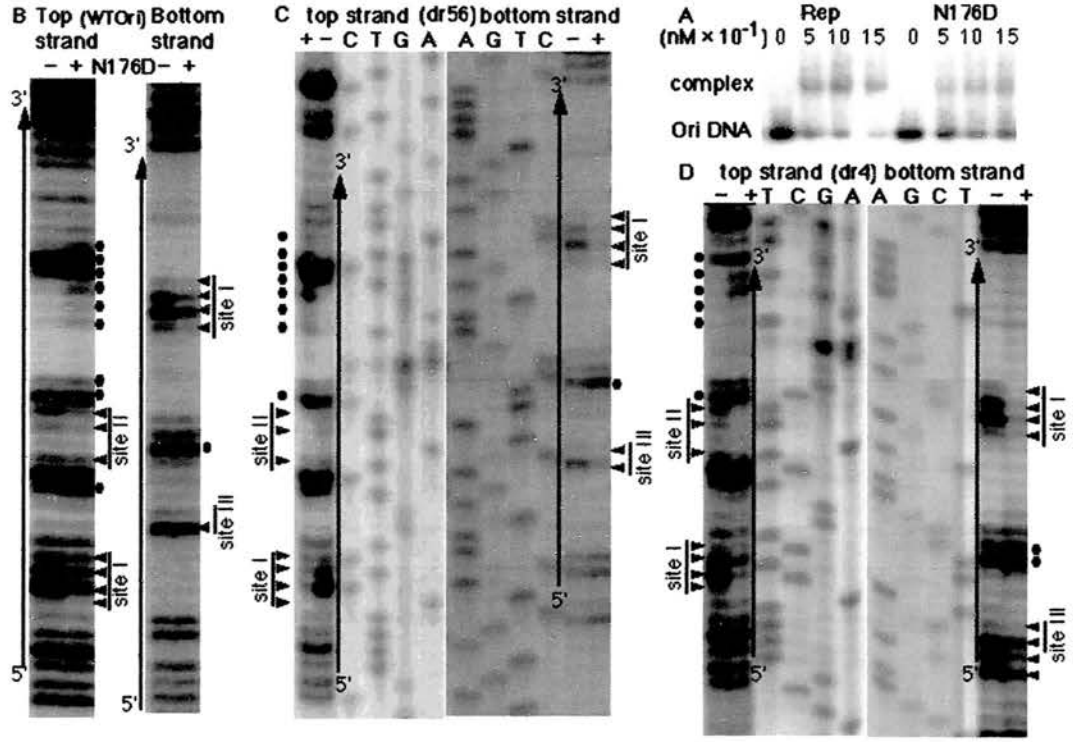

B Top (WTOri) Bottom
strand strand
− + N176D− +
3'
5'
site I
site II
site III
C top strand (dr56) bottom strand
+ − C T G A A G T C − +
A
(nM × 10⁻¹) 0 5 10 15 0 5 10 15
Rep
N176D
complex
Ori DNA
D top strand (dr4) bottom strand
− +T C G A A G C T − +

Fig. 5

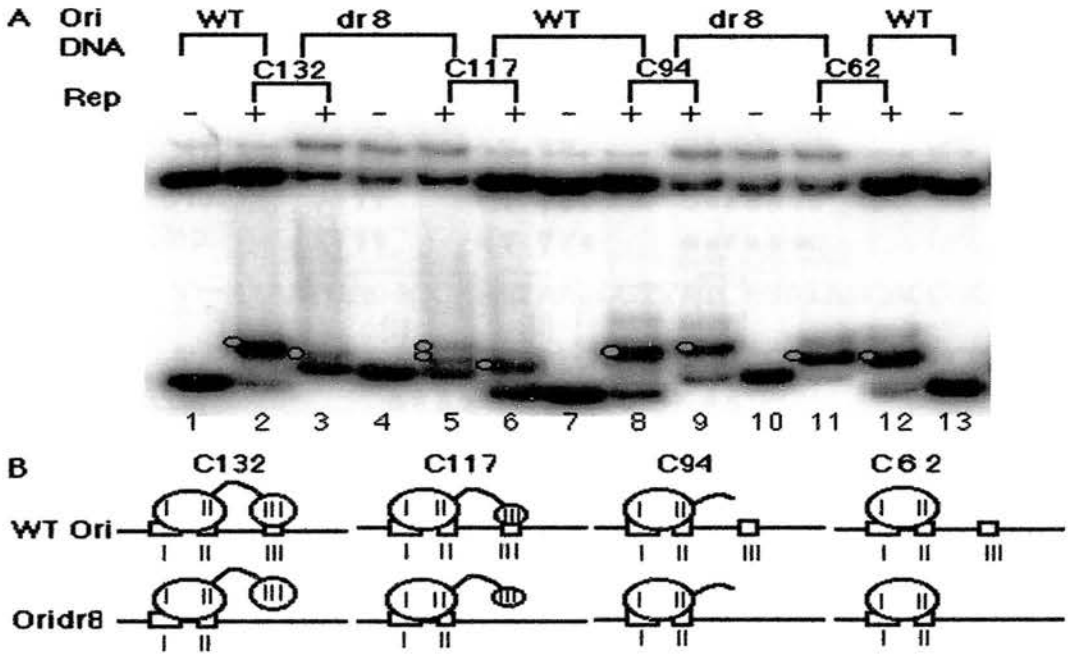

A
Ori
DNA
Rep
WT
dr8
WT
dr8
WT
C132
C117
C94
C62
1 2 3 4 5 6 7 8 9 10 11 12 13
B
C132
C117
C94
C62
WT Ori
Oridr8

Fig. 6

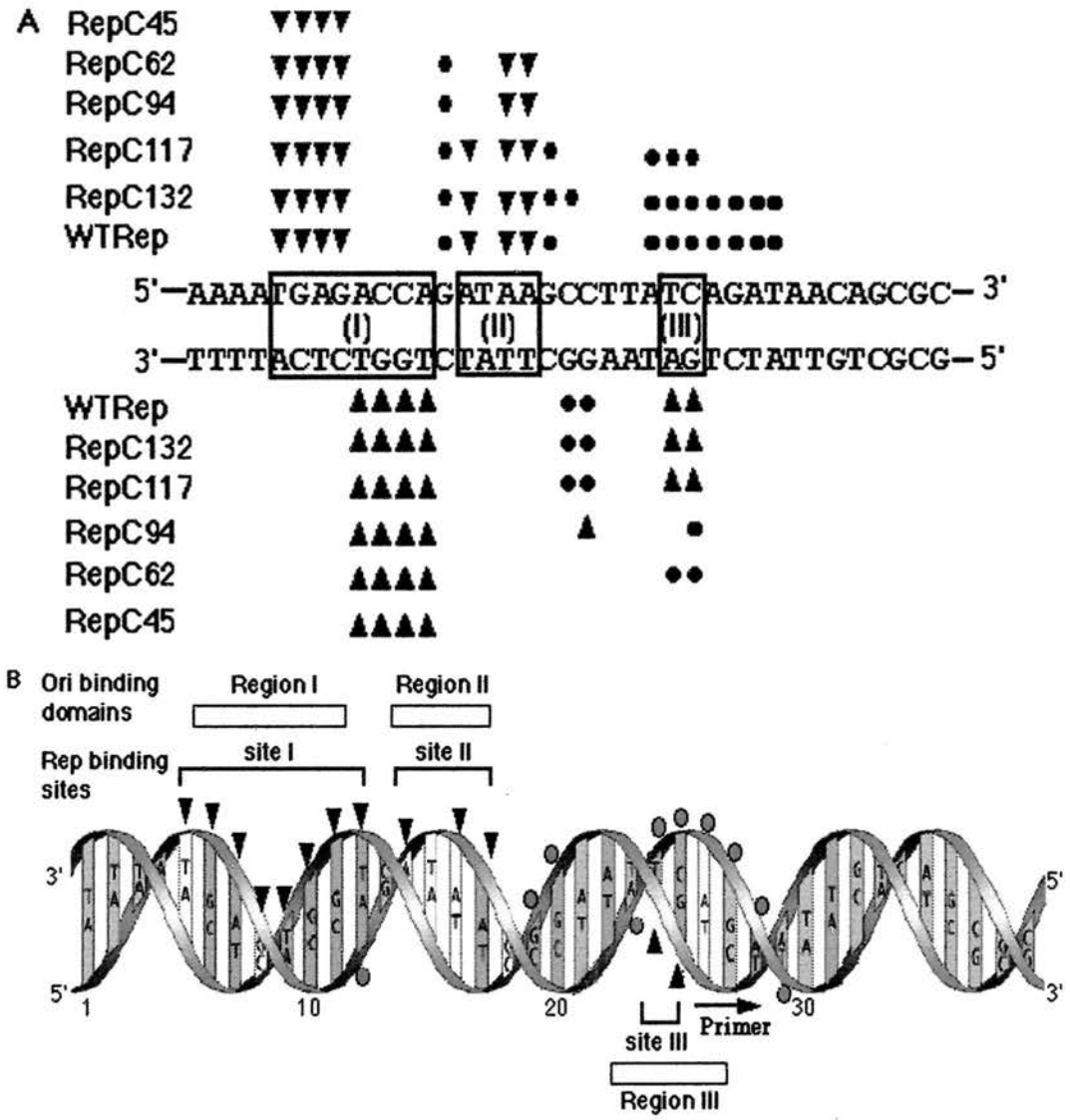

A
RepC45
RepC62
RepC94
RepC117
RepC132
WTRep
5'—AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGC— 3'
(I)
(II)
(III)
3'—TTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCG— 5'
WTRep
RepC132
RepC117
RepC94
RepC62
RepC45
B
Ori binding domains
Region I
Region II
Rep binding sites
site I
site II
3'
5'
1
10
20
30
5'
3'
site III
Primer
Region III

Fig. 7

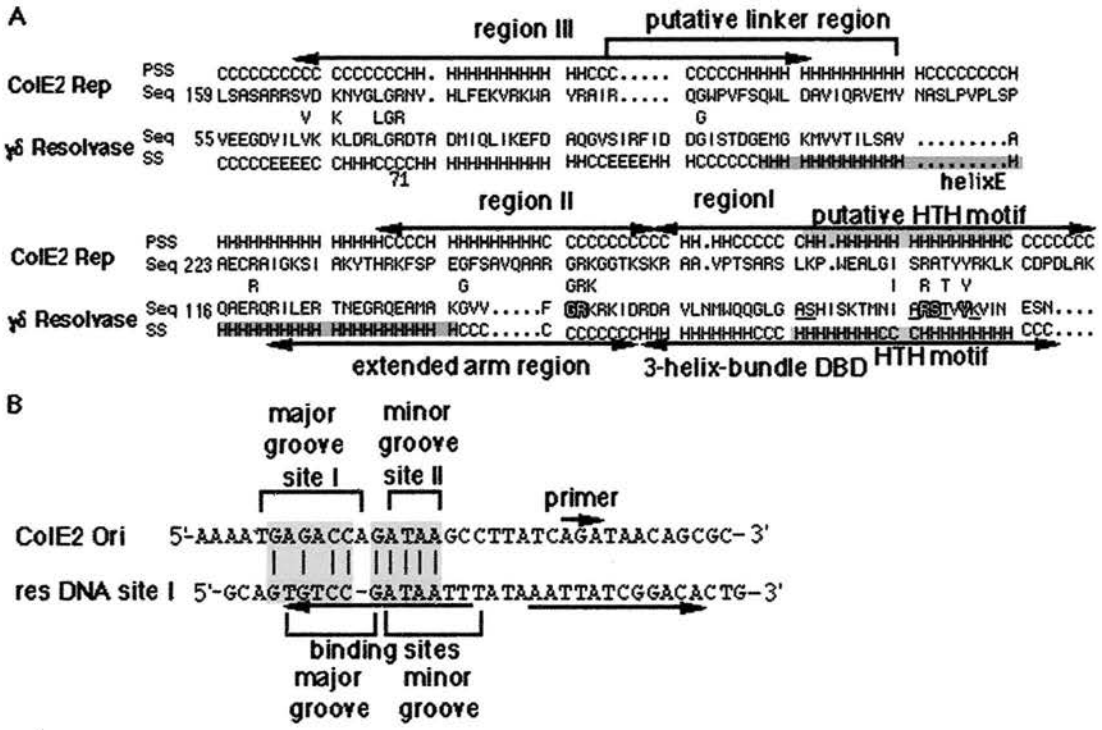

A
region III
putative linker region
ColE2 Rep
γδ Resolvase
region II
regionI
putative HTH motif
helixE
extended arm region
3-helix-bundle DBD
HTH motif
B
major groove
minor groove
site I
site II
primer
ColE2 Ori
res DNA site I
binding sites
major groove
minor groove

Fig. 8

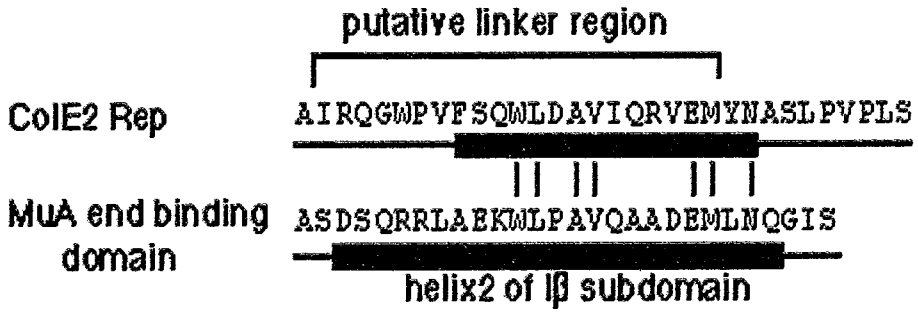

putative linker region
ColE2 Rep
AIRQGWPVFSQWLDAVIQRVEMYNASLPVPLS
MuA end binding domain
ASDSQRRLAEKWLPAVQAADEMLNQGIS
helix2 of Iβ subdomain

Fig. 9

IV Characterization of the region important for dimerization of ColE2-P9 Rep

Abstract

The ColE2 Rep proteins are present mostly in a dimeric form with some multimers larger than dimers in solution, while they bind to Ori as monomers. Detailed analyses of the interaction of Rep and Ori using various mutant Rep proteins were done, and three regions (I, II and III) in the C terminal region of the ColE2 Rep protein important for binding to the origin were identified. However, The regions important for dimerization of Rep proteins is not known. Here I identified the region important for dimerization of Rep. It was located within the area containing regions II and III important for Rep-Ori interaction. This raised a possibility that dimerization of Rep affected the interaction of regions II and III of Rep with Ori and this might explain why Rep binds to Ori as a monomer but not as a dimer.

MATERIALS AND METHODS

Strain. The *E. coli* strains used was BL21(DE3) (Novagen).

Construction of plasmids. To construct pRepΔC10+Km expressing the Rep proteins without His_6-tag (RepΔC10), the *BstE*II-*Sca*I fragment of pETE2+T7+site RIICm (M. Yagura, unpublished data) containing chloramphenicol resistance gene was cloned between the *BstE*II-*Sca*I sites of pACYC177 (1), followed by selfligation after digesting with *Pst*I and filled with T4 DNA polymerase, then the *Bsiw*I-*Bgl*II fragment of this plasmid containing the chloramphenicol resistance gene was replaced with the PCR fragments (amplified with oligonucleotides KM1F and KM3R from the template DNA of pRepΔC10+Ap) digested with *Bsiw*I and *Bgl*II enzymes.

Purification of Rep proteins. The proteins were purified from BL21 (DE3) cells essentially as described (2), except for the step of expression. Cells were grown for 3 hr at 37°C in 2 ml terrific broth (3) supplemented with 20 μg/ml kanamycin for the cells expressing the RepΔC10 proteins alone and with 50 μg/ml ampicillin and 20 μg/ml kanamicin for the cells coexpressing the proteins of RepΔC10 and Rep fragments or of wild type Rep and mutant Rep with a stop codon. The culture was diluted to 1 to 2×10^7 cells/ml with 50 ml fresh medium, grown at 25°C to 1×10^8 cells/ml and treated with 0.4 mM IPTG for additional 4hr.

Western blot analysis. Samples for Western blotting were separated in parallel on a 15% SDS-PAGE and proteins were transferred to a polyvinyliden difluoride sheet (PVDF; Atto). The sheet was treated with the anti-T7 tag antibody (Novagen) at a 1:8000 dilution for 1h and then with NBT/BCIP (Roche) at room temperature as recommended by the supplier (Novagen manual).

Results and discussion

To identify the region important for dimerization of Rep proteins, the Rep protein without His_6-tag (RepΔC10) or with a stop codon (RepΔC39) was coexpressed with a series of partial Rep fragments lacking the N terminal portions of the Rep protein and with His_6-tag (Fig. 1 A). RepΔC10 lacking the C terminal 10-amino-acid region of Rep alone did not specifically bind to the NTA-Ni resin (Fig. 1 B, lane 3). However, when it was coexpressed with RepC117 and purified from NTA-Ni resin, it was obtained from the elution (Fig. 1 B, lane 6). This suggested that the N-terminal region to the left of position 180 and the C-terminal region to the right of position 287 (Fig. 1 A) of Rep are unnecessary for dimerization of Rep proteins. To delimit the dimerization region of Rep, we used a deletion mutant Rep protein named RepΔC39 lacking the C terminal 39-amino-acid region of Rep and coexpresssed it with wild type Rep. After purified from NTA-Ni resin, RepΔC39 was also obtaited from the elution (Fig. 1 B, lane 9). This suggested that the region to the right of position 258 (Fig. 1 A) was unnecessary for dimerization of Rep proteins. Based on the results above, I conclude that the region important for dimerization of Rep proteins is located to the right of position 180 and down to position 258.

The region important for dimerization of Rep proteins is located within the area containing regions II and III important for Rep-Ori interaction (Fig. 2). This raised a possibility that dimerization of Rep affected the interaction of regions II and III of Rep with Ori and this might explain why Rep binds to Ori as a monomer but not as a dimer.

The deletion mutant Rep lacking the N terminal portions of the Rep protein with the deletion end in this region like RepC94 and RepC36 constructed before can be used to delimit this region further by future analyses.

References

1. **Chang, A. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J Bacteriol **134:**1141-56.

2. **Yagura, M., S. Y. Nishio, H. Kurozumi, C. F. Wang, and T. Itoh.** 2006. Anatomy of the replication origin of plasmid ColE2-P9. J Bacteriol **188:**999-1010.

3. **Takechi, S., and T. Itoh.** 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. Nucleic Acids Res **23:**4196-201.

Figure legends

Fig. 1 Determination of the regions important for dimerization of Rep proteins. A. Schematic representations of the wild type and partial fragments of Rep (solid bars) used. B. The western blot analysis was carried out with anti-T7 tag antibody. Total solution (lane 1, 4, 7); Supernatant (lane 2, 5, 8); Elution (lane 3, 6, 9).

Fig. 2 Schematic representation of the wild type Rep protein. Regions I, II and III important for binding to Ori identified by DMS protection assay are indicated by gray bars. A putative linker was indicated. The region important for dimerization of Rep is indicated by a two-headed arrow.

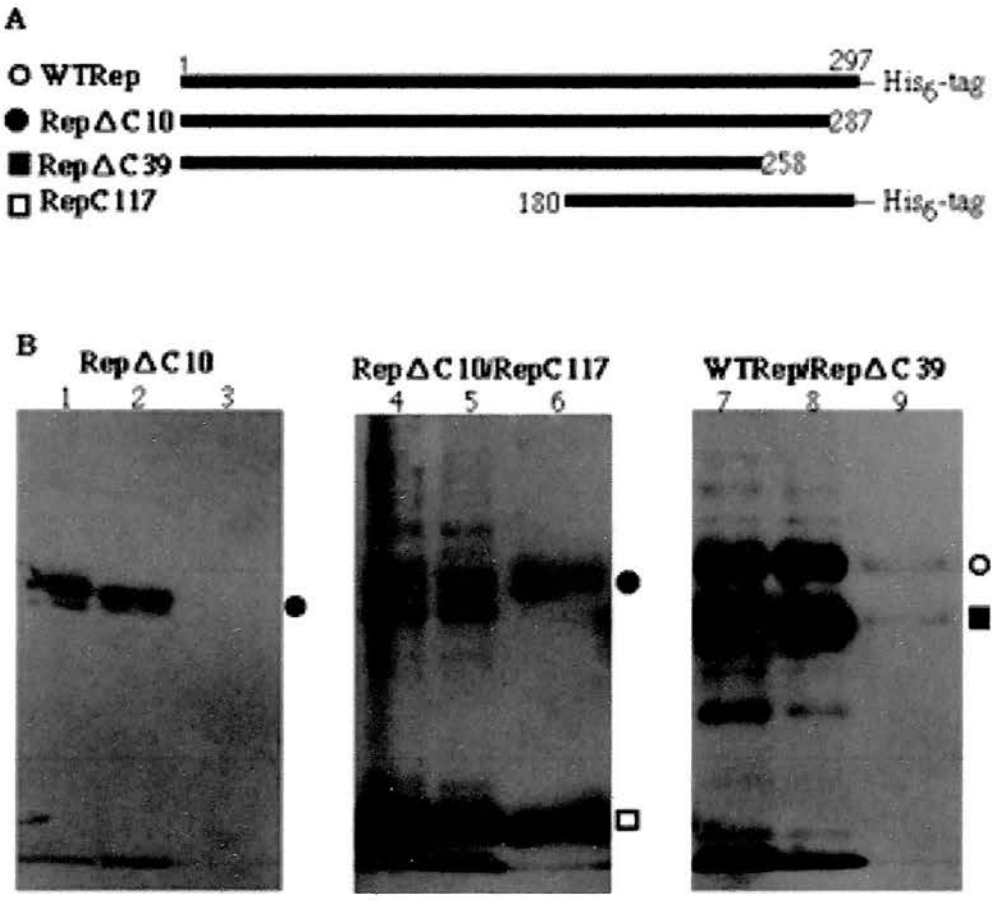

A
WTRep
RepΔC10
RepΔC39
RepC117
1
297
287
258
180
His6-tag
His6-tag
B
RepΔC10
RepΔC10/RepC117
WTRep/RepΔC39
1 2 3
4 5 6
7 8 9

Fig. 1

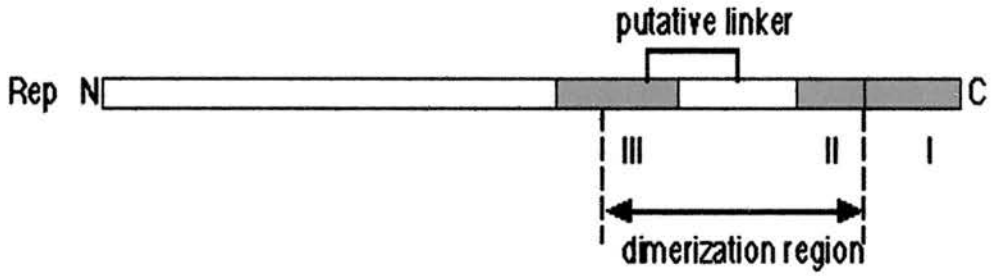


Fig. 2

V Mechanism of origin melting of ColE2-P9 plasmid

Abstract

Origin melting is essential for the initiation of DNA replication and requires binding of initiator proteins. For many bacterial plasmids, the melting of origin in DNA generally requires combined actions of plasmid encoded initiator protein and host factors. While origin melting of ColE2-P9 (ColE2) plasmid only requires its own initiator protein (Rep). Rep binds to Ori as a monomer and causes a weak bend (21°) at the origin DNA. Rep induces a localized melting of the origin DNA in a supercoiling-dependent way. The melting region lies in the right half of Ori and overlaps the binding site III for Rep and the template sequence for the primer synthesis. Rep binds specifically to the nontemplate strand of Ori, which is essential for full melting of origin. Based on the results presented in this paper, I proposed a model for the mechanism of origin melting.

Introduction

For various bacterial chromosomal and plasmid origins of DNA replication, the first step of initiation requires binding of the initiator protein that serves to melt the DNA duplex in an AT-rich region (1, 34). This process has been extensively studied in Escherichia coli. In the presence of HU or IHF protein, and a millimolar level of ATP, binding of the DnaA protein to the 9-mer boxes (called DnaA boxes) in OriC leads to unwinding of the AT-rich region containing the 13-mers (5, 9, 15), and then the ATP-bound DnaA proteins bind to the single-stranded 6-mer sequences 5'-AGATCT in the unwound region, stabilizing the single-stranded regions (39). The DnaA protein acts as a replisome organizer for the subsequent loading of proteins required for replication process (35). For many bacterial plasmids such as F (19, 47), R6K (23, 29), RK2 (22) and P1 (36), Initiation of DNA replication involves binding of the plasmid encoded initiator proteins to the iterons and forms an open complex at their cognate origins to recruit the replisome (7). The opening of the AT-rich segments of the origins generally requires combined actions of plasmid-encoded initiator protein and host factors containing DnaA, HU protein or both.

The plasmid ColE2-P9 (ColE2) is a circular duplex DNA molecule of ~7 kb (13). It is present in ~10-15 copies per host chromosome(2, 14). The initiator protein (Rep; 35 kDa) of plasmid ColE2 is the only plasmid-specified trans-acting factor required for the initiation of plasmid replication (20, 21, 45). Initiation of plasmid replication also requires host DNA polymerase I along with other factors (18, 40), but not RNA polymerase and DnaG primase (18, 41). The minimal origin (Ori) of the ColE2 DNA replication consists of 32 bp, where DNA replication initiates (44). The ColE2 Rep protein binds specifically to the origin, as revealed by a filter binding assay ((20)) and EMSA (44). In an *in vitro* ColE2 DNA replication system with crude extracts of E. coli cells ColE2 DNA replication starts at a unique site and proceeds unidirectionally (18, 41). The ColE2 Rep protein is unique among other initiator proteins in that it is also a plasmid-specific primase. It synthesizes a 3-nucleotide primer RNA at a specific site in the origin, which has a unique structure of 5' ppApGpA (42). Host DNA polymerase I specifically uses the primer RNA to start DNA synthesis, and then form a D-loop structure, into which various replication proteins of *E.coli* like DnaB helicase, DnaG

primase and DNA polymerase III holoenzyme, were introduced to continue replication of ColE2 DNA. The ColE2 Rep proteins are mainly present as dimers in solution (30) while they bind to Ori as monomers (Han, submitted). The ColE2 Rep protein has a potential HTH DNA binding motif based on amino acid sequence homology to some known *E. coli* DNA binding proteins (20). Detailed analyses of the interaction of Rep and Ori using various mutant Rep proteins were done (11), and three regions (I, II and III) in the C terminal region of the ColE2 Rep protein. important for binding to the origin were identified (Fig. 1 A). Two of them (regions I and II) interact with the origin DNA from the adjacent major (site I) and minor (site II) grooves in the left half of Ori and the remaining region III interacts with the site III from the major groove next to the template sequence for the primer synthesis, which is one and a half turn apart from site I on the opposite surface of the origin DNA (Fig. 1 B). Sites I, II and III identified by the DMS protection assay (11) are included within the three Rep protein binding elements (sites a, b and c, respectively) proposed by *in vivo* analyses of mutant origins with deletions and single-base-pair substitutions (44). A linker region connects the regions II and III and provides Rep an extended conformation suitable for Ori binding as a monomer. The interaction of region III of Rep with site III of Ori might be required to position the primase domain properly. Synthesis of the primer by the ColE2 Rep protein requires opening of the DNA in the origin to expose the template strand.

In this paper, I examined how the interaction of ColE2 Rep and Ori leads to opening of the origin DNA. I found that ColE2 Rep protein induces localized melting of the origin DNA around the primer sequence by itself, but only very weak bending of the origin is introduced by binding of Rep protein. I further found that Rep binds to the nontemplate strand of Ori specifically, which is important for full opening of the origin DNA. Adding these new data to what is already known about Rep-Ori interaction, I proposed a model for the mechanism of Origin melting of ColE2 plasmid.

Materials and Methods

Bacterial and plasmids. The *E.coli* K-12 strain used was BL21 (DE3) (Novagen). D110: rnh::Tn3 (polA1. end I ⁻. thy⁻). Plasmids used have been described elsewhere ((11)) except for those described below.

Construction of plasmid for bending assay. PbluescriptII KS+ was digested with *Pst*I(filled by T4 polymease) and *Sma*I, followed by selfligation to construct pBlueΔPS. The PCR fragment (amplified with –21M13, M13 oligonucleotides, and pBlueΔPS as a template) digested with *Kpn*I (filled by T4 polymerase) and *Sac*I was inserted between the sites *sac*II and *Sac*I of pBlueΔPS to construct pKA, which contains a new EndoR cleavage site (*Sma*I). pKAE2Ori used for bending assay was constructed by insertion of 47bp DNA fragment with ColE2 Ori in the *Sma*I site of pKA.

Construction of the plasmid for in vivo $KMnO_4$ footprint assay. The 2517bp *Ssp*I-*Ssp*I fragment of pACYC184 containing tetracycline gene was cloned between the *Ssp*I-*Ssp*I sites of pBlue22+wtOri (44) to construct pHME2Ori+tet. To costruct pACX43ΔWT (Iwayama and Itoh, unpublished) the 1.5-kb *BamH*I-*Hind*III fragment of pBX243 (46) carrying the ColE2 *rep* gene was inserted between the *BamH*I and *Hind*III sites (within the tetracycline-resistance gene) of pACYC184 (6). To construct pAKX43ΔWT (Matsumura and Itoh, unpublished) the 3.6-kb *EcoR*I-*Sal*I fragment of pACX43ΔWT carrying the p15A replicon and the ColE2 *rep* gene was ligated with the 1.5-kb *EcoR*I-*Sal*I fragment of pKC7 (37) carrying the kanamycin-resistance gene. PAKX43ΔWT was digested with *Bgl*II and *BamH*I, followed by selfligation to construct pAKX43ΔWT-Rep.

Enzymes, and chemicals. Chemicals, enzymes and antibiotics used were from commercial sources. [α-^{32}P] dATP (3,000 Ci/mmol), [α-^{32}P] dCTP (3,000 Ci/mmol) and [γ-^{32}P] ATP (3,000 Ci/mmol) were obtained from Amersham Biosiences.

Purification of Rep protein. The ColE2 Rep protein and its fragments with a His_6-tag at the C-terminus were purified from BL21 (DE3) cells carrying derivatives of pET21a(+) producing these peptides as described (11).

***In vitro* $KMnO_4$ Footprint assay.** Either supercoiled or linearized pBlue22+WTOri (0.1

pmol) was incubated for 15 min at RT with wild type or mutant Rep proteins in 27 μl of reaction mixtures containing the binding buffer (200mM Tris-HCl pH7.5, 10mM $MgCl_2$, 100mM KCl, 0.2mM EDTA, 10% ethylene glycol), 0.04 μg/μl BSA and 0.02 μg/μl salmon DNA. Reaction of modification was started by adding 3 μl of $KMnO_4$ (10nM) at room temperature. After 2 min, reactions were stopped by adding 30ul of stop solution (3M β-mercaptoethanol, 40mM EDTA, 0.6M sodium acetate) and ethanol precipitated. Modified nucleotides were mapped by primer extension as described (11).

***In vivo* $KMnO_4$ Footprinting.** pHME2Ori+tet with tetracycline gene and ColE2 Ori was introduced into D110 *polA*1 cells carrying either pAKX43ΔWT (with the kanamycin resistance gene and the *rep* gene) or pAKX43ΔRep (lacking the *rep* gene). Cells were plated onto LB agar medium containing kanamycin (20 μg/ml) and tetracycline (10 μg/ml). After incubation for 12h at 37°C, single colony were grown for 12h at 37°C in a 2 ml of LB midium supplemented with kanamycin (20 μg/ml) and tetracycline (10 μg/ml), diluted 100 fold with 50 ml of fresh medium, grown to $1\sim2\times10^8$cells/ml. The cells were collected by centrifugation at 4000rpm for 5 min at RT, resuspended in 5 ml of LB medium diluted 10 fold, followed by treatment with rifampicin (0.2 mg/ml) for 5 min prior to adding $KMnO_4$ (10mM). After incubation for 2 min at 37°C, the reaction was stopped by adding 100 μl of β-mercaptoethanol, the cells were harvested and plasmid DNA was isolated using alkaline lysis technique. Modified nucleotides were mapped by primer extention as described (11).

EMSA for bending assay. DNA fragments with (134 bp) or without (83 bp) origin were generated by digesting pKAE2Ori or pBlue22+wtori (44) with seven ristriction enzymes or *BssH*II and 3'-end labeling with [α-^{32}P]dATP or [α-^{32}P]dCTP, respectively, and filling with Klenow fragment at room temperature for 30 min. Labeled fragments were purified by PCR purification Kit (QIAGEN). Five nM labeled origin DNA fragment was incubated for 15 min at 25°C with 100 nM of wild type Rep or control DNA fragments without origin in 27 μl reaction mixtures containing the binding buffer (200 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 10% ethylene glycol), 40 μg/ml BSA and 20 μg/ml salmon DNA. Aliquots were loaded onto 8% polyacrylamide (with 1/75 ratio of bisacrylamide/acrylamide)-0.5×TBE gels (20×20 cm^2). Electrophoresis was run at 4°C at 420

V (constant voltage) after prerun for 10 min. DNA was visualized by using Fuji BAS1500 phosphorimager and image reader v 1.7J software.

Specific ssDNA binding by Rep and its fragments. Oligonucleotides with origin, 5'-gat ctc gca aaa tga gac cag ata agc ctt atc aga taa cag cgc cct ttt-3' (nontemplate strand) and 5'-aaa agg gcg ctg tta tct gat aag gct tat ctg gtc tca ttt tcg ctc tag-3' (template strand) were labeled with [γ-^{32}P] ATP (3,000 Ci/mmol) and T4 polynucleotide kinase at 37°C for 30 min. Ten nM labeled oligonucleotide was incubated with wild type Rep (0, 100,150 nM) or RepC132 (0, 25 , 50, 100nM) or RepC117 (0, 100, 200 nM), and the reaction and eletrophoresis were done as described above.

RESULTS AND DISCUSSION

ColE2 Rep protein melts the origin DNA on the localized region in a supercoiling-dependent way. I used $KMnO_4$ to probe for the origin DNA melting by the ColE2 Rep protein. $KMnO_4$ oxidizes thymines and to a lesser extent cytosines at unwound or sharply distorted DNA sites (3, 12, 38). The modified nucleotides can be detected easily by primer extension, because the oxidized residues are unable to be copied by DNA polymerases (16).

To investigate the role of Rep for origin DNA melting under a physiological condition, I first used supercoiled pblue22+wtori DNA. The plasmid DNA was incubated with or without purified Rep protein and subjected to $KMnO_4$ treatment (Fig. 2 and Fig.3C). The bases 21T (T at position 21), 22T, 24T, 25C and 29T on the top strandand and the bases 23T, 26T, 27C and 28T on the bottom strand became reactive to $KMnO_4$ oxidation upon the Rep protein addition. In addition, 23A, 26A, 28A and 30A on the top strand and 24A and 25G on the bottom strand seemed to be reactive to $KMnO_4$ oxidation. These sites are located in a region overlapping the sites of primer synthesis, suggesting that the region was unwound upon binding of Rep to Ori. DMS protection assay also supported this result (Fig.1, (11)). DMS modifies bases on dsDNA and ssDNA differently. It methylates the N3 position of adenine from the minor groove and the N7 position of guanine from the major groove on dsDNA substrates (31), whereas on ssDNA, it modifies the N1 position of adenine and, to a lesser extent, the N3 position of cytosine (32). In the presence of the Rep protein, the enhanced modification of the bases 23A, 25C, 26A and 28A on the top strand were detected (Fig. 1C) in the distorted region revealed by the $KMnO_4$ footprint assay. The absence of enhanced modification on the bottom strand in the corresponding region might due to the binding of Rep to the bottom strand of site III in this region. The results of the DMS protection assay together with the result of $KMnO_4$ modification described above indicated that Rep melts the origin DNA in a specific region.

The melting of the origin DNA by the initiator proteins generally depends on DNA supercoiling strictly (4, 48). To test whether DNA supercoiling is required for ColE2 origin DNA melting, we performed the $KMnO_4$ footprint using the linear origin DNA and purified Rep protein (Fig. 3A). Only the bases 24T and 25C on the top strand and 24A and 25G on the bottom strand became reactive to $KMnO_4$ oxidation in the presence of Rep and the signals

were very weak. This suggested origin DNA melting in and around the site of primer synthesis by Rep depends on DNA supercoiling.

Comparison of the origin DNA melting by binding of Rep *in vivo* and *in vitro*. For most of the initiator proteins, melting of the origin DNA requires coexistence of cofactors like IHF, HU and SSB (4, 48). I then performed the in vivo $KMnO_4$ footprint assay of the ColE2 origin DNA in living cells to compare with the result of $KMnO_4$ footprint with Rep alone. To accumulate plasmids at the initiation step, an *E. coli* host with a polA mutation was used. PolA-defective cells carrying pHME2Ori+tet with the plasmid ColE2 Ori and pAKX43ΔWT expressing the Rep proteins or pAKX43ΔWT-Rep, in which the *rep* gene was removed were incubated to a log phase and subjected to $KMnO_4$ treatment. Plasmid DNA was extracted and the bases sensitive to $KMnO_4$ on two strands were examined (Fig. 3 B and 3 C). The bases 21T, 22T, 24T and 25C on the top strand and the bases 23T, 27C, 28T, 30T and 31T on the bottom strand became reactive to $KMnO_4$ oxidation. In addition, 23A and 26A on the top strand and 20G, 24A, 25G, 29A and 32G also became reactive to $KMnO_4$ oxidation. The bases modified from position 28 to the position 30 on the top strand in the *in vitro* assay (Fig. 2 and Fig. 3C) were not detected in this *in vivo* assay. When the concentration of the Rep protein is decreased in the *in vitro* assay, the modified signals in the bases from the position 28 to the position 30 on the top strand also weakened or disappeared (data not shown). Therefore, I interpreted the *in vivo* result is due to the high plasmid copy number of pblue22+wtori which resulted in the high ratio of origin DNA not bound by Rep. I conclude that the bases modified on the top stand in the *in vivo* assay were almost identical with those in the *in vitro* assay. On the other hand, the modified signals on the bottom strand extended downstream by four bases to the position 32 in the *in vivo* assay compared with the *in vitro* assay. This extension might result from the primer synthesis by the Rep protein in the *in vivo* assay, which did not occur in the *in vitro* assay. Much weaker enhanced signals on the bottom strand than that of the top strand in the *in vivo* assay might be due to the Rep binding to the template DNA to synthesize the primer. These results suggested that there were likely no other factors stimulating the Rep mediated melting of the origin DNA.

Regions of Ori and Rep sufficient for melting of the origin DNA. To understand how

Rep binding to Ori induce melting of the origin DNA, I first examined the relative importance of portions of the origin sequence for KMnO4 sensitivity using supercoiled forms of four plasmids containing the intact origin or truncated origins (Fig. 4). Oridr56 and Oridr4 failed to initiate DNA replication of the origin, but retained the incB incompatibility activities (44), and bound by RepC132 with a similar affinity and mode to those of the intact origin (11). RepC132 was used in this experiment. The $KMnO_4$ oxidation patterns of Oridr56 and Oridr4 were identical to that obtained using the intact origin DNA (Fig. 4). These results confirmed the conclusion that they lost the activity to initiate DNA replication at the origin after the step of Rep-Ori interaction, and further suggested the sequence lacking in them was important for initiation of DNA replication after the step of Rep mediated DNA opening. Both Oridr56 and Oridr4 contained the site III of Ori, whereas another truncated origin named Oridr8 lacking the site III of Ori failed to melt the origin at the specific sites (Fig. 4). These results might suggest a possibility that binding to the site III by Rep is required for opening of origin DNA. However, I can not exclude the possibility that weak affinity of RepC132 and Oridr8 (11) resulted in failure of DNA opening of Oridr8.

I then used a series of mutant Rep proteins to examine their ability to melt the origin DNA (Fig. 2), whose affinity with the origin DNA had been examined (Fig. 1 C). RepC132 with the origin binding affinity similar to that of the wild type Rep protein showed similar modification signals against $KMnO_4$ to those by the intact Rep protein (Fig. 2), suggesting that RepC132 retain the origin melting activity as efficient as that of the wild type Rep protein. The melting activity of RepC117 decreased dramatically, although its origin binding activity of it was not so weak, and RepC94, RepC62 and RepC45 lacking region III of Rep which interact with site III of Ori (Fig. 1 A) were unable to melt the origin DNA (Fig. 2), even though RepC94 and RepC62 bound to the intact origin efficiently as revealed by EMSA (11). This supported the notion above that the interaction of region III of Rep with site III of Ori is important for the origin melting. Our results of the DMS footprint analysis (Fig. 1) also supported the results of $KMnO_4$ modification above. RepC132 showed enhanced signals by DMS modification similar to those with the intact Rep protein on the top strand in the melted region and the same region RepC117 exhibited weaker enhanced signals, whereas in the case

of RepC94, RepC61 and RepC45, no enhanced signals were detected in the melted region (Fig. 1 C). The difference of Ori melting level by RepC132 and RepC117 suggested that the region containing amino acid residues from positions 165 to 180 is required for efficient Ori melting.

The functional role of the interaction of region III of Rep and site III of Ori for origin melting. The results obtained above showed that the interaction of the region III of the Rep and the site III of Ori is important for melting the origin DNA. To gain further insight into the melting mechanism of the origin DNA, I tried to investigate the functional role of this interaction. Sequence specific DNA binding proteins frequently induce conformational change of DNA by bending at the site of the interaction (8). In some prokaryotic and eukaryotic genomes, the sharp bending of the origin may facilitate duplex opening (5). I then examined bending of the ColE2 origin DNA by Rep (Fig. 5). I used a set of seven DNA fragments with equal in lengths carrying the origin sequence at varying portions relative to the ends of the fragments (see Materials and methods). The electrophoretic mobilities of these fragments were then compared with and without the Rep protein. All the fragments used showed the same mobility in the absence of the Rep protein, indicating a lack of intrinsic curvature at the origin sequence (Fig. 5A). In the presence of the Rep protein, some complexes DNA fragments with the origin sequence nearer to the end than others migrated faster (Fig. 5B). The relative eletrophoretic mobility (μM/μE) was measured for *EcoRI/XbaI* and *EcoRI/XhoI* and calculated bending angle was 21.7° (±0.52°, three independent experiments). The bent zone was located just upstream from the site III of Ori (Fig. 5B). The results above indicate that the binding of the wild type Rep protein to the origin induced only a weak bending of the origin DNA.

I then examined whether the interaction of region III of Rep with site III of Ori induces bending of the origin DNA and which regions of Rep protein induces bending of origin DNA by using various partial fragments of Rep. As expected, RepC132 with the origin binding affinity similar to that of the intact Rep protein (Fig. 1 C) induced similar extent bending to that of the intact Rep protein (Fig. 5C). Unexpectedly RepC94 and RepC61 lacking region III of Rep and defective in binding to the site III of Ori (Fig. 1 A and 1 C) induced bending of the

origin DNA similar to that by the intact Rep protein (Fig. 5C). This suggested that the interaction with sites I and II of Ori by Rep is sufficient for inducing bending of the origin DNA. Moreover, this interaction induced destabilization of double stranded DNA at site III of Ori (stimulation of 24A and 25G) as shown by DMS protection assay (Fig. 1).

Based on the results above, I propose that the role of the interaction of region III of Rep and site III of Ori for origin melting is to stabilize the conformation change of the origin DNA strands containing site III caused by bending induced by interaction of the site I and site II of the origin by rep and induces the primary melting of ori DNA.

The Rep protein binds to the ssDNA of the origin. The opening of the origin DNA results in partial exposure of the ssDNA. I therefore examined the affinity of the Rep protein for the ssDNA of the origin by EMSA. The intact Rep protein bound specifically to the nontemplate strand of the origin DNA (Fig. 6A, Top), as it formed a specific complex in the presence of single strand DNA of M13 as the nonspecific competition. To identify the region of Rep important for specific binding to the nontemplate strand of the origin, various partial fragments of Rep were used for EMSA. Very similar results were obtained with RepC199, RepC169 (data not shown) and RepC132 as compared with the intact Rep protein (Fig. 6B, Top), however no complex was detected by using RepC117 (Fig. 6C, Top). These observations, along with the results reported before (11), that RepC117 bound to the sites I, II and III of Ori (Fig. 1 C), suggested that the region from positions 166 to 180 might be important for specific binding to the nontemplate strand of Ori. Thus the reason for the wilder opening induced by RepC132 than that induced by RepC117 might be due to the presence of the additional region in RepC132 which bound specifically to nontemplate strand of Ori to stabilize the opened DNA structure of Ori as well as the interaction of the region III of Rep and the site III of Ori as we proposed above. ColE2 Rep is a plasmid specific primase and it synthesizes a 3-nucleotide primer RNA (42). It is evident that a primase has to interact specifically with the DNA template to synthesize primers (27, 33). I then examined the interaction of Rep and the template strand of the origin DNA (Fig. 6). Contrary to our expectation, the wild type Rep bound specifically to template strand of the origin DNA with much less affinity than to nontemplate strand of the origin DNA (Fig. 6A). Such a weak

affinity might be enough for a primase, because a primase only needs to bind to template DNA for a short period of time, which is sufficient for synthesizing primers, and an interaction with too much strength might inhibit processive primer synthesis. Binding of Rep to the nontemplate strand of the origin DNA might possibly help to position the primase domain properly on the template strand to synthesize primers efficiently.

Possible role for the ssDNA binding activity of Rep. It seems to be common for the initiator proteins that stabilize the unwound region by binding to the single-stranded DNA like those of pAMβI (28), DnaA of *E.coli* (39), SV40 (43). However, the single-stranded DNA binding by the initiator is mainly nonspecific. ATP-bound DnaA of *E.coli* binds cooperatively and with sequence specificity to single-stranded DNA of AT-rich region of OriC (39). Binding of ATP-bound DnaA to top strand is more effective than to bottom strand. ColE2 Rep also binds specifically to nontemplate strand of Ori DNA, however, the binding affinity of Rep and template strand is very faint (Fig. 6), and unlike additional DnaA proteins of *E.coli* binding to ssDNA of melting region, the region in the same protein of ColE2 Rep (which binds to dsDNA of Ori) binds to ssDNA of Ori. ColE2 Rep is also a plasmid specific primase. All DNA primases, whether from bacteriophage, viral, prokaryotic, or eukarytic sources, share a conserved amino acid sequence, which is referred to as the zinc-binding motif (17, 25) The zinc motif in the T7 gene 4 protein is involved in template recognition (24, 26). However, no zinc-binding motif was detected in ColE2 Rep sequence. I proposed that Rep binding specifically to the nontemplate strand of Ori is important for stabilizing the unwound Ori DNA and positioning the N terminal region of Rep probably containing the primase domain closely to the template strand, since the binding affinity of Rep and template strand is very faint.

A model for origin melting. ColE2 Rep binds, bends and melts the origin in a supercoiling-dependent way, and only monomers of Rep bind to Ori (10), although they are mainly present as dimers in solution (Matsumoto, unpublished). The located melting region of Ori contains the primer sequence. Based on the observations in this paper, I proposed a model

for the mechanism of origin melting. Regions I and II of Rep in a single Rep protein bind to sites I and II of Ori and bends origin DNA, which induces the destabilization of the double-stranded DNA of site III of Ori. The interaction of Region III of the same Rep protein and site III of Ori is to stabilize the conformation change of site III and induces the primary melting of origin DNA in a supercoiling-dependent manner, then a region (from positions 166 to 180) within the region III of the same Rep protein binds to the nontemplate strand of Ori to induce full melting of Ori and helps to position the primase domain locating in the N terminal region of Rep on the template strand properly to synthesize a primer stably. The finding in this study should provide insights into the mechanism of DNA replication initiation of ColE2-P9 and ColE2-related plasmids.

.

Figure legends

Fig. 1. Three binding regions (I, II and III) of Rep and the corresponding binding sites (I, II and III) of Ori important for Rep-Ori interaction. A. Schematic representation of the wild type and partial C-terminal fragments of Rep. The positions of the deletion ends of Rep fragments are indicated by hooked arrows. Regions I, II and III important for binding to Ori identified by DMS protection assay (han et al, in press) are indicated by grey bars. A putative linker was indicated. (B) Rep-binding sites I, II and III (brackets) of Ori identified by DMS protection assay (Han, In press) are shown on a B-form DNA double helix structure and regions I, II and III of Rep interacting with sites I, II and III of Ori are indicated by gray squares. The site of the primer synthesis is also indicated by an arrow. C. Summary and schematic representation of DMS protection assays (Han et al, In press) of Rep protein and its partial fragments on the ColE2 origin. A schematization of 37bp origin region is presented and three sites I, II and III bound by Rep are indicated by double-headed arrows. Nucleotides protected by Rep against DMS modification are indicated by triangles and those showing increased DMS sensitivity upon Rep binding are indicated by circles.

Fig. 2. Reactivity of ColE2 Ori to $KMnO_4$ in the presence of the wild type and partial fragments of Rep (In vitro). Supercoiled plasmid DNA carrying Ori was incubated with (WT, 100nM, 200nM; C132, 25nM, 50nM; C117, 100nM, 200nM; C94, 200nM; C62, 200nM) or without Rep and its partial fragments and treated with $KMnO_4$ (10 nM) as described in Materials and Methods. Reactive bases on the top and bottom strands were mapped by primer extension using FITC-labeled primers on a sequencing gel. Control sequencing ladders (A. G. C. T) were obtained by using the same primers. The 37-bp Ori region was indicated with an arrow in each gel.

Fig. 3. In vivo and in vitro $KMnO_4$ sensitivity of ColE2 origin. A. In vitro, linear DNA containing Ori was incubated with (200nM) or without Rep and treated with $KMnO_4$ (10nM) as described in Materials and Methods. B. In vivo, growing D110 cells of deficient in polI (*polA*1) carrying pAKX43ΔWT (with the rep gene) or a derivative pAKX43ΔWT-Rep

(without the rep gene) were treated with $KMnO_4$ (10nM) as described in Materials and Methods. Modified bases on the top and bottom strands of A and B are mapped as in Fig.2 and indicated by short arrows. The 37-bp origin region is indicated with an arrow in each gel. C. Summary and schematic representation of $KMnO_4$ sensitivity of ColE2 origin in the presence of the wild type and its partial fragments of Rep. Nucleotides showing increased $KMnO_4$ sensitivity upon Rep binding are indicated by arrows. The primer sequence was indicated by short arrows. Three sites I, II and III important for Rep binding were indicated with two-headed arrows.

Fig. 4. $KMnO_4$ oxidation patterns of truncated origins (Top strand). A. Truncated origin DNA (dr56 ordr4) containing three binding sites of Ori for Rep or lacking binding site III of Ori (dr8) were incubated with RepC132 (50nM) and treated with $KMnO_4$ (10nM). Modified bases on the top strand were mapped as in Fig. 2. B. A comparison of the $KMnO_4$ modification for the wild type and the truncated origins is represented schematically. Modified bases were indicated by black bars.

Fig. 5. Origin DNA bending assay. Set of DNA fragments of similar length but with the origin located at various positions relative to the fragment ends were used to perform EMSA. A. As a negative control, DNA fragments generated by restriction enzymes as indicated were incubated without Rep to reveal origin DNA do not possess any intrinsic bend. B. The same DNA fragments used in A (B の 1~7 番は A の 1~7 と同じ DNA 断片を使っている) were incubated with Rep (100nM). For each lane the relative mobility μ of DNA was calculated and plotted against the distance from half position of DNA fragments to the *Sac*I site of pKAE2Ori. The bending angle α obtained by using the equation $Cos\alpha/2=\mu M/\mu E$ is indicated. The bending center deduced is boxed. C. A comparison of origin bending by the wild type Rep and its partial fragments. The wild type Rep (WT, 100nM), RepC132 (25nM), RepC94 (100nM) and RepC62 (50nM) were used to perform EMSA.

Fig. 6. Specific single-stranded Ori DNA binding of ColE2 Rep protein and its partial fragments analyzed by EMSA. Various of concentrations of wild type rep and its partial

fragments indicated at the top of each panel were incubated with end labeled single-stranded Ori DNA (Top, nontemplate strand; Bottom, template strand) in the presence of excess M13 single-stranded DNA. The DNA-protein complexes and free DNA fragments were analyzed by electrophoresis on a 4% polyacrylamide gel. The positions of DNA-protein complexes and free DNA are indicated by black triangles. A, the wild type Rep (0, 100, 150nM); B, RepC132 (0, 25, 50, 100nM); C, RepC117 (0, 100, 200nM).

Fig. 7. Model of Ori-DNA melting of ColE2 plasmid.

Step 1. Region I and II of Rep bind to sites I and II of Ori and bend Ori DNA, which induces destabilization of dsDNA of site III of Ori (see the result of DMS footprint of RepC62).

Step 2. The interaction of region III of Rep and site III of Ori stabilizes the conformational change of dsDNA of site III caused by bending at step 1 and induces the primary melting of origin DNA. (RepC117 binds to site III of Ori with a weaker affinity than that of the wild type Rep. It was unable to bind specifically to the ssDNA of Ori. Nevertheless, it can melt the origin DNA with a very low efficiency as compared with that of the wild Rep).

Step 3. Rep binds specifically to the nontemplate strand of Ori to induce full opening of the dsDNA strands of Ori and stabilizes the unwound Ori DNA. (RepC132 binds specifically to the nontemplate strand of Ori and it melts dsDNA of Ori as efficiently as the wild type Rep).

Step 4. The binding of Rep to the nontemplate strand of Ori might also help to position the primase domain probably located in the N terminal region of Rep closely to the template strand to synthesize a primer efficiently, since the specific binding affinity of Rep to the template strand is very faint.

References

1. **Baker, T. A., and S. P. Bell.** 1998. Polymerases and the replisome: machines within machines. Cell **92:**295-305.
2. **Bazaral, M., and D. R. Helinski.** 1968. Circular DNA forms of colicinogenic factors E1, E2 and E3 from Escherichia coli. J Mol Biol **36:**185-94.
3. **Borowiec, J. A., L. Zhang, S. Sasse-Dwight, and J. D. Gralla.** 1987. DNA supercoiling promotes formation of a bent repression loop in lac DNA. J Mol Biol **196:**101-11.
4. **Boulikas, T.** 1996. Common structural features of replication origins in all life forms. J Cell Biochem **60:**297-316.
5. **Bramhill, D., and A. Kornberg.** 1988. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome. Cell **52:**743-55.
6. **Chang, A. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J Bacteriol **134:**1141-56.
7. **del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas.** 1998. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev **62:**434-64.
8. **Dickerson, R. E.** 1999. in Oxford Handbook of Nucleic Acid Structure (Neidle, S., cd) pp. 145-197, Oxford Science, Oxford.
9. **Gille, H., and W. Messer.** 1991. Localized DNA melting and structural pertubations in the origin of replication, oriC, of Escherichia coli in vitro and in vivo. Embo J **10:**1579-84.
10. **Han, M., K. Aoki, M. Yagura, and T. Itoh.** 2007. The ColE2-P9 Rep protein binds to the origin DNA as a monomer. Biochem Biophys Res Commun **353:**306-10.
11. **Han, M., M. Yagura, and T. Itoh.** 2007. Specific interaction between the initiator protein (Rep) and origin of plasmid ColE2-P9. J Bacteriol **189:**1061-71.
12. **Hayatsu, H., and T. Ukita.** 1967. The selective degradation of pyrimidines in nucleic

acids by permanganate oxidation. Biochem Biophys Res Commun **29:**556-61.

13. **Herschman, H. R., and D. R. Helinski.** 1967. Comparative study of the events associated with colicin induction. J Bacteriol **94:**691-9.

14. **Horii, T., and T. Itoh.** 1988. Replication of ColE2 and ColE3 plasmids: the regions sufficient for autonomous replication. Mol Gen Genet **212:**225-31.

15. **Hwang, D. S., and A. Kornberg.** 1992. Opening of the replication origin of Escherichia coli by DnaA protein with protein HU or IHF. J Biol Chem **267:**23083-6.

16. **Ide, H., Y. W. Kow, and S. S. Wallace.** 1985. Thymine glycols and urea residues in M13 DNA constitute replicative blocks in vitro. Nucleic Acids Res **13:**8035-52.

17. **Ilyina, T. V., A. E. Gorbalenya, and E. V. Koonin.** 1992. Organization and evolution of bacterial and bacteriophage primase-helicase systems. J Mol Evol **34:**351-7.

18. **Itoh, T., and T. Horii.** 1989. Replication of ColE2 and ColE3 plasmids: in vitro replication dependent on plasmid-coded proteins. Mol Gen Genet **219:**249-55.

19. **Kawasaki, Y., F. Matsunaga, Y. Kano, T. Yura, and C. Wada.** 1996. The localized melting of mini-F origin by the combined action of the mini-F initiator protein (RepE) and HU and DnaA of Escherichia coli. Mol Gen Genet **253:**42-9.

20. **Kido, M., H. Yasueda, and T. Itoh.** 1991. Identification of a plasmid-coded protein required for initiation of ColE2 DNA replication. Nucleic Acids Res **19:**2875-80.

21. **Kingsbury, D. T., and D. R. Helinski.** 1970. DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem Biophys Res Commun **41:**1538-44.

22. **Konieczny, I., K. S. Doran, D. R. Helinski, and A. Blasina.** 1997. Role of TrfA and DnaA proteins in origin opening during initiation of DNA replication of the broad host range plasmid RK2. J Biol Chem **272:**20173-8.

23. **Kruger, R., I. Konieczny, and M. Filutowicz.** 2001. Monomer/dimer ratios of replication protein modulate the DNA strand-opening in a replication origin. J Mol Biol **306:**945-55.

24. **Kusakabe, T., K. Baradaran, J. Lee, and C. C. Richardson.** 1998. Roles of the helicase and primase domain of the gene 4 protein of bacteriophage T7 in accessing

the primase recognition site. Embo J **17:**1542-52.

25. **Kusakabe, T., A. V. Hine, S. G. Hyberts, and C. C. Richardson.** 1999. The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition. Proc Natl Acad Sci U S A **96:**4295-300.

26. **Kusakabe, T., and C. C. Richardson.** 1997. Gene 4 DNA primase of bacteriophage T7 mediates the annealing and extension of ribo-oligonucleotides at primase recognition sites. J Biol Chem **272:**12446-53.

27. **Kusakabe, T., and C. C. Richardson.** 1996. The role of the zinc motif in sequence recognition by DNA primases. J Biol Chem **271:**19563-70.

28. **Le Chatelier, E., L. Janniere, S. D. Ehrlich, and D. Canceill.** 2001. The RepE initiator is a double-stranded and single-stranded DNA-binding protein that forms an atypical open complex at the onset of replication of plasmid pAMbeta 1 from Gram-positive bacteria. J Biol Chem **276:**10234-46.

29. **Lu, Y. B., H. J. Datta, and D. Bastia.** 1998. Mechanistic studies of initiator-initiator interaction and replication initiation. Embo J **17:**5192-200.

30. **Matsumoto, K., and T. Itoh. unpublished data.**

31. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc Natl Acad Sci U S A **74:**560-4.

32. **Melnikova, A. F., R. Beabealashvilli, and A. D. Mirzabekov.** 1978. A study of unwinding of DNA and shielding of the DNA grooves by RNA polymerase by using methylation with dimethylsulphate. Eur J Biochem **84:**301-9.

33. **Mendelman, L. V., B. B. Beauchamp, and C. C. Richardson.** 1994. Requirement for a zinc motif for template recognition by the bacteriophage T7 primase. Embo J **13:**3909-16.

34. **Messer, W.** 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev **26:**355-74.

35. **Messer, W. a. C. W.** 1996. Initiation of chromosome replication, p. 1579-1601. *In* F.C. Neidhardt, Curtiss, R.,III, J. Ingraham, E.C.C.Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (eds), Escherichia coli and

Salmonella: Cellular and Molecular Biology. ASM Press, Washington D.C.

36. **Park, K., S. Mukhopadhyay, and D. K. Chattoraj.** 1998. Requirements for and regulation of origin opening of plasmid P1. J Biol Chem **273:**24906-11.

37. **Rao, R. N., and S. G. Rogers.** 1979. Plasmid pKC7: a vector containing ten restriction endonuclease sites suitable for cloning DNA segments. Gene **7:**79-82.

38. **Rubin, C. M., and C. W. Schmid.** 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucleic Acids Res **8:**4613-9.

39. **Speck, C., and W. Messer.** 2001. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. Embo J **20:**1469-76.

40. **Tacon, W., and D. Sherratt.** 1976. ColE plasmid replication in DNA polymerase I-deficient strains of Escherichia coli. Mol Gen Genet **147:**331-5.

41. **Takechi, S., and T. Itoh.** 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. Nucleic Acids Res **23:**4196-201.

42. **Takechi, S., H. Matsui, and T. Itoh.** 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. Embo J **14:**5141-7.

43. **Wu, C., R. Roy, and D. T. Simmons.** 2001. Role of single-stranded DNA binding activity of T antigen in simian virus 40 DNA replication. J Virol **75:**2839-47.

44. **Yagura, M., S. Y. Nishio, H. Kurozumi, C. F. Wang, and T. Itoh.** 2006. Anatomy of the replication origin of plasmid ColE2-P9. J Bacteriol **188:**999-1010.

45. **Yasueda, H., T. Horii, and T. Itoh.** 1989. Structural and functional organization of ColE2 and ColE3 replicons. Mol Gen Genet **215:**209-16.

46. **Yasueda, H., S. Takechi, T. Sugiyama, and T. Itoh.** 1994. Control of ColE2 plasmid replication: negative regulation of the expression of the plasmid-specified initiator protein, Rep, at a posttranscriptional step. Mol Gen Genet **244:**41-8.

47. **Yoshimura, S. H., R. L. Ohniwa, M. H. Sato, F. Matsunaga, G. Kobayashi, H. Uga, C. Wada, and K. Takeyasu.** 2000. DNA phase transition promoted by replication initiator. Biochemistry **39:**9139-45.

48. **Zannis-Hadjopoulos, M., and G. B. Price.** 1999. Eukaryotic DNA replication. J Cell Biochem **Suppl 32-33:**1-14.

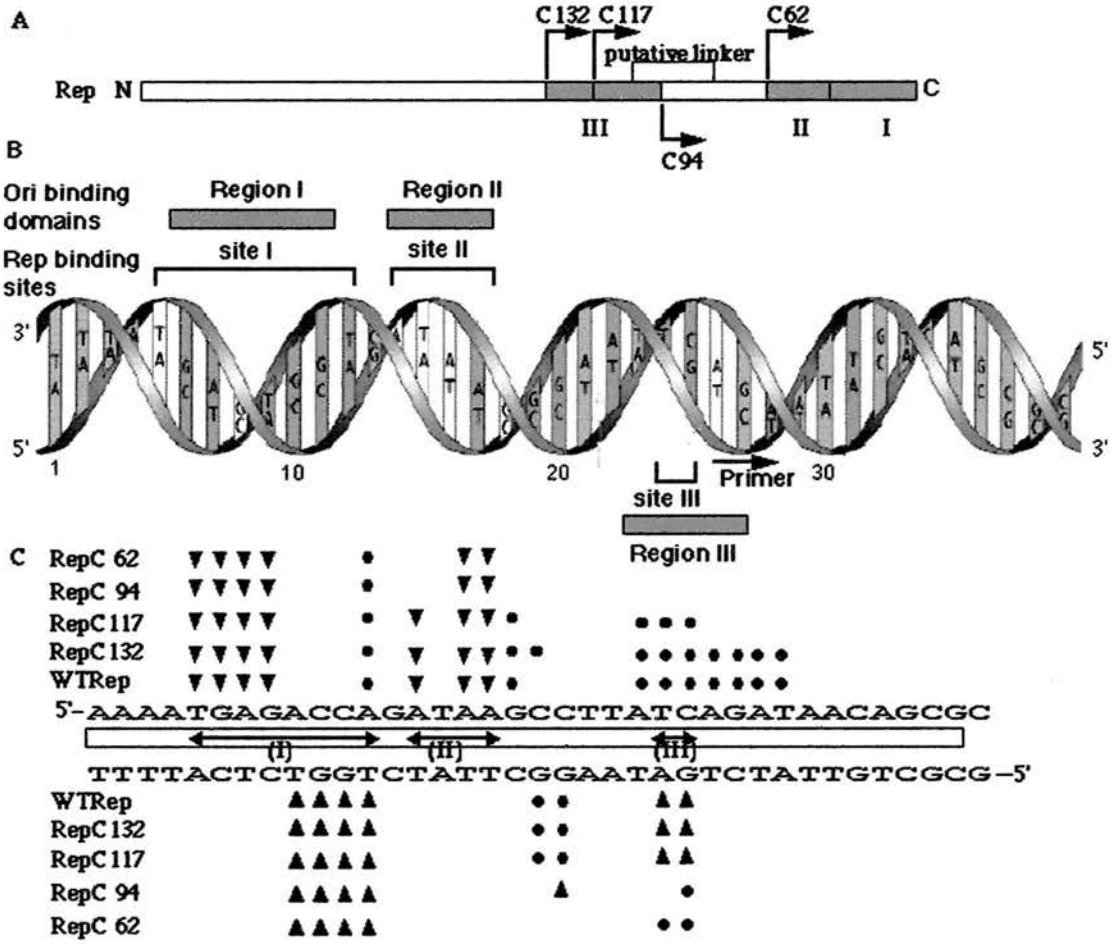

A
C132 C117
C62
putative linker
Rep N
C
III
II
I
C94
B
Ori binding domains
Region I
Region II
Rep binding sites
site I
site II
3'
5'
5'
3'
1
10
20
30
Primer
site III
Region III
C
RepC 62
RepC 94
RepC 117
RepC 132
WTRep
5'-AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGC
(i)
(ii)
(iii)
TTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCG-5'
WTRep
RepC 132
RepC 117
RepC 94
RepC 62

Fig. 1

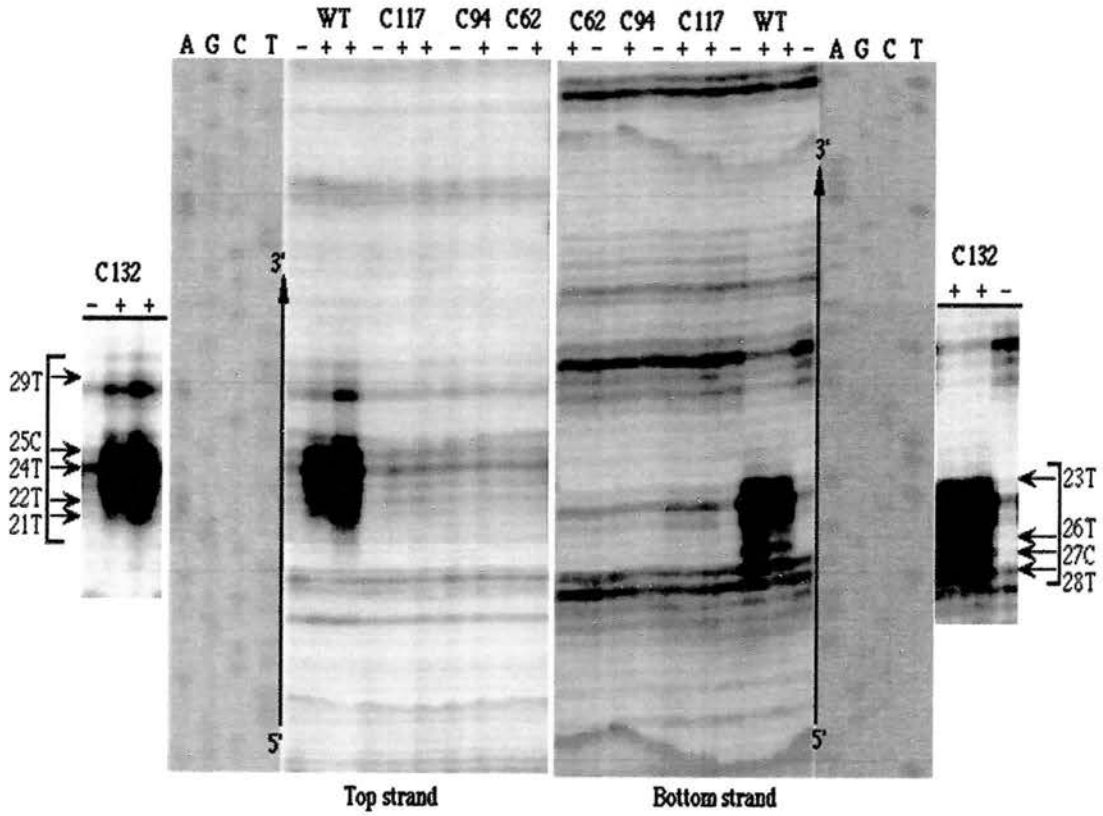

WT C117 C94 C62 C62 C94 C117 WT
A G C T - + + - + + - + - + + - + + - + + - A G C T
C132
- + +
29T
25C
24T
22T
21T
3'
5'
3'
5'
C132
+ + -
23T
26T
27C
28T
Top strand
Bottom strand

Fig. 2

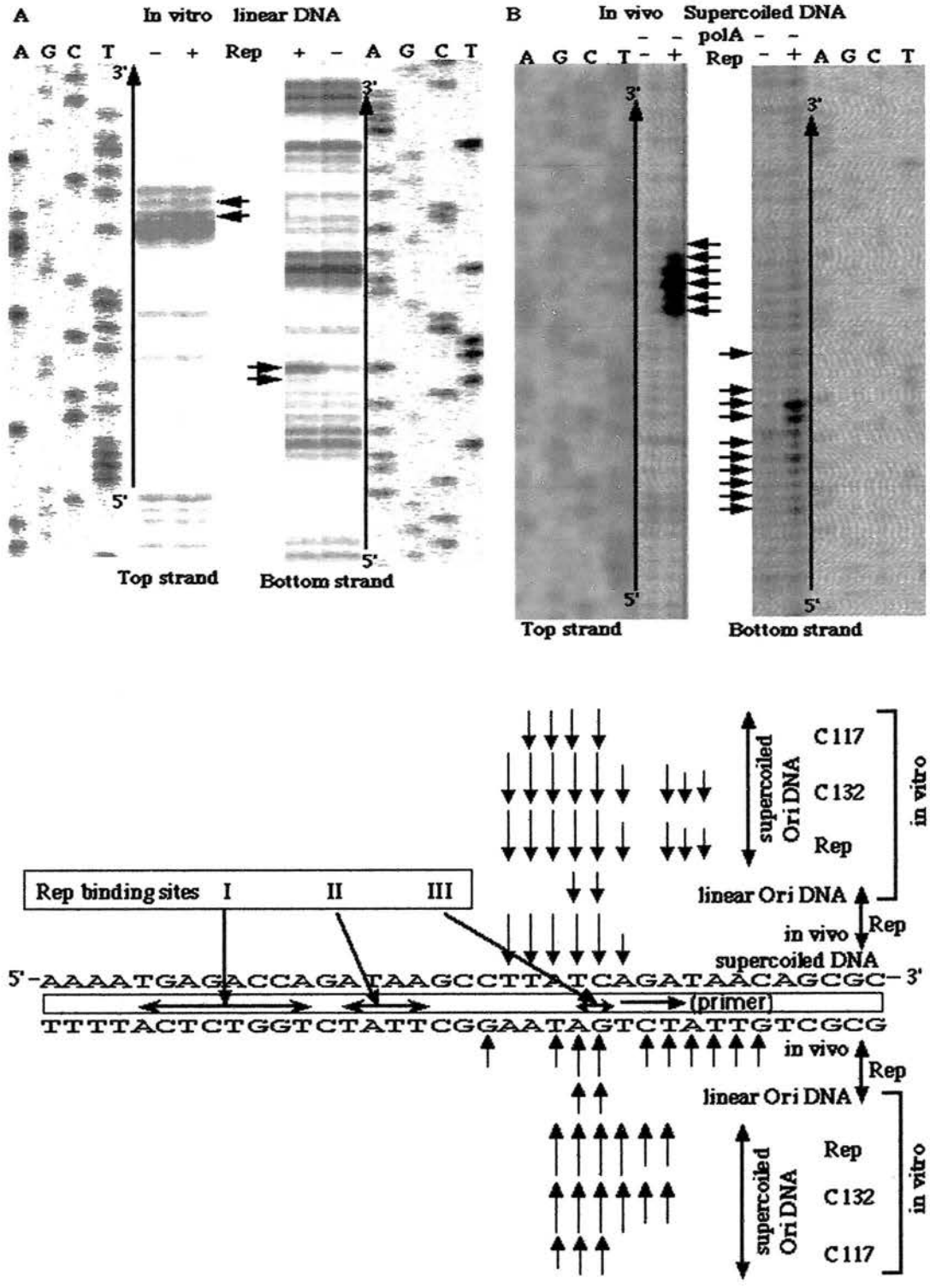

A
In vitro linear DNA
A G C T – + Rep + – A G C T
3'
5'
Top strand
Bottom strand
B
In vivo Supercoiled DNA
polA – –
A G C T – + Rep – + A G C T
Top strand
Bottom strand
supercoiled Ori DNA
C117
C132
Rep
in vitro
linear Ori DNA
in vivo Rep
supercoiled DNA
Rep binding sites I II III
5'-AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGC-3'
(primer)
TTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCG
in vivo Rep
linear Ori DNA
Rep
C132
C117
supercoiled Ori DNA
in vitro

Fig. 3

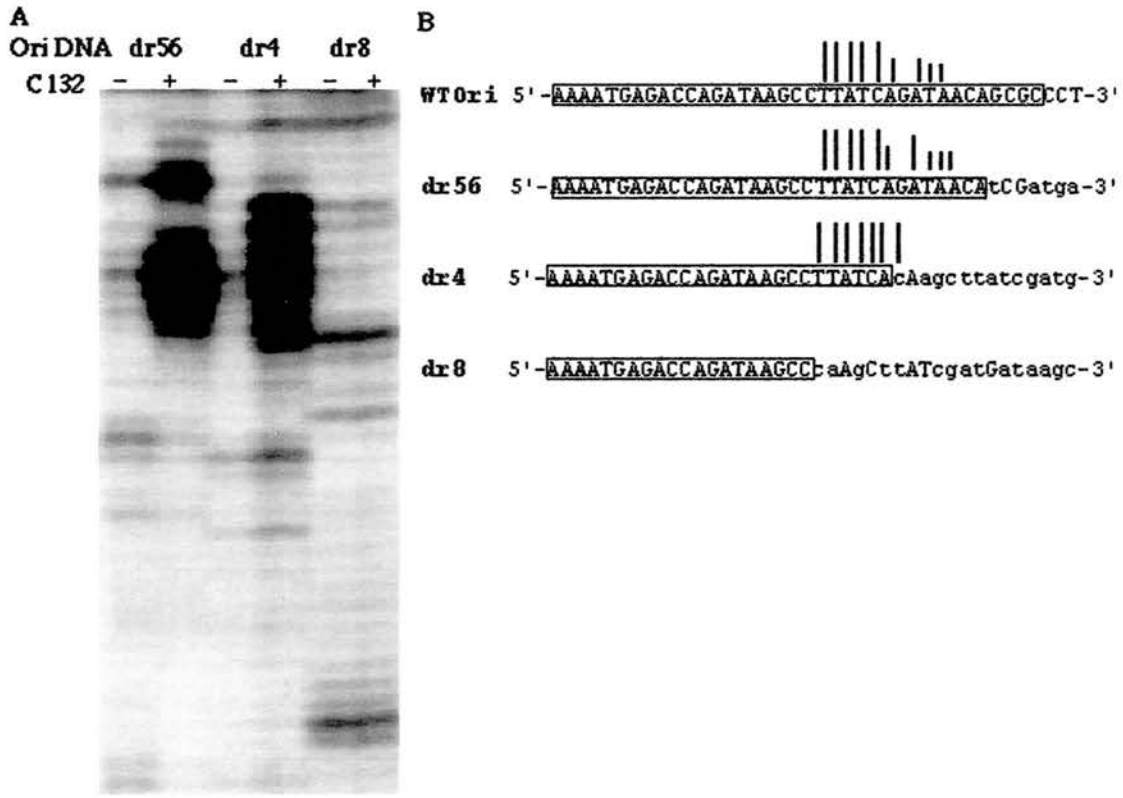

A
Ori DNA dr56 dr4 dr8
C132 - + - + - +
B
WTOri 5'-AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGCCCT-3'
dr56 5'-AAAATGAGACCAGATAAGCCTTATCAGATAACAtCGatga-3'
dr4 5'-AAAATGAGACCAGATAAGCCTTATCAcAagcttatcgatg-3'
dr8 5'-AAAATGAGACCAGATAAGCCtaAgCttATcgatGataagc-3'

Fig. 4

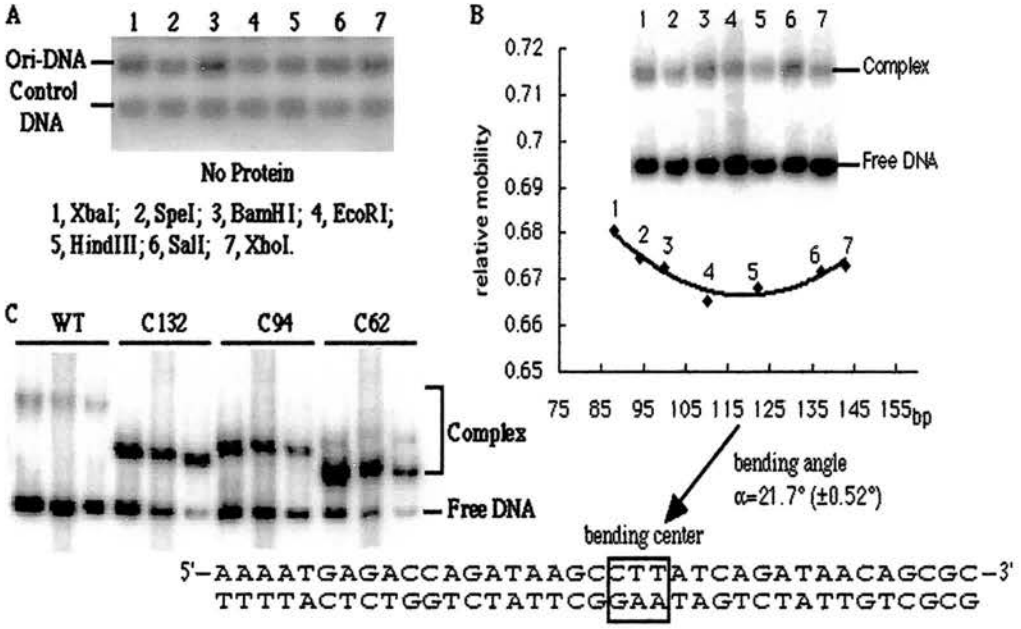

A
1 2 3 4 5 6 7
Ori-DNA
Control DNA
No Protein
1, XbaI; 2, SpeI; 3, BamHI; 4, EcoRI;
5, HindIII; 6, SalI; 7, XhoI.
B
1 2 3 4 5 6 7
Complex
Free DNA
relative mobility
0.72
0.71
0.7
0.69
0.68
0.67
0.66
0.65
75 85 95 105 115 125 135 145 155bp
bending angle
α=21.7° (±0.52°)
bending center
C
WT
C132
C94
C62
Complex
Free DNA
5'-AAAATGAGACCAGATAAGCCTTATCAGATAACAGCGC-3'
TTTTACTCTGGTCTATTCGGAATAGTCTATTGTCGCG

Fig. 5

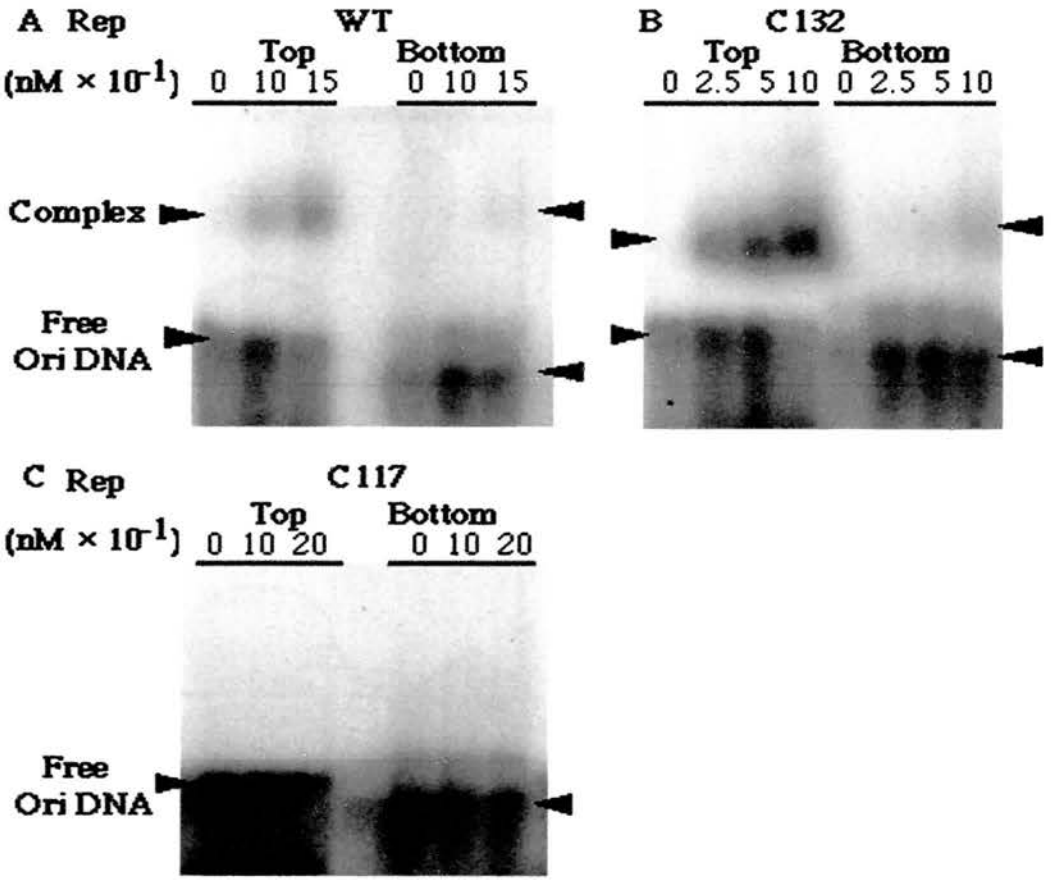

A Rep
WT
Top
Bottom
(nM × 10⁻¹) 0 10 15 0 10 15
Complex
Free
Ori DNA
B
C132
Top
Bottom
0 2.5 5 10 0 2.5 5 10
C Rep
C117
Top
Bottom
(nM × 10⁻¹) 0 10 20 0 10 20
Free
Ori DNA

Fig. 6

Model. 1

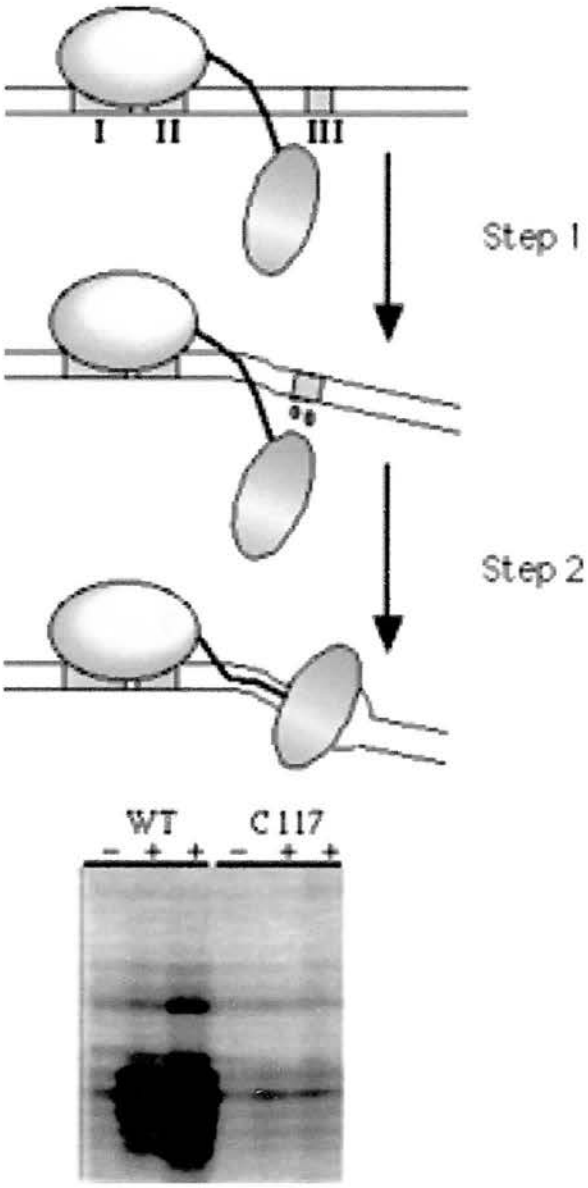


Fig.7

Model. 2

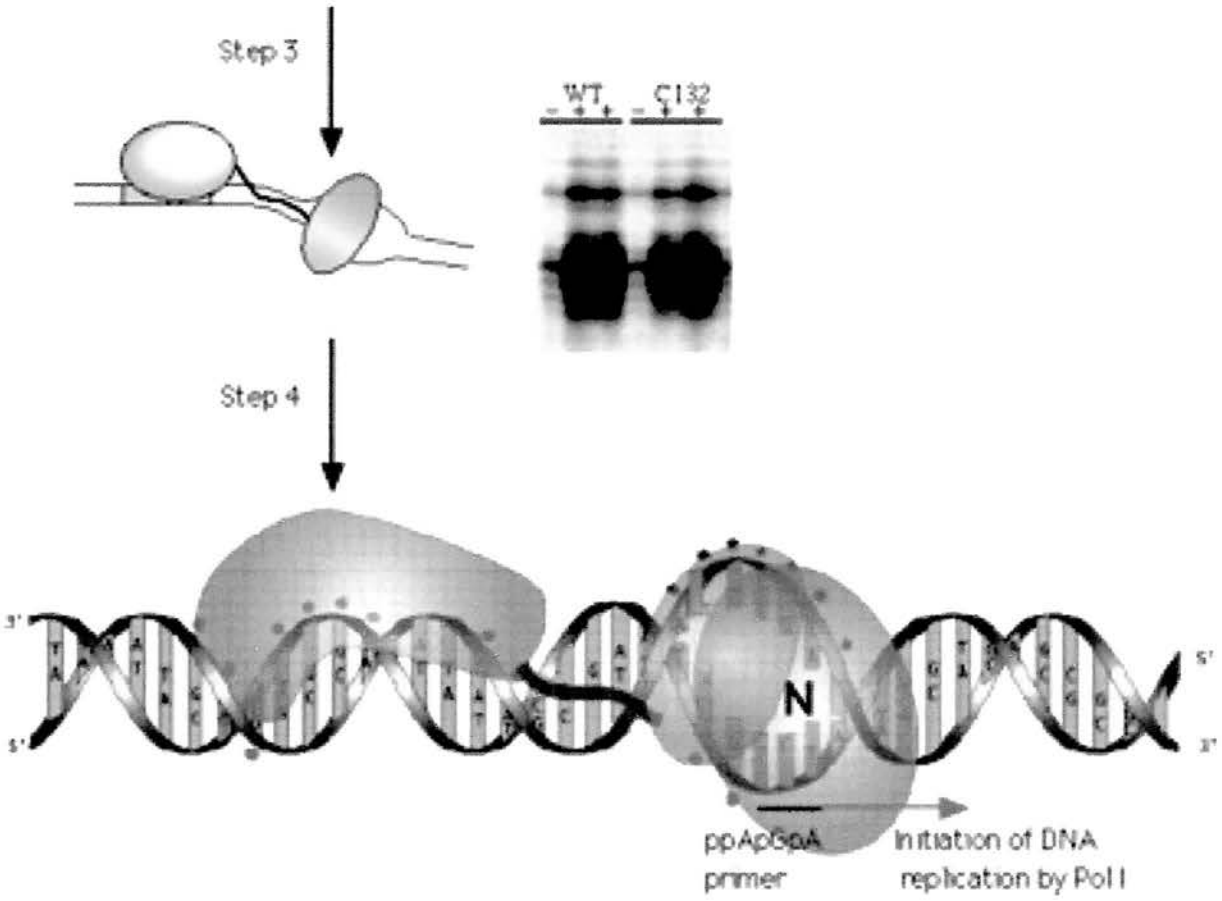

Step 3
WT
C132
Step 4
N
ppApGpA
primer
Initiation of DNA
replication by Pol I

Fig. 7

VI Acknowledgments

I am grateful to all the members of our research group for their discussions. I also thank M. Yagura and K. Aoki for providing materials for this study and S. Nishio for the figure of DNA double helix structure. I would also like to thank M. Yagura and S. Nishio for helpful advice on this study. I wish to express my gratitude especially to Prof. T. Itoh for his ardent guidance. Profs. M.Inuzuka and T. Ichino and Assistant Professors S. Asami, H. Kubo and K. Takata are thanked for their helpful comments on my academic dissertation.

I thank G. Minote and Dr. N. Hayashida in the Research Center for Human and Environmental Sciences of Shinshu University for radioisotope experiments and for providing us an opportunity to use Shimadzu DNA sequencer.

VII. Publications

1. HAN, Man, YAGURA, Masaru and ITOH, Tateo (Dept. Biol.Sci., Fac. Sci., Shinshu Univ.) Analysis of the functional domains of the ColE2 Rep protein. Genes Genet. Syst. (2000) **75** 1C-2.

2. HAN, Man, YAGURA, Masaru, AOKI, kazuteru and ITOH, Tateo (Dept. Biol.Sci., Fac. Sci., Shinshu Univ.) Molecular mechanism of early steps of ColE2 plasmid replication initiation. Genes Genet. Syst. (2004) **79** 1A-03.

3. M. Yagura, S. Nishio, M. Han, T. Nagase, K. Aoki, Department of Biological Science, Shinshu University, 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan. Mechanism of initiation of plasmid ColE2 DNA replication. *Abstracts/plasmid 53 (2005) 23-96.*

4.M. Han, M. Yagura, and T. Itoh. Specific Interaction between the Initiator Protein (Rep) andOrigin of Plasmid ColE2-P9. J. Bacteriol. Feb. (2007). **189:**1061-1071.

5.M. Han, K. Aoki, M. Yagura, and T. Itoh. The ColE2-P9 Rep protein binds to the origin DNA as a monomer. Biochem. Biophys. Res. Commun. 353 (2007) 306-310.

6.M. Han, K. Aoki, M. Yagura, and T. Itoh. Mechanism of origin melting of ColE2-P9 plasmid. In preparition.