

# RepA and DnaA proteins are required for initiation of R1 plasmid replication *in vitro* and interact with the *oriR* sequence

(DNA replication origin/initiation protein/*Escherichia coli* replication proteins/DNase I-protection analysis)

HISAO MASAI AND KEN-ICHI ARAI

Department of Molecular Biology, DNAX Research Institute of Molecular and Cellular Biology, 901 California Avenue, Palo Alto, CA 94304

Communicated by Masayasu Nomura, March 11, 1987

**ABSTRACT** RepA, an initiation protein of R1 plasmid replication, was purified from an *Escherichia coli* strain overproducing the protein. The purified RepA protein specifically initiated replication *in vitro* of plasmid DNA bearing the replication origin of R1 plasmid (*oriR*). The replication, strictly dependent on added RepA protein, was independent of host RNA polymerase but required other host replication functions (DnaB and DnaC proteins, the single-stranded-DNA-binding protein SSB, and DNA gyrase). The replication was also completely dependent on the host DnaA function. In filter binding assays in high salt (0.5 M KCl) conditions, RepA specifically binds to both supercoiled and linear plasmid DNA containing the *oriR* sequence, whereas it binds to nonspecific DNA in low salt. DNase I-protection studies on a linearized DNA fragment revealed that DnaA protein specifically binds to a 9-base-pair DnaA-recognition sequence ("DnaA box") within *oriR* only when RepA is bound to the sequence immediately downstream of the DnaA box. These results indicate that initiation of R1 plasmid replication is triggered by interaction of RepA and DnaA proteins with the *oriR* sequence.

Initiation of R1 plasmid replication requires the plasmid-encoded protein RepA (1, 2). R1 plasmid is replicated *in vitro* when supercoiled template DNA is added to a high-speed supernatant (fraction I) of gently lysed, plasmid-free *Escherichia coli* cells (3). The replication is sensitive to both chloramphenicol and rifampicin, indicating that ongoing protein synthesis is required (3). Subsequently, it was shown that R1 plasmid is replicated in the absence of protein synthesis if the RepA protein is presynthesized (4). RepA, identified as a 33-kDa protein *in vitro*, is cis-acting, being preferentially utilized by the replication origin on the same DNA molecule (1, 5). However, it activates replication origins in trans when the cis origin is deleted (1). The *oriR* sequence, defined *in vitro* by a trans-complementation assay, is sufficient for initiation of R1 plasmid replication *in vitro*. R1 plasmid replication was previously thought to be DnaA-independent because temperature-sensitive *dnaA* mutations can be integratively suppressed by R1 plasmid (6) or by the closely related plasmid R100 (7). The significance of a 9-base-pair (bp) DnaA-recognition sequence ("DnaA box") within *oriR* (8) has not been clear.

Here we report the purification of RepA protein from an overproducing strain. We found that the purified RepA protein specifically bound to the *oriR* sequence and promoted efficient replication of R1 plasmid *in vitro*. Replication in this system, specific for the template bearing the *oriR* sequence, was dependent on the host DnaA function; DnaA protein specifically binds to the DnaA box in *oriR* in the presence of RepA protein bound to the *oriR* sequence.

## MATERIALS AND METHODS

***E. coli* Strains and Plasmids.** The strains and sources were C600, W3110, and HB101 from laboratory stocks; WM433 (*dnaA204*) and WM434 (*dnaA205*) from W. Messer (Max Planck Institute) (9); PC2 (*dnaC2*) from J. A. Wechsler (Columbia University) (10); FA22 (*dnaB*) from I. Herskowitz (University of California, San Francisco) (11); JC206 (*ssb*) from the *E. coli* genetic stock center (Yale University) (12); and  $\lambda$  lysogens carrying the temperature-sensitive *ci* repressor gene *cI857*, MZ-1 (D. Court, unpublished), from D. Bramhill (Stanford University). The overproducing vector pNT26CII was generously provided by K. Shigesada (Kyoto University) (unpublished). pMOB45 (1, 13), a 10.5-kbp mini-derivative of the runaway replication mutant of R1 plasmid, was primarily used as a template for *in vitro* reactions. Other plasmids are described in the text and the legend to Table 1.

**Reagents and Buffers.** Creatine phosphate, creatine kinase, and polyethylene glycol 8000 were from Sigma; [*methyl*-<sup>3</sup>H]dTTP (40–50 Ci/mmol; 1 Ci = 37 GBq) and [*alpha*-<sup>32</sup>P]dTTP (600 Ci/mmol) were from Amersham. Buffer A was 25 mM Hepes-KOH, pH 7.6/40 mM KCl/10 mM MgCl<sub>2</sub>/2 mM dithiothreitol/10% (wt/vol) sucrose, and buffer B was 25 mM Hepes-KOH, pH 6.8/10 mM MgCl<sub>2</sub>/2 mM dithiothreitol/10 mM 2-mercaptoethanol/6 M urea.

***E. coli* Replication Proteins.** Purified DnaA protein and DnaC protein were provided, respectively, by K. Sekimizu and J. Kobori (Stanford University). DnaB protein and single-stranded-DNA-binding protein (SSB) were purified as described (14, 15).

**Assay for DNA Replication.** The standard reaction mixture (25  $\mu$ l) contained the following components: Hepes-KOH (pH 7.6), 40 mM; KCl, 50 mM; MgCl<sub>2</sub>, 9 mM; dithiothreitol, 4 mM; creatine phosphate, 40 mM; creatine kinase, 110  $\mu$ g/ml; ATP, 2 mM; GTP, CTP, and UTP, 500  $\mu$ M each; dATP, dGTP, dCTP, and dTTP, 100  $\mu$ M each, with [*methyl*-<sup>3</sup>H]dTTP at 35 cpm per pmol of total deoxynucleotide; sucrose, 4%; polyethylene glycol 8000, 1% (wt/vol); DNA template, 3600 pmol (deoxynucleotide equivalent); *E. coli* enzyme fraction from strain C600 or W3110 (fraction II) (16), 80–100  $\mu$ g; and RepA protein, 2.5  $\mu$ g. The reaction mixtures were assembled at 0°C and incubated for 60 min at 30°C. Total nucleotide incorporation was measured by determining radioactivity after trichloroacetic acid precipitation. The products of *in vitro* reactions were analyzed as described (1).

**DNase I Protection Analysis ("Footprinting").** Plasmid pHM1289, constructed by inserting a fragment containing *oriR* [positions 1256–1623 in the coordinates of Ryder *et al.* (2)] at the *Bam*HI site of pUC8, was linearized by cleavage with *Eco*RI restriction endonuclease, and the 3' ends were filled in with [*alpha*-<sup>32</sup>P]dATP and the Klenow fragment of DNA polymerase. The product was digested by *Hind*III, and the end-labeled insert fragment was recovered from an agarose gel after electrophoresis. RepA and/or DnaA protein of the

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

Abbreviation: SSB, single-stranded-DNA-binding protein.

amount indicated (in Fig. 5) was bound to the end-labeled fragment in a reaction mixture (50  $\mu$ l) containing 40 mM Hepes-KOH (pH 7.6), 10 mM Mg(OAc)<sub>2</sub>, 40 mM KCl, 4 mM dithiothreitol, 2 mM ATP, 4% sucrose, 0.5 mM CaCl<sub>2</sub>, and 2.5  $\mu$ g of bovine serum albumin. After incubation at 30°C for 15 min, the reaction mixture was incubated with 5 ng of DNase I (Boehringer Mannheim) at 30°C for 1 min. The digested fragment was precipitated with ethanol, resuspended in 3  $\mu$ l of formamide dye (17), and loaded on a sequencing gel for electrophoresis. Maxam-Gilbert (17) sequencing reactions were performed on the same labeled fragment to generate size markers.

## RESULTS

**Purification of the RepA Protein from the RepA-Overproducing Strain.** The *Sau3AI* fragment encompassing the entire *repA* coding region but lacking codons for the two N-terminal amino acids (nucleotides 400–1256) was subcloned into an overproducing vector, pNT26CII, that carries a heat-inducible *P<sub>L</sub>* promoter of bacteriophage  $\lambda$  and the efficient ribosome binding site of  $\lambda$  *cII* (K. Shigesada, unpublished data). The resulting plasmid, pHM5449, overproduced RepA protein in *E. coli* strain MZ-1 after temperature shift to 42°C (Fig. 1A). However, the overproduced RepA protein was not solubilized by treatment of the cells with lysozyme and Triton X-100 (Fig. 1B, fraction I). When the detergent-lysed pellet was resuspended and sonicated, more than 95% of the RepA protein remained in the pellet (Fig. 1B, pellet). Guanidine hydrochloride (3 M) solubilized the RepA protein, leaving contaminating proteins largely insoluble (Fig. 1C, lane 1). The RepA protein was selectively precipitated in 6 M urea at low salt concentration, yielding the active RepA protein >95% pure (Fig. 1C, lane 2).

**Assay for Activity to Promote the Replication of pMOB45 DNA.** An enzyme fraction (fraction II) from *E. coli*, prepared by the method of Fuller *et al.* (16), was supplemented with each protein fraction prepared as above. Consistent with the gel electrophoresis profile of the RepA protein, vigorous activity was detected in the guanidine hydrochloride extract of the pellet. Replication of pMOB45 DNA in fraction II was absolutely dependent on added RepA fraction. The response to increasing amounts of purified RepA was sigmoidal, indicative of a cooperative action of the RepA protein (Fig. 2). Addition of KCl (to 40 mM) stimulated the replication 2- to 3-fold; polyethylene glycol 8000 stimulated 1.5- to 2-fold at 1–2% but was severely inhibitory above 4%. The Mg<sup>2+</sup> concentration optimum was rather sharp around 9 mM. Omission of ATP together with the ATP-regenerating system (creatine phosphate and creatine kinase) completely abolished the reaction, indicating the requirement for ATP. Sucrose was generally stimulatory (by 20–30%) over a concentration range of 1–8%, whereas glycerol at 0.4% inhibited the reaction by more than 70% (data not shown). DNA synthesis was linear for 60 min after a lag of 5–10 min, reaching a level of 60% of the input DNA in some experiments (data not shown).

**Replication Is Specific for the *oriR* Sequence.** The specificity of the replication for the *oriR* sequence was demonstrated by the failure of those plasmids that lack the entire *oriR* sequence (pRP825) or carry a deletion within *oriR* (pHM648) to serve as templates (Table 1). These plasmids were inert as templates in a fraction I system and were not maintained as independent replicons *in vivo* (unpublished data). Plasmid DNA carrying only the *oriR* sequence (pRP814-8) served as a template, demonstrating that *oriR* is the minimal sequence requirement for the replication in this fraction II system.

**Characterization of RepA-Dependent Replication *In Vitro*: Requirement for Host DnaA Function.** Replication in the fraction II system was resistant to both rifampicin and chloramphenicol, indicating that neither the function of host

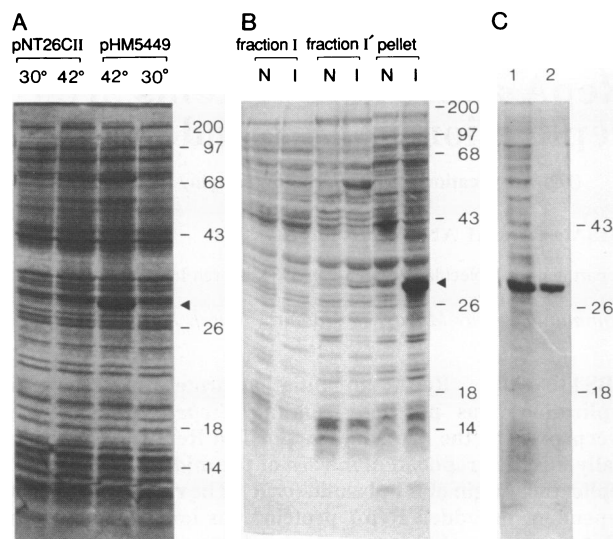


FIG. 1. Overproduction and purification of RepA protein. Proteins were analyzed by NaDodSO<sub>4</sub>/polyacrylamide gel electrophoresis in 15.0% (A) or 12.5% (B and C) gels and detected by staining with Coomassie brilliant blue R-250. Positions and sizes (kDa) of standard proteins run in the same gel are given to the right of each panel. In A and B, arrowheads mark the positions of RepA protein. (A) pHM5449 and the vector pNT26CII were introduced into *E. coli* strain MZ-1. The strains were grown in L broth at 30°C to OD<sub>600</sub> 0.5. Then a portion of each culture was transferred to 42°C to activate the  $\lambda$  *P<sub>L</sub>* promoter and the remainder was kept growing at 30°C. After 1 hr, the cells were harvested and the whole-cell extract was loaded on a gel. (B) The cells of 1-liter cultures of MZ-1(pHM5449), either induced at 42°C or not, were resuspended in 50 mM Tris-HCl, pH 7.5/10% sucrose and frozen in liquid nitrogen. After the cells were thawed at 10°C, lysozyme (0.2 mg/ml), MgCl<sub>2</sub> (10 mM), KCl (40 mM), and dithiothreitol (2 mM) were added and the mixture was incubated at 0°C for 20 min. The cells were lysed by 0.1% Triton X-100 and the supernatant from centrifugation at 150,000  $\times$  *g* (fraction I) was obtained. The pellet was resuspended in 10 ml of buffer A with sonication and ultracentrifuged at 150,000  $\times$  *g* for 15 min. The supernatant (fraction I') was removed, and the pellet was resuspended in 5 ml of buffer A (pellet). Approximately 100  $\mu$ g of protein was loaded on each lane. N, noninduced; I, induced. (C) The pellet suspension prepared from induced MZ-1(pHM5449) as described in B was mixed with an equal volume of 6 M guanidine hydrochloride and ultracentrifuged at 150,000  $\times$  *g* for 10 min. The supernatant (lane 1), containing the RepA protein, was dialyzed against buffer B. The RepA protein was specifically precipitated during the dialysis; the precipitate, collected by centrifugation, was resuspended in 1/10th the volume of buffer A containing 3 M guanidine hydrochloride and dialyzed against buffer B containing 0.5 M ammonium sulfate. The final preparation (lane 2) contains RepA protein that is more than 95% pure.

RNA polymerase nor concomitant protein synthesis is required, whereas inhibition by nalidixic acid and novobiocin implicated DNA gyrase (data not shown). Analysis of replication intermediates indicated that replication initiates at or near *oriR* and proceeds unidirectionally (data not shown). Host replication proteins, including DnaB, DnaC, and SSB, are required for the replication, as indicated by the inactivity of extracts prepared from the mutant strains and the recovery of activity by addition of the corresponding purified protein. On the other hand, protein *i*, the product of *dnaT* (18), is not required (Table 2). Unexpectedly, the fraction II from *dnaA* mutants *dnaA204* and *dnaA205* was inert for R1 plasmid replication. Replication was restored when purified DnaA protein was added (Fig. 3). Only two or three DnaA molecules per template are required to achieve maximal R1 plasmid replication; in contrast, maximal replication of *oriC* requires 20–40 DnaA molecules per *oriC* template (8).

**Interaction of RepA and DnaA with the *oriR* Sequence.** Filter

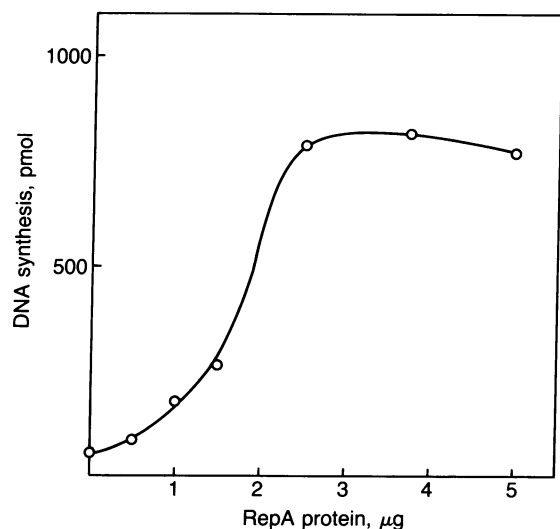


FIG. 2. Requirement for RepA protein in the replication of pMOB45 DNA *in vitro*. Fraction II from *E. coli* strain C600 was supplemented with various amounts of purified RepA protein in the standard reaction mixture with supercoiled pMOB45 DNA as template. DNA synthesis is expressed as pmol of dNMP incorporated.

binding assays using  $^3\text{H}$ -labeled supercoiled or linear plasmid DNA indicated that RepA protein binds efficiently to DNA containing the *oriR* sequence. At high salt (0.5 M KCl), nonspecific DNA binding was completely abolished and only plasmid containing *oriR* was bound to RepA protein (data not shown).

DNase I footprinting analyses of a linear DNA fragment indicated the extensive interaction of RepA protein with the *oriR* sequence. The strongest protection was detected on the 21-bp sequence (positions 1436–1456) immediately downstream of the DnaA box, with enhanced cleavage at the deoxyadenosine residue at position 1435 (Fig. 4A). Protection was detected over nearly 100 bp of sequence downstream of the DnaA box, although less protection was observed over the sequence distal to the box. No protection was detected upstream as far as position 1434. Use of more DNase I in the cleavage reaction revealed the presence of hypersensitive sites within the protected region (Fig. 4B). Enhanced cleavage was detected between the three blocks of the protected

Table 1. Template specificity of replication *in vitro*

Template	DNA synthesis, pmol
pMOB45	630
pKA1	620
pREP803	384
pRP845	380
pHM648 ( <i>oriR</i> $\Delta$ )	45
pRP825 ( <i>oriR</i> $^-$ )	32
pRP814-8 ( <i>oriR</i> )	230
pUC13	30

Standard reaction mixtures containing 3 nmol (as nucleotide) of template were incubated at 30°C for 60 min. Rifampicin (100  $\mu\text{g}/\text{ml}$ ) was also included to suppress the replication from the pUC8 origin. Values are averages of duplicate experiments. pKA1 (14) is a pMOB45 derivative carrying the *E. coli* gene *dnaB*. pREP803 and pRP845 (1) are hybrid plasmids between pUC8 and R1; pREP803 carries  $\approx 2$  kb downstream of *oriR*, whereas pRP845 carries only 12 bp downstream of *oriR*. pHM648 is identical to pRP845 except that it has a 23-bp deletion within *oriR* (positions 1548–1570); due to the deletion, pHM648 lost replication activity *in vivo*. pRP825 lacks the entire *oriR* sequence (deletion of the region downstream of position 1434). pRP814-8 carries *oriR* (positions 1406–1611) on pUC8 vector.

Table 2. Replication of R1 plasmid in extracts from *dna* mutants

Source of fraction II	Protein added	DNA synthesis, pmol
FA22 ( <i>dnaB</i> )	None	60
	DnaB (50 ng)	590
PC2 ( <i>dnaC2</i> )	None	19
	DnaC (120 ng)	657
PC7 ( <i>dnaC7</i> )	None	17
	DnaC (60 ng)	471
JC206 ( <i>ssb</i> )	None	47
	SSB (150 ng)	175
UT205 ( <i>dnaT1</i> )	None	639
	Protein i (50 ng)	555

Fraction II was prepared from the indicated *dna* mutant strains. Cultures were grown at 30°C. Standard reaction mixtures containing 3600 pmol (as nucleotide) of pMOB45 were incubated with or without the addition of purified replication proteins. Values with added replication proteins are those at the maximal stimulation.

regions (positions 1436–1456, 1459–1470, and 1482–1496). Alternate enhancement and protection appeared at 8- to 12-bp intervals between positions 1496 and 1522 (Fig. 4B). The cleavage pattern in the region downstream of 1523 was not influenced by RepA protein.

Incubation of the fragment with a  $>100$ -fold molar excess of DnaA protein did not influence the cleavage pattern except for the appearance of enhanced cleavage at positions 1423 and 1424 (Fig. 4C). However, with increased amounts of RepA protein in the presence of a constant amount of DnaA protein, protection of the 9-bp sequence TTATCCACA (positions 1427–1435) and enhanced cleavage at positions 1423, 1424, and 1435 were detected (Fig. 4C). These results indicate that binding of RepA protein to *oriR* is required for efficient interaction of DnaA protein with the DnaA box.

## DISCUSSION

Replication of R1 plasmid is strictly dependent on the plasmid-encoded protein RepA and the replication origin (*oriR*) both *in vivo* and *in vitro*. In this paper, we report that RepA, purified from an overproducing strain, initiates replication *in vitro* of R1 plasmid and its derivatives carrying the

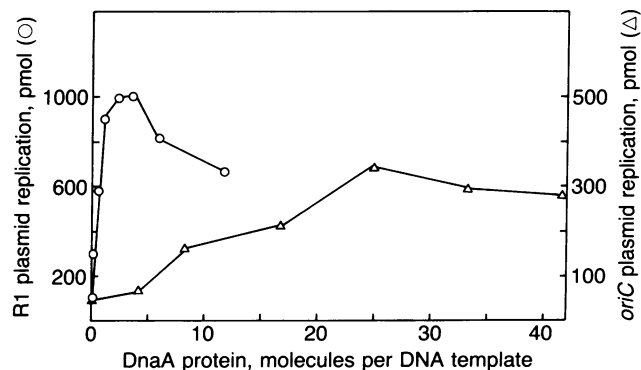


FIG. 3. DnaA function is required for R1 and *oriC* plasmid replications *in vitro*. Fraction II from *E. coli* strain WM433 (*dnaA204*) was incubated under the standard conditions with various amounts of purified DnaA protein. pMOB45 DNA (3600 pmol of nucleotides; 10.5 kbp, 0.17 pmol of circular plasmid molecules) or M13 *oriC26* DNA (600 pmol of nucleotides; 12.3 kbp, 0.024 pmol of circular molecules) was used as template. For *oriC* replication, 7.0% polyvinyl alcohol was used instead of 1.0% polyethylene glycol (16). The amounts of DnaA protein required for maximal replication of R1 and *oriC* plasmids were 0.4 pmol and 0.6 pmol, respectively. Similar results were obtained with extract from WM434 (*dnaA205*).

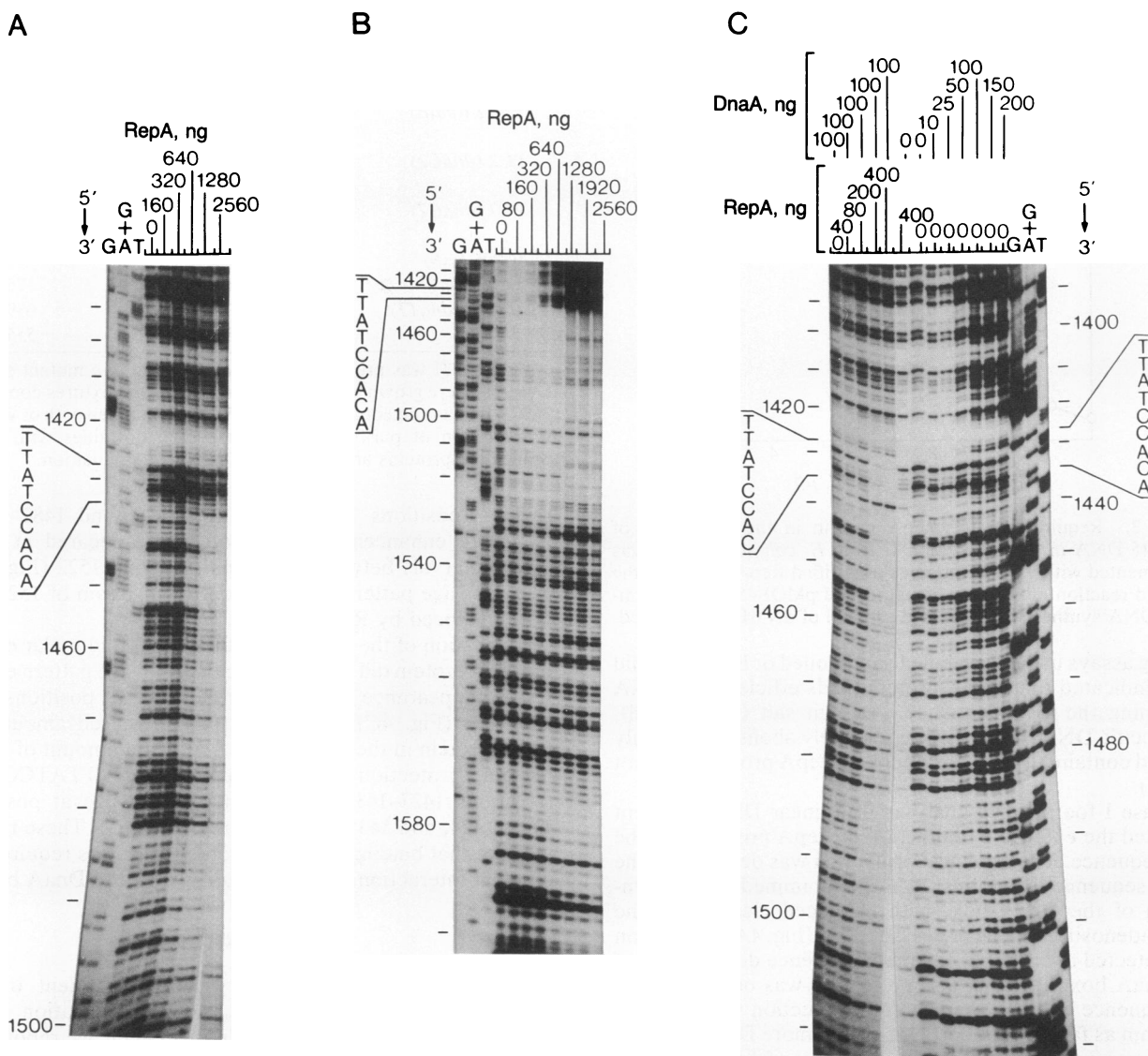


FIG. 4. Footprint of RepA and/or DnaA bound to *oriR*. Ten nanograms (A and B) or 6 ng (C) of the *EcoRI*–*HindIII* fragment of pHM1289 (Materials and Methods), labeled at the 3' end of the *EcoRI* site (position 1623), was incubated with various amounts of RepA and/or DnaA proteins (shown above individual lanes) and subjected to partial Dnaase I digestion. Samples were analyzed by electrophoresis in 6% (A and C) and 8% (B) sequencing gels. In B, digestions were with 20 ng of DNase I to emphasize hypersensitive sites. At this level of DNase I, the average size of digested fragments is too small to analyze the protection over a range longer than 150 bp. The sequence numbering is based on the Maxam–Gilbert (17) sequencing ladders (G, G+A, and T).

*oriR* sequence and that RepA binds specifically to *oriR*. The replication, requiring host replication functions such as DnaB, DnaC, SSB, and DNA gyrase, was also completely dependent on DnaA function. The requirement for DnaA protein was unexpected, since integrative suppression of *dnaA* mutations has been reported for R1 plasmid (6) and the closely related IncFII plasmid R100 (7). This may be explained if R1 plasmid is a composite replicon consisting of a DnaA- and RepA-dependent replicon and a second, *dnaA*-independent replicon. An alternative explanation is that the residual DnaA activity of *dnaA* mutant strains permits replication of R1 plasmid but not that of the host chromosome; this is supported by the fact that R1 plasmid replication *in vitro* requires very small amounts of DnaA protein compared to *oriC* plasmid. pSC101 (19), P1, and F (20, 21) plasmids are known to require DnaA function for their replication *in vivo*. All of these plasmids have one or two DnaA boxes near the origin (2, 8, 22–24), which may be significant in the initiation of replication.

Specific binding of the purified RepA protein to both

supercoiled and linear DNA containing the *oriR* sequence was detected in filter binding assays in the presence of high salt (0.5 M KCl). Footprinting analyses revealed that the RepA protein specifically binds to the sequence immediately downstream of the DnaA box within *oriR* on a linear DNA fragment (Fig. 5). Protection covers 87 bp downstream of the *dnaA* box at a saturating level of the RepA protein, with the strongest protection observed over the 21-bp sequence at positions 1436–1456. No protection was observed downstream of position 1523 or upstream of the DnaA box. Host DnaA protein, which has been shown to bind specifically to DnaA boxes within *oriC* or at other locations (8), did not bind very efficiently to the DnaA box within *oriR*. However, the DnaA box was protected from DNase I digestion when DnaA protein was bound to a fragment in the presence of RepA protein. This may indicate that the binding of RepA to *oriR* facilitates the recognition of the DnaA box by DnaA protein by inducing a conformational change within the *oriR* DNA sequence. Alternatively, the interaction of DnaA protein with the DnaA box sequence may be facilitated by direct inter-

