J. Biochem. 130, 399–405 (2001)

The Structure and Function of the Replication Initiator Protein (Rep) of pSC101: An Analysis Based on a Novel Positive-Selection System for the Replication-Deficient Mutants

Tsukasa Fueki and Kazuo Yamaguchi[1]

Institute for Gene Research, Kanazawa University, Kanazawa 920-0934

Received May 23, 2001; accepted June 19, 2001

Plasmid pSC101 encodes a 37.5 kDa Rep (RepA) protein, which binds to three 21-base repeats (DR-1, DR-2, and DR-3) in the replication origin region (*ori*) of the plasmid to initiate replication. Rep also binds to two palindromic sequences (IR-1 and IR-2) which overlap the *rep* promoter. The binding of Rep to IR-2 represses the production of Rep itself. It is highly likely that the balance of these functions of Rep plays a major role in controlling the copy number of pSC101. In this study, we developed a positive-selection system for replication-deficient mutants of the initiator protein. This system can be applied to the study of other replication systems by changing *ori* and *rep* of pSC101 to the corresponding genes. Thirty-four replication-deficient (Ini$^-$) mutants were isolated with this system, and analyzed as to the relation between the structure and function of the Rep protein. Seventeen of these 34 Ini$^-$ mutants were found to lack auto-repressor activity as well as initiator activity. DNA sequence analysis showed that one-third (from the C-terminus) of Rep is dispensable for the auto-repressor activity, while the initiator activity seems to require the whole protein.

Key words: DNA replication, initiator protein, plasmid, positive selection of mutants, pSC101.

Plasmid pSC101 encodes a 1.3 kb segment which is essential for its autonomous replication (*1, 2*) (Fig. 1). The segment is roughly divided into two parts, *ori*, the replication origin, and the *rep* gene, which is translated into the 37.5 kDa Rep (RepA) protein. The *ori* region contains two binding boxes for host-encoded proteins, IHF and DnaA, respectively (*3, 4*), separated by an AT-rich region, followed by three 21-base repeats (direct repeats; DR-1, DR-2, and DR-3) (*1, 5–8*). The *ori* region must be in *cis*, whereas the *rep* region can be either *in cis* or *in trans* for the autonomous replication of the plasmid (*1*).

It is suggested that binding of Rep to the DRs in *ori* is a key step in the initiation of plasmid replication (*9, 10*). On the other hand, Rep also binds to two inverted repeats (IR-1 and IR-2), which are located between DR-3 and the *rep* gene, overlapping the *rep*-promoter. The binding of Rep to IR-1 is suggested to enhance the replication of the plasmid (*12*) and the binding to IR-2 is suggested to repress the production of Rep itself, which interferes with the transcription (*13–15*) (Fig. 1). Several mutations on Rep were found to increase the copy number of the plasmid (*8, 16–19*), suggesting a regulatory role for the protein in the control of the copy number of pSC101. Rep was also found to have different DNA-binding properties in monomeric and dimeric forms *in vitro*; the monomer binds to DRs and the dimer to IRs selectively (*11, 20*). It is highly likely that the balance between the monomeric initiator and the dimeric auto-repressor is the first regulatory factor controlling the copy number. In addition, Rep is also a negative control factor, as overproduction of Rep represses the replication of the plasmid *in vivo* (*18, 19, 21*).

Our next concern is the relation between the function and structure of Rep itself. In the present study, we developed a screening system for the positive selection of replication-deficient *rep* mutants (Ini$^-$ mutants). With this system, we isolated 34 Ini$^-$ mutants, of which 17 also lack auto-repressor activity. DNA sequence analysis of these mutants showed that one-third (from the C-terminus) of Rep is dispensable for the auto-repressor activity, while the replication initiation activity seems to require the whole protein.

MATERIALS AND METHODS

Bacterial Strains and Plasmids—Escherichia coli K-12 strains JM109 *recA1 Δ(lac proAB) endA1 gyrA96 thi1 hsdR17 supE44 relA1*/F′ *traD36 proAB lacI^q-Z ΔM15* (*22*), HI1006 F$^-$ *araD139 Δ(ara leu)7697 lacX74 galU galK strA recA1* (*23*), and MA194 *mutT* (*24*) were used. The plasmids used in this study are listed in Table I.

Plasmid Construction—Plasmid pλCI was constructed by cloning the *Pst*I–*Bgl*II fragment of pMY12-6Amp^r containing λ cI857 into the *Pst*I–*Xba*I site of pUC19 (the *Bgl*II and *Xba*I ends were blunt-ended by end-filling reaction with Klenow-fragment and then ligated). Plasmid pλ101 was constructed by cloning the *Hinc*II–*Nsp*(7524)I fragment of pKMY213dr10P containing pSC101 *ori* but lacking *rep* into the *Sma*I site of pλCI [the *Nsp*(7524)I end was blunt-ended

[1] To whom correspondence should be addressed. Tel: +81-76-265-2771, Fax: +81-76-234-4535, E-mail: yamaguti@kenroku.kanazawa-u.ac.jp

Abbreviations: DR, direct repeats; IR, inverted; Amp, ampicillin; Km, kanamycin; Cm, chloramphenicol.

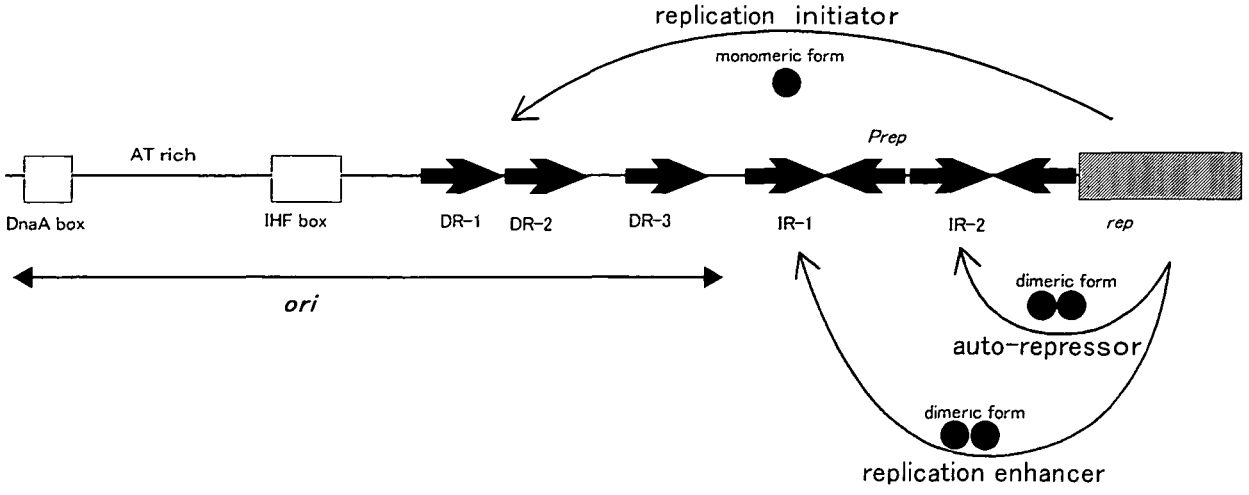

Fig. 1. **The structure of the segment essential for the autonomous replication of pS101 and the functions of the Rep protein.** Boxes represent the DnaA protein–binding sequence (DnaA box), the integration host factor protein–binding sequence (IHF box), and the structure gene (*rep*) of Rep. The arrows represent three direct repeats (DR-1, DR-2, and DR-3) and two inverted repeats (IR-1 and IR-2). P*rep*, the *rep* promoter; AT rich, AT-rich stretches.

TABLE I. **Plasmids used.**

Plasmid	Properties	Ref./source
pMY12-6Amp^r	pBR322*ori*, P_L, P_R, λ cI857	*13*
pUC19	pUC*ori*, Amp^r	*25*
pKMY213dr10P	pBR322*ori*, pSC101*ori*	*12*
pUC4K	pUC*ori*, Km^r	Pharmacia
pFF5	pBR322*ori*, F*ori*(f5), *lac*, Amp^r	Imai
pUC-431/585	pUC*ori*, P*rep*	*26*
pKMY292dl29	pBR322*ori*, *rep*	*13*
pCM4	pBR322*ori*, CAT	Pharmacia
pBR322	pBR322*ori*, Tet^r, Amp^r	*27*
pλCI	pUC*ori*, λ cI857	This work
pλ101	pUC*ori*, λ cI857, pSC101 *ori*	This work
pλ101KM^r	pUC*ori*, λ cI857, pSC101 *ori*, Km^r	This work
pF	pUC*ori*, F*ori*	This work
pTF001	pSC101 *ori*, F*ori*, λ cI857, Km^r	This work
pP_L	pUC*ori*, Amp^r, λ P_L	This work
pP_LCM^r	pUC*ori*, Amp^r, λ P_L-CAT	This work
pP_LCM^r-Finc	pUC*ori*, Amp^r, λ P_L-CAT, F_{incB}	This work
pTF002	pBR322*ori*, Amp^r, λ P_L-CAT, F_{incB}	This work
pTF002-P_R	pBR322*ori*, Amp^r, λ P_RP_L-CAT, F_{incB}	This work
pTF003	pBR322*ori*, pUC*ori*, Amp^r, λ P_RP_L-CAT, F_{incB}	This work
pUC-rep	pUC*ori*, Amp^r, *rep*	This work
pTF005	pBR322*ori*, Amp^r, λ P_RP_L-CAT, F_{incB}, *rep*	This work

by mung-bean nuclease digestion]. Plasmid pλ 101Km^r was constructed by cloning the *Pst*I fragment of pUC4K containing the kanamycine resistant marker into the *Pst*I site of pλ101. Plasmid pF was constructed by cloning the *Sal*I–*Xho*I fragment of miniF(f5) containing the miniF basic replicon (F*ori*) into the *Sal*I site of pUC19.

Plasmid pTF001 was constructed by combining two *Sac*I–*Sph*I fragments containing Km^r, λcI857, and pSC101 *ori* from pλ 101Km^r, and F*ori* from pF, respectively. Plasmid pP_L was constructed by cloning the *Bgl*II–*Bam*HI fragment of pMY12-6Amp^r containing the P_L promoter into the *Bam*HI site of pUC19. Plasmid pP_LCm^r was constructed by cloning the *Bam*HI fragment of pCM4 containing the *cat* gene into the *Bam*HI site of pP_L, located downstream of the P_L promoter. Plasmid pP_LCm^r-Finc was constructed by cloning the *Sma*I fragment of pF containing *incB* in F*ori* into the *Sma*I site of pP_LCm^r.

Plasmid pTF002 was constructed by combining the *Hin*dIII–*Kpn*I fragment of pP_LCm^r-Finc, and the *Hin*dIII–*Ava*I fragment of pBR322 containing the pBR322 *ori* region and the *bla* gene (the *Kpn*I and *Ava*I ends were blunt-ended by mung-bean nuclease digestion and ligated). Plasmid pTF-002-P_R was constructed by insertion of the synthetic λ P_R promoter fragment into the *Sal*I–*Sph*I site of pTF002, located upstream of the P_L promoter.

Plasmid pTF003 was constructed by combining pUC-431/585 containing the *rep*-promoter and pTF002-P_R at the *Hin*dIII and *Sph*I sites of each plasmid. Plasmid pUC-rep was constructed by cloning the *Bam*HI–*Hin*dIII fragment of pKMY292dl29, containing *rep* without its promoter, into the corresponding site of pUC19.

Plasmid pTF005 was constructed by combining pTF003 and pUC-rep after removing their pUC*ori* and the Amp^r marker by *Hin*dIII and *Kpn*I digestion.

Hydroxylamine Mutagenesis—pTF005 was digested with *Hin*dIII and *Kpn*I. The smaller fragment (*rep*) was incubated in 0.7 M NH_2OH·HCl at 65°C for 180 or 240 min. After dialysis against 10 mM Tris, 20 mM NaCl, and 1 mM EDTA (pH 8.0), DNA was retrieved by ethanol precipitation. Then the mutated fragments were re-ligated to the larger fragment.

mutT Mutagenesis—Plasmid pTF005 was introduced into *E. coli* MA194 *mutT*. Soon after the transformation, the cells were divided into 20 portions and each portion was incubated in 8 ml of LB-Amp at 37°C for 17 h (till the stationary phase) independently (I). Each culture (5 μl) was transferred to 8 ml of fresh LB-Amp and then cultured at 37°C for another 12 h (II). Plasmid DNA was extracted from each culture at (I) and (II).

Auto-Repressor Assay—HI1006 carrying pZZ1 (R6K*ori*, Km^r, P*rep-lacZ*) was transformed with Ini$^-$ mutants. The transformants were streaked on MacConkey plates as positive and negative controls. The β-galactosidase activities were determined as described previously (*28*).

DNA Sequence Analysis—DNA sequence analysis was performed with a auto-sequencer (A.L.F. DNA sequencerII: Amersham Pharmacia Biotech) using standard procedures.

Immunodetection of the Rep Protein—The immunodetection of Rep was performed as described previously (*19*). The protein was immunostained with biotinylated goat anti-rabbit IgG antibodies (Peroxidase Vectastain ABC; Vector Labs) and Immunostaining HRP-1000 (Konika).

Antibiotics—The concentrations of antibiotics used were as follows: ampicillin 75 mg/liter, chloramphenicol 50 mg/liter, and kanamycine 15 mg/liter.

RESULTS

In order to specify the functional domain related to the initiator activity, it is necessary to obtain as many Ini⁻ *rep* mutants as possible. The problem is how to select the mutants that do not replicate. A negative selection system would be easier to construct. However, the chance of such mutations in *rep* may be so small that the mutants would be overlooked. Therefore we constructed a system for positive selection of Ini⁻ mutants on plates. Our strategy is as follows. The system consists of two plasmids: pTF001 and pTF005 (Fig. 2). Plasmid pTF001 was constructed to have the λ cI repressor gene, Km^r marker, and two replicational origins, of which the basic replicon of miniF (*Fori*) is fully functional and pSC101*ori* lacks the *rep* gene. pTF005 was constructed to have pBR322 *ori*, the *cat* gene under the control of λ P_L and P_R promoters, the *incB* gene of miniF, the Amp^r marker, and the *rep* gene of pSC101.

pTF001 alone in a cell replicates stably from F*ori*. The phenotype of the cell is therefore Km^r. Then, once pTF005 is introduced with pTF001, the replication from F*ori* of pTF001 is prevented by F*incB* of pTF005, while in cells provided with Rep *in trans* by pTF005, the replication starts from pSC101*ori* of pTF001. As a result, these two plasmids can only co-exist in a cell when Rep from pTF005 is active in initiating replication. When these two plasmids co-exist in a cell, λ repressors from pTF001 block the expression of *cat* of pTF005, leading to the phenotype Cm^s, Amp^r, Km^r. Then, if mutated, Rep of pTF005 loses its ability to initiate replication, pTF001 is flipped out from the cell, and the phenotype becomes Cm^r, Amp^r, Km^s. Therefore, in this system, the Ini⁻ mutants are positively selected on LB-Cm plates.

First, the proposed system was checked with pTF005 and pTF005-Δrep as negative and positive controls, respectively. Either pTF005 or pTF005-Δrep was introduced with pTF001 and then the cells were spread on LB-Amp plates.

TABLE II. **Properties of Ini⁻ mutants.**

MS	MT	Base change	Rpr	Mutagen
3	F	GAA (Glu) → *AA	–	HA
8	F	AAA (Lys) → *AA	–	HA
35	M	AAC (Asn) → CAC (His)	++	*mutT*
46	F	CGG (Arg) → C*G	–	HA
49	F	TCG (Ser) → TC*	–	HA
52	M	TAT (Tyr) → GAT (Asp)	++	*mutT*
59	F	ATG (Met) → AT*	–	HA
68	F	GGT (Gly) → *GT	–	HA
70	M	TTA (Leu) → TTC (Phe)	++	*mutT*
86	F	AAT (Asn) → *AT	–	HA
97	N	TGG (Trp) → TAG (STOP)	–	HA
104	F	TCA (Ser) → T*A	–	HA
111	M	GTT (Val) → GGT (Gly)	++	*mutT*
116	F	ATA (Ile) → AT*	–	HA
117	F	TTG (Leu) → T**	–	HA
119	M	TAT (Tyr) → GAT (Asp)	++	*mutT*
121	M	TTC (Phe) → TGC (Cys)	++	*mutT*
134	M	GTT (Val) → ATT (Ile)	+	HA
191	F	TTG (Leu) → T*G	–	HA
208	F	GTT (Val) → G*T	–	HA
211	N	CGA (Arg) → TGA (STOP)	–	*mutT*
221	M	CAA (Gln) → CAC (His)	++	*mutT*
224	F	CTA (Leu) → *TA	–	HA
227	N	CAA (Gln) → TAA (STOP)	++	HA
233	F	GAA (Glu) → *AA	++	HA
236	F	AAC (Asn) → AA*	++	HA
248	F	ACA (Thr) → A*A	++	HA
259	F	GGA (Gly) → G*A	++	HA
264	F	CTA (Leu) → *TA	++	HA
287	F	AGT (Ser) → A*T	+	HA
291	M	CTC (Leu) → TTC (Phe)	++	HA
294	F	TCG (Ser) → TC*	+	HA
295	M	TTC (Phe) → TCC (Ser)	++	*mutT*
306	M	CTA (Leu) → CGA (Arg)	–	*mutT*

MS, mutation site (the codon numbered from the 5′ end of *rep*). MT, mutation type: F, frame shift; M, missense; N, nonsense. Base change, *deleted base. Rpr, repressor activity: ++, 55–120 Miller units of β-galactosidase; +, 350–500 Miller units; –, 1,500–2,000 Miller units. Mutagen: HA, hydroxylamine.

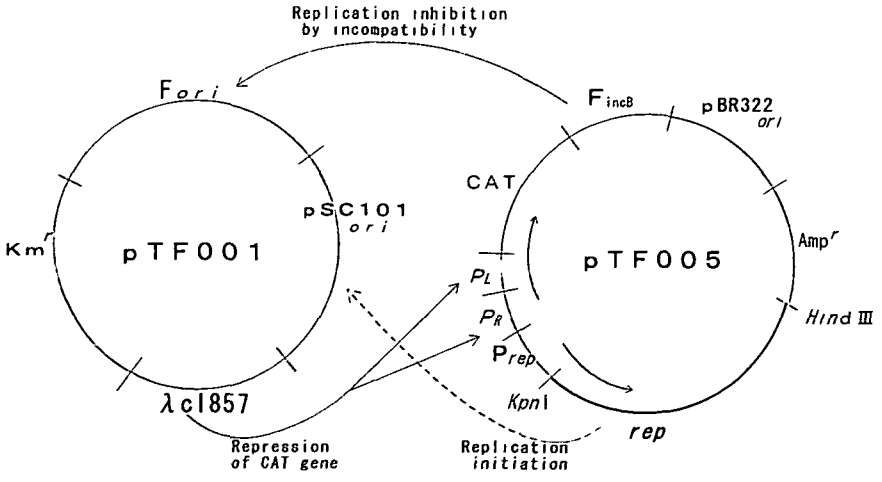

Fig. 2. **A scheme of the positive selection system designed to isolate Ini⁻ mutants.** F*ori*, the basic replicon of miniF; F*incB*, the *incB* region of miniF.

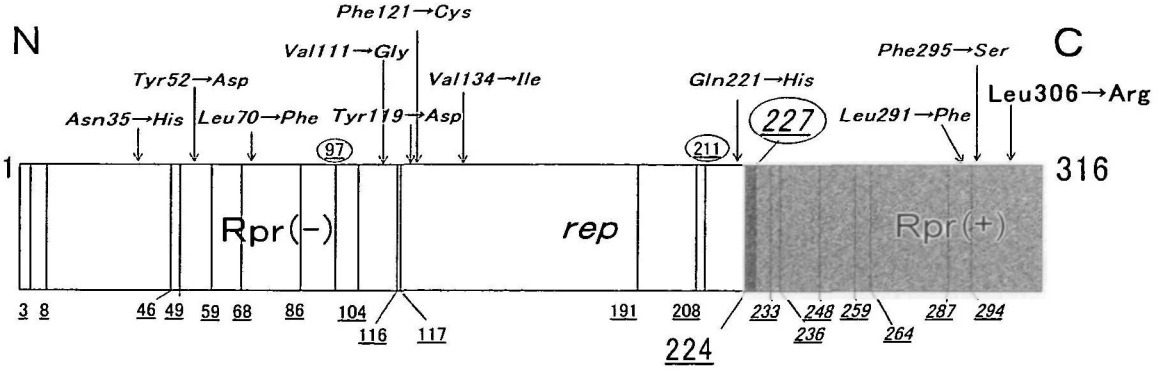

Fig. 3. **Properties of Ini⁻ mutations in *rep*.** The numbers, 1 to 316, show the codons numbered from the 5′-terminus of the *rep* gene presented as a box, although the first Met is removed in the purified Rep (*10*). Circled numbers above the box and the underlined numbers below the box denote the codons where the nonsense and frame-shift mutations occurred, respectively. Missense mutations are presented above the box. The mutations in italic letters maintain the auto-repressor activity (Ini⁻ Rpr⁺). □, The region indispensable for the auto-repressor activity; ▨, The region dispensable for the auto-repressor activity; ■, the switch for the auto-repressor activity.

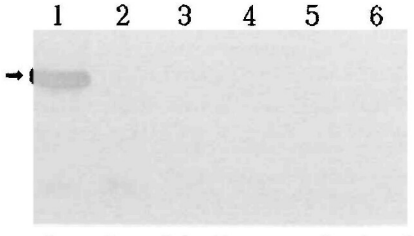

Fig. 4. **Immunodetection of the Rep proteins.** Total proteins (50 μg) extracted from cells carrying a plasmid with the wild *rep* or a mutated *rep* were electrophoresed and then immunodetected with anti-Rep antibodies. The arrow indicates the Rep protein. Lane 1, wild type *rep* (pTF005); lane 2, *rep227*; lane 3, *rep236*; lane 4, *rep259*; lane 5, *rep294*; lane 6, *rep306*.

Twenty Amp^r colonies were randomly selected from each group, and streaked onto both LB-Cm and LB-Km+Amp plates. All the colonies from pTF001/pTF005 were Cm^s and Km^r, and those from pTF001/pTF005-Δrep were Cm^r and Km^s without exception. Also, when pTF005 was introduced into cells with pTF001, the increase in the amount of pTF001 DNA per cell was confirmed on gel-electrophoresis, showing the change in the replicational origin of pTF001 from *Fori* to pSC101*ori*. When pTF005-Δrep was introduced into cells with pTF001, the disappearance of pTF001 from the cells containing pTF005-Δrep was confirmed. Thus, the positive selection system was proved to function as designed with high reliability.

Then, chemically (hydroxylamine) or biologically (*mutT*) mutated pTF005 was introduced into cells with pTF001, and the cells were spread on LB-Cm plates. Plasmids were extracted from the Cm^r colonies and re-introduced into newly prepared competent cells carrying pTF001 for confirmation of the phenotype. From a chemically mutagenized pTF005 group, 24 Ini⁻ mutants were isolated. Ten mutants were also obtained by *mutT* mutagenesis. These 34 mutants (Table II) were subsequently subjected to repressor-activity assays. Sixteen of them were found to lack auto-repressor activity (Rpr⁻) (Table II and Fig. 3). DNA sequence analysis of the mutants obtained on hydroxylamine mutagenesis showed that almost all the mutations were conversions from GC to AT, resulting in the introduction of a stop codon or deletion of one base pair (a frame shift mutation), the exception (134Val to Ile) being a missense mutation. On the other hand, 9 of the 10 mutations obtained on *mutT* mutagenesis were missense ones, as expected.

The map positions of the nonsense and frame-shift mutations indicated that initiator activity is lost when only 22 amino acids from the C-terminus of Rep are deleted (*rep294*), while repressor activity remains even with the deletion of about 1/3 of the C-terminus. The switch for the auto-repressor activity was found to be located between 224Leu and 227Gln. The point mutations (Ini⁻ Rpr⁺) have a tendency to gather in the first half (between codons 35 and 134) from the N-terminal end, where the mutations of most of the copy number mutants of pSC101 were reported (*16–19*). Only one missense mutant (*rep306*) was found to lack both initiator and repressor activity. Very interestingly, the mutation was located at codon 306 (306Leu), *i.e.* in the region dispensable for the auto-repressor activity. The same mutation (306Leu to Arg) was detected in Ini⁻ Rpr⁻ mutants from independent *mutT* cultures.

It is plausible that the Ini⁻ Rpr⁻ phenotype of mutant Rep proteins is due to rapid degradation of the protein within cells. We performed immunodetection of some mutant Rep proteins with anti-Rep antibodies. However, no Rep306 protein was detected as well as other mutant proteins carrying mutations in the C-terminal region (Fig. 4). Since mutant Rep proteins (Rep227, Rep236, and Rep259) have normal repressor activities, our anti-Rep antibody would only recognize the C-terminal region, regardless of the polyclonal antibody.

DISCUSSION

A novel Selection System for Mutants Deficient in Replication—The positive selection system we designed to isolate Ini⁻ mutants in this study showed very high reliability and worked remarkably well. The estimated frequency of Ini⁻ mutations on *mutT* mutagenesis was below 1/10^6 cells and we would have needed thousands of plates for this study, if we had adopted the negative selection system, with a high possibility of overlooking the targets. The system can easily be applied to the study of similar plasmids

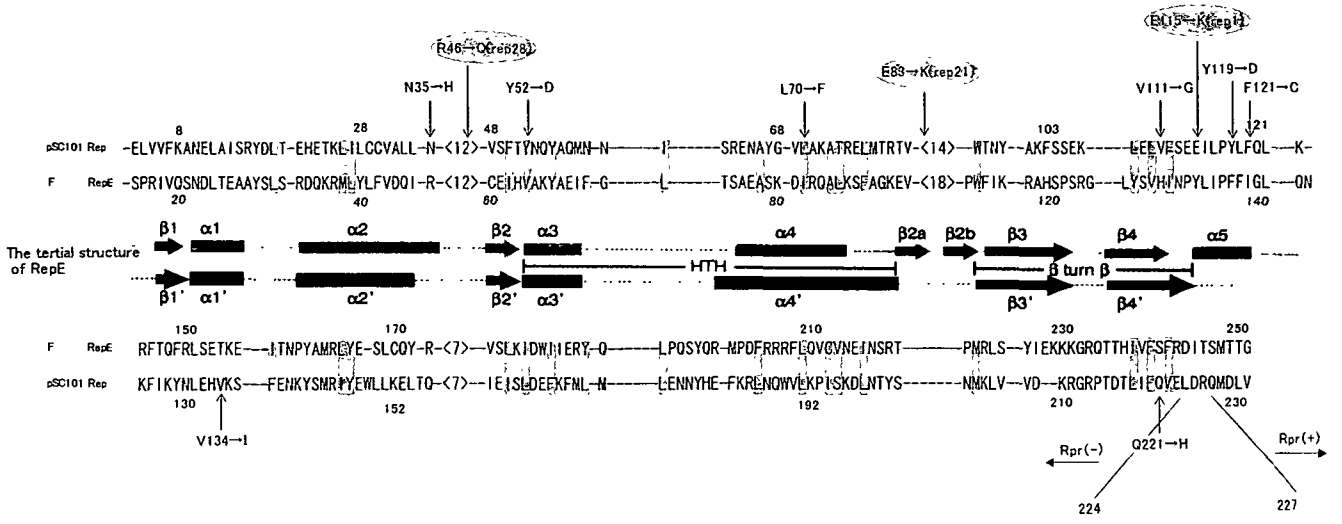

Fig. 5. **The presumed three-dimensional locations of Ini⁻ missense mutations, based on comparison with the tertiary structure of RepE.** The tertiary structure of RepE is schematically indicated in the centre, coordinated with its amino acid sequences. Note the N-terminal (residues 15–144: above) and C-terminal (residues 145–246: below) are related by 2-fold symmetry. The α-helical segments are shown as cylinders and the β-strands as arrows. The conserved locations of the hydrophobic amino acid residues between Rep (outside) and RepE (inside) are shown by shaded boxes. The numbers in < > indicate the numbers of residues that are not displayed. Hyphens indicate gaps. The N-terminal sequences of both proteins and the C-terminal sequence of Rep are not shown. The locations of Ini⁻ mutations relative to the tertiary structure of RepE are based on the locations of the conserved hydrophobic residues (shaded boxes) determined by twofold symmetry (*31*). Ini⁻ missense mutations of Rep are shown above or below the sequence as well as the high copy number mutations (*rep28*, *rep21*, and *rep1*; shaded circles). The switch for the auto-repressor activity is located between residues 224 and 227 of Rep.

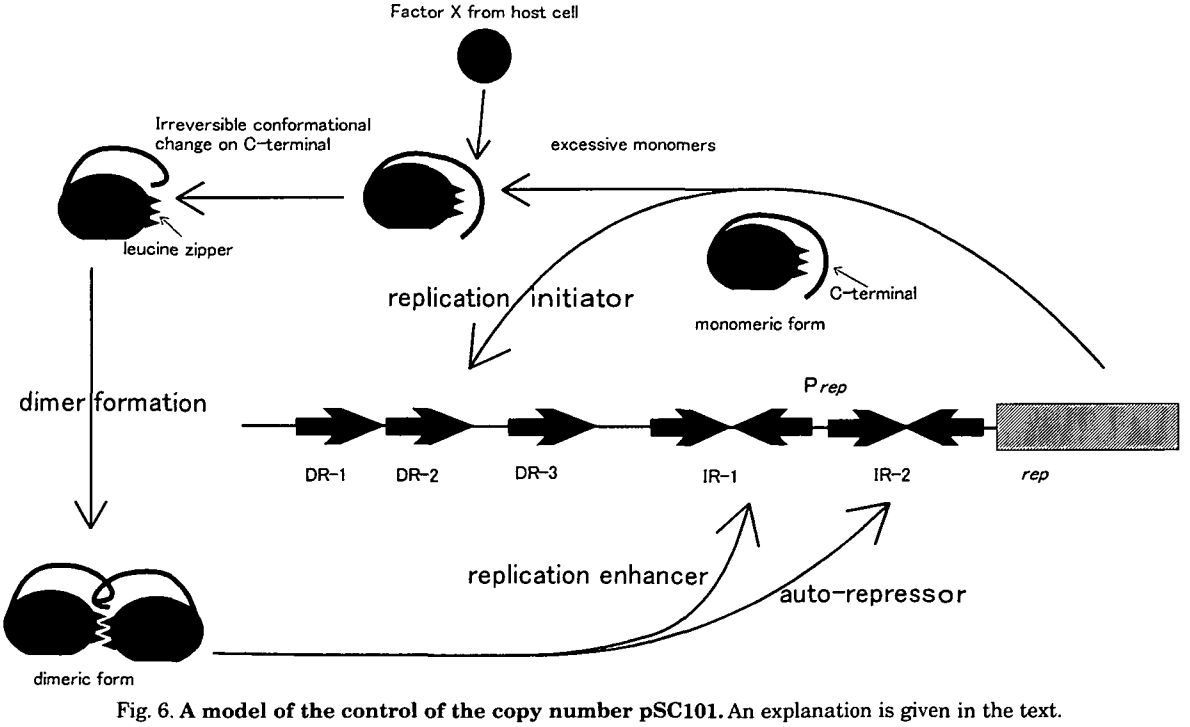

Fig. 6. **A model of the control of the copy number pSC101.** An explanation is given in the text.

by changing pSC101*ori* of pTF001 and pSC101*rep* of pTF-005 to the corresponding genes.

Repressor Domain of Rep—For the initiator activity, almost the whole protein seems to be required since the loss of 16 amino acids from the C terminal end (Fig. 3), and deletion of more than 12 amino acids from the N terminal end prevented replication (*11*). While about 1/3 of the sequence starting from the C-terminus of Rep (up to codon 227) was dispensable for the auto-repressor activity. A previous study on the binding properties of dimeric Rep for IRs *in vitro* showed that a protein retaining 227 amino acids from the N terminus was functional but one retaining 211

amino acids was not (*11*). The present study more strictly narrowed this boundary down to between residues 224 and 227. Near the switch for auto-repressor activity, a consensus amino-acid motif, -KRGRP-, that binds to an AT-rich DNA sequence (*29*) is found (codon 210–214). This region may participate in the binding of Rep to IR-DNA. An alternative explanation of the Ini$^-$ Rpr$^-$ phenotype of mutant Rep proteins is rapid degradation of the protein within cells. In fact, Rep224 retaining 224 amino acids from the N-terminus was not immunodetected (unpublished data). However, this does not necessarily prove the instability of the protein, since other mutant Rep proteins with normal repressor activity are also not recognized by the antibody (Fig. 4). It is most likely that Rep224 is defective in the repressor activity and that this characteristic is not due to rapid degradation, although the possibility could not be completely ruled out.

The mutation sites of most of the Ini$^-$ Rpr$^+$ point mutants isolated in this study were found to be distributed in the first half (between codons 35 and 134 from the N-terminus) of the Rep protein, where the sites of mutation reported for the high copy number mutants are found (*16–19*). This region may have an important function in initiating the replication of the plasmid. As the mutants with point mutations in this region retain the auto-repressor activity, it is unlikely that this region participates in the DNA binding or dimerization of Rep itself. We speculate that the region participates in the interaction with host-encoded replicational proteins. In fact, the region near codon 120 was reported to interact with host-encoded DnaB helicase (*30*).

Recently, crystal structure analysis of the initiator protein (RepE) of F factor was reported (*31*) (Fig. 5) (Accession nos. for MMDB and PDB, 12768 and 1REP, respectively). RepE is about 2/3 the size of Rep, and has dual functions as a replication initiator and an auto-repressor. The conserved locations of the hydrophobic residues related by two-fold symmetry in the two proteins have been pointed out. The residues form hydrophobic cores in the tertiary structure and it was suggested that these proteins could have very similar three-dimensional structures. The switch for the auto-repressor activity on Rep is located between 224Leu and 227Gln, which almost coordinates with the end of β4′ in the proposed three-dimension model of RepE.

Adopting the proposed model of RepE for Rep, most of the point mutations [such as 35Asn to His(α2), 52Tyr to Asp(β2-α3), 111Val to Gly(β4), 119Tyr to Asp(α5), 121Phe to Cys(α5), and 134Val to Ile(α1′)] are found on the opposite side of the DNA binding domains along with the copy number mutations [46Arg to Gln(β2), 83Glu to Lys(β2a), and 115Glu to Lys(α5)] reported previously. This is consistent with our presumption that these mutations affect the interaction between Rep and host-encoded proteins. It is also interesting that the only Ini$^-$Rpr$^+$ mutation in the DNA binding domain (α4) is responsible for conservation of the hydrophobicity (70Leu to Phe), whereas most of the other mutations change the polarity. The reason why with this mutation the auto-repressor activity is conserved, although it is located in the DNA binding domain, may be as follows: The consensus sequence, symbolized as $GGN_2TAGN_{2\text{-}4}AT$, has a 4-base spacer between TAG and AT in DR, whereas IR has 2- and 3-base spacers in its pseudo-symmetric sequences (*10*). The change of 70Leu to Phe does not influence the recognition of IR by the Rep dimer, but the increased mass of the side chain may prevent monomeric Rep from recognizing the DR sequence with the 4-base spacer.

The "Extra" C-Terminal Region of Rep—Thus, we are now beginning to comprehend the relation between the function and structure of Rep down to 2/3 (from the N-terminal end) of the sequence. However, the structure and function of the "extra" 1/3 remains an enigma, although it probably takes part in replication initiation, as the deletion mutants and the three point mutants as to this region failed to allow pSC101 to replicate. As it is dispensable for the auto-repressor activity, it is unlikely that this region participates in the DNA-binding. Also, it has been reported that deletion mutants as to the C-terminal region bind to DR DNA *in vitro* (*11*). From the observation that deletion or frame shift mutants as to the C-terminus fail to interact with polyclonal antibodies of Rep (Fig. 4), it is supposed to be located at the surface of the protein.

The striking observation in this study that a single point mutation at codon 306 (Leu to Arg) prevents Rep from functioning as not only the replication initiator but also the auto-repressor may provide a clue as to the function of the C-terminal region. Although the possibility cannot be excluded at this stage, it is unlikely that the mutation destroyed any of the positive functions as a repressor, since the mutated site is located in the "extra" region. One possibility is that the mutation destroyed the internal switch converting Rep to the repressor mode from the initiator mode. The monomeric form (initiator form) and dimeric form (repressor form) of Rep were once reported to be in equilibrium (*32*). However, DR sequences, the binding sites of the Rep monomer, are responsible for expression of the incompatibility but the IR sequences, the binding sites of the Rep dimer, are not (*7*). Extra copies of IR cancel the autoregulation of Rep synthesis resulting in an increase in the concentration of Rep in the cell, but extra copies of DR do not (unpublished observation). If they are in equilibrium, an explanation for above the findings is required.

On the assumption that the mutation at codon 306 destroyed the switch that converts Rep from a monomer to a dimer, we propose a simple model for control of the copy number of pSC101 (Fig. 6). The increase in cell volume with growth lowers the concentration of dimeric Rep, leading to Rep being produced in the monomeric form *via* autoregulation. The newly produced "extra" region of the C-terminus hinders, for instance, the leucine-zipper region in the N-terminal domain to prevent dimerization, so Rep remains in the monomeric, *i.e.*, initiator, form. Some of the initiators bind to DRs to begin replication to maintain a certain concentration of the plasmid in a cell. The C-terminal domain of the excess initiator is quickly and irreversibly modified, for instance, folded by factor X from the host, which recognizes Rep at the amino-acid sequence near codon 306, to allow Rep to dimerize and make it inactive as to replication. Folding of the C-terminal domain of the dimeric form might drastically differ from that of the monomeric form, since the mutation at codon 306 (Rep306) led to inability in immunodetection (Fig. 4), although again the possibility of rapid degradation of the protein cannot be completely ruled out. This model has yet to be tested experimentally. We are currently performing experiments to examine the validity of this model.

REFERENCES

1. Yamaguchi, K. and Yamaguchi, M. (1984) The replication origin of pSC101: Replication properties of a segment capable of autonomous replication. *J. Gen. Appl. Microbiol.* **30**, 347–358
2. Linder, P., Churchward, G., and Caro, L. (1983) Plasmid pSC101 replication mutants generated by insertion of the transposon Tn1000. *J. Mol. Biol.* **170**, 287–303
3. Fuller, R.S., Funnel, B.E., and Kornberg, A. (1984) The DnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* **38**, 889–900
4. Gamas, P., Burger, A.C., Churchward, G., Caro, L., Galas, D., and Chandler, M. (1986) Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. *Mol. Gen. Genet.* **204**, 85–89
5. Churchward, G., Linder, P., and Caro, L. (1983) The nucleotide sequence of replication and maintenance function encoded by plasmid pSC101. *Nucleic Acids Res.* **11**, 5645–5659
6. Vocke, C. and Bastia, D. (1983) Primary structure of the essential replicon of the plasmid pSC101. *Proc. Natl. Acad. Sci. USA* **80**, 6557–6561
7. Yamaguchi, K. and Yamaguchi, M. (1984) The replication origin of pSC 101: The nucleotide sequence of the ori region. *Gene* **29**, 211–219
8. Armstrong, K.A., Acosta, R., Lender, E., Machida, Y., Pancotto, M., McComick, M., Ohtsubo, H., and Ohtsubo, E. (1984) A 37 × 10^3 molecular weight plasmid-encoded protein is required for replication and copy number control in the plasmid pSC101 and its temperature-sensitive derivative pHS1. *J. Mol. Biol.* **175**, 331–347
9. Vocke, C. and Bastia, D. (1983) Primary structure of the essential replicon of the plasmid pSC101. *Proc. Natl. Acad. Sci. USA* **80**, 6557–6561
10. Sugiura, S., Tanaka, M., Masamune, Y., and Yamaguchi, K. (1990) DNA binding properties of purified replication initiator protein (Rep) encoded by plasmid pSC101. *J. Biochem.* **107**, 369–376
11. Manen, D., Gonzalez, U.L.C., and Caro, L. (1992) Monomers and dimers of the RepA protein in plasmid pSC101 replication: Domains in RepA. *Proc. Natl. Acad. Sci. USA* **89**, 8923–8927
12. Ohkubo, S. and Yamaguchi, K. (1995) Two enhancer elements for DNA replication of pSC101, *par* and a palindromic binding sequence of the Rep protein. *J. Bacteriol.* **177**, 558–565
13. Yamaguchi, K. and Masamune, Y. (1985) Autogenous regulation of synthesis of the replication protein in plasmid pSC101. *Mol. Gen. Genet.* **200**, 362–367
14. Linder, P., Churchward, G., Xia, G.X., Yu, Y.Y., and Caro, L. (1985) An essential replication gene, *repA*, of plasmid pSC101 is autoregulated. *J. Mol. Biol.* **181**, 383–393
15. Vocke, C. and Bastia, D. (1985) The replication initiator protein of plasmid pSC101 is a transcriptional repressor of its own cistron. *Proc. Natl. Acad. Sci. USA* **82**, 2252–2256
16. Ishige, F., Kobayashi, Y., Sugiura, S., Masamune, S., and Yamaguchi, K. (1987) Properties of the replication protein in copy number mutants of plasmid pSC101. *J. Pharmacobio-Dyn.* **10**, 39
17. Xia, G.X., Manen, D., Goebel, T., Linder, P., Churchward, G., and Caro, L. (1991) A copy number mutant of plasmid pSC101. *Mol. Microbiol.* **5**, 631–640
18. Ingmer, H. and Cohen, S.N. (1993) Excess intracellular concentration of the pSC101 RepA protein interferes with both plasmid DNA replication and partitioning. *J. Bacteriol.* **175**, 7834–7841
19. Wadood, A., Dohmoto, M., Sugiura, S., and Yamaguchi, K. (1997) Characterization of copy number mutants of plasmid pSC101. *J. Gen. Appl. Microbiol.* **43**, 309–316
20. Murakami, A., Sugiura, S., and Yamaguchi, K. (1997) DNA bending of pSC101 *ori* induced by Rep, a replication initiator protein and IHF. *J. Gen. Appl. Microbiol.* **43**, 189–197
21. Furuno, S., W.-Murakami, Y., T.-Suzuki, N., and Yamaguchi, K. (2000) Negative control of plasmid pSC101 replication by increased concentrations of both initiator protein and iterons. *J. Gen. Appl. Microbiol.* **46**, 29–37
22. Messing, J. (1983) New M13 vectors for cloning. *Methods Enzymol.* **101**, 20–78
23. Casadaban, M.J. and Cohen, S.N. (1980) Analysis of gene control signals by DNA fusion and cloning in *E. coli*. *J. Mol. Biol.* **138**, 179–207
24. Akiyama, M., Horiuchi, T., and Sekiguchi, M. (1987) Molecular cloning and nucleotide sequence of the *mutT* mutator of *Escherichia coli* that causes A:T to C:G transversion. *Mol. Gen. Genet.* **206**, 9–16
25. Yanish-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119
26. Fueki, T., Sugiura, S., and Yamaguchi, K. (1996) Open strands adjacent to iterons promote the binding of the replication initiator protein (Rep) of pSC101 to the unit sequence of the iterons *in vitro*. *Biochim. Biophys. Acta* **1305**, 181–188
27. Boliver, F., Rodrigues, R.L., Green, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Grosa, J.H., and Falkow, S. (1977) Construction and characterization of new cloning vehicles. II. A multi-purpose cloning system. *Gene* **2**, 95–113
28. Miller, J. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
29. Churchill, M.E.A. and Travers, A.A. (1991) Protein motifs that recognize structural features of DNA. *Trends Biochem. Sci.* **16**, 92–97
30. Datta, H.J., Khatri, G.S., and Bastia, D. (1999) Mechanism of recruitment of DnaB helicase to the replication origin of the plasmid pSC101. *Proc. Natl. Acad. Sci. USA* **96**, 73–78
31. Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C., and Miki, K. (1999) Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 Å resolution. *EMBO J.* **18**, 4597–4607
32. Ingmer, H., Fong, E.L., and Cohen, S.N. (1995) Monomer-dimer equilibrium of the pSC101 RepA protein. *J. Mol. Biol.* **250**, 309–314