Molecular analysis of the replication region of the conjugative *Streptococcus agalactiae* plasmid pIP501 in *Bacillus subtilis*. Comparison with plasmids pAMβ1 and pSM19035

Sabine Brantl[1], Detlev Behnke[1] and Juan C.Alonso*
Max-Planck-Institut für molekulare Genetik, Ihnestrasse 73, D-1000 Berlin 33, FRG and [1]Zentralinstitut für Mikrobiologie und experimentelle Therapie der AdW, Beutenbergstraße 11, DDR-6900 Jena, GDR

Received May 12, 1990; Accepted July 19, 1990 EMBL accession no. X17655

ABSTRACT

The large conjugative plasmid pIP501 was originally isolated from *Streptococcus agalactiae*. To study the molecular basis of pIP501 replication we determined the nucleotide sequence of a 2.2 kb DNA segment which is essential and sufficient for autonomous replication of pIP501 derived plasmids, in *Bacillus subtilis* cells. This region can be divided into two functionally discrete segments: a 496 bp region (*oriR*) that acts as an origin of replication, and a 1488 bp segment coding for an essential replication protein (RepR). The RepR protein, which has a molecular mass of 57.4 kDa, could complement in trans a thermosensitive replicon bearing the pIP501 origin. Chimeric Rep proteins and replicons were obtained by domain swapping between *rep* genes of closely related streptococcal plasmids belonging to the *inc*18 group (pIP501, pAMβ1 and pSM19035). The chimeras were functional in *B. subtilis*.

INTRODUCTION

The genomic organization and protein sequence conservation of small, high copy number, non-conjugative plasmids of Gram-positive bacteria led to their classification into four distinct families (1–3). All plasmids replicate via a rolling circle mechanism like the single-stranded *Escherichia coli* bacteriophages (see 3, 4 for reviews). Cloning of large DNA fragments (> 6 kb) into such plasmids has been found to cause high structural instability of the recombinant molecules (5, 6). In order to overcome this problem we, therefore, searched for plasmids which have a different mode of replication in *Bacillus subtilis*. Large conjugative plasmids like the broad host range plasmid pIP501 (7) which were originally isolated from various streptococcal species have been used to construct smaller derivatives suitable as cloning vectors in these bacteria (8). Plasmid pIP501 confers resistance to erythromycin (Em) and chloramphenicol (Cm) and is considered to be evolutionarily related to pAMβ1 (9) and to the non-conjugative plasmid pSM19035 (10). The replication region of pIP501 (30.2 kb) was localized by deletion analysis on a 2.1 kb fragment (8). Plasmids carrying this region together with antibiotic resistance markers can be introduced into *B. subtilis* by transformation (8).

In this paper we report the molecular analysis of the replication functions of pIP501 and their functional relationship to those of plasmids pAMβ1 and pSM19035.

MATERIALS AND METHODS

Bacterial strains, phage and plasmids

B. subtilis strain YB886 *trp*C2 *met*B5 *amy*E *xin*-1 *att*SPb and it isogenic derivative *rec*E4 (YB1015) have been previously described (11).

The following *E. coli* strains were used: BL21DE3 *hsd*S20 *gal* [λ-T7RNApol] (12) and JM103 [Δ(*lac pro*)] *thi str*A *sup*E *end*A *sbc*B *hsd*R F'*tra*D36 *pro*AB *lac*Iq λM15 (13).

Bacteriophage SPP1v40 was used for plasmid transduction (14). The plasmids used are listed in Table 1.

Growth, transformation and transduction

Cultures of *B. subtilis* were grown and transformed as described by Rottländer and Trautner (22). *E. coli* cells were transformed essentially as described by Maniatis et al. (23). When necessary, plasmids were moved into the desired genetic background by transduction with bacteriophage SPP1v40 as described by Deichelbohrer et al. (24). A transduction enhancement was observed when homology was provided between the phage and the plasmid (24; Table 3). If not stated otherwise the temperature of cell growth was 37°C.

DNA preparation and manipulation

Plasmids were purified as previously reported (25, 26).

Restriction endonucleases and DNA modification enzymes were obtained commercially and used as specified by their suppliers.

* To whom correspondence should be addressed

Table 1: Plasmids used

Plasmid	Description	Reference
E. coli		
pUC118	multiple purpose cloning vehicle Ap^R,	15
pUC18	multiple purpose cloning vehicle Ap^R,	13
pLysE	pACYC184 based-T7 lysozyme gene,Cm^R	16
pT712	pUC12 based vector with a T7 promoter, Ap^R	GIBCO-BRL
B. subtilis		
pC194	naturally occuring,Cm^R	17
pSA0331	a Tsr derivative of pT181,Tc^R	18
pGB301	a pIP501 derivative, Cm^R	8
pGB354	a pIP501 derivative, Cm^R	8
pGB3631	a pIP501 derivative, Cm^R	19
pDB1021	a pSM19035 derivative, EmR	20
pBT21	a pDB1021 derivative,PmR	Ceglowski, unpubl.
pIL252	a pAMβ1 derivative, EmR	21

Abbreviations: Ap, ampicillin; Cm,chloramphenicol; Em, erythromycin; Pm, phleomycin; Tc, tetracycline.

Homo- and heteroplasmid growth.

The thermosensitive replicon (pSA0331, Tc^R) containing a DNA fragment from plasmid pIP501 was introduced into the *rec*E4 strain by transformation; whereas the helper plasmid (pGB354, Cm^R) was introduced into the same cells by transduction. Cells bearing both plasmids were grown at 30°C with selection for the presence of both antibiotic resistance markers to about 1.0×10^8 colony-forming units/ml. After diluting the culture 10^4-fold in TY media without antibiotics, an aliquot was shifted to 47°C and both cultures (30°C and 47°C) were allowed to grow for further 12 to 20 generations. The cells were then spread onto plates containing tetracycline (Tc) and incubated at 30°C or 47°C.

Plasmid transduction and the incompatibility testing were done as previously described (27).

DNA sequence determination

For DNA sequencing of the pIP501 replication region a 2.2 kb *Kpn*I-*Eco*RI DNA fragment derived from plasmid pGB3631 was cloned into pUC118 and pUC119. Nested deletions of the cloned *Kpn*I-*Eco*RI fragment were generated with exonuclase III (28) and used as templates. The DNA sequence was determined by the dideoxynucleotide chain-termination method of Sanger et al. (29) as outlined in the legend to Fig. 1. Both strands of the *Kpn*I-*Eco*RI DNA fragment were sequenced. Several specific primers were synthesized on an applied Biosystem DNA synthesizer.

Construction of new plasmids

From previous deletion analysis the replication functions were known to be localized on the smaller 2.2 kb *Kpn*I-*Eco*RI DNA fragment of plasmid pGB3631. This DNA fragment was cloned into plasmid pUC118 or pUC119119, generating plasmids pUC118-F and pUC119-F.

Plasmid pF was derived by joining pUC118-F cleaved with *Sma*I to the tetracycline-resistant (Tc^R) thermosensitive replicon of plasmid pSA0331 cleaved with *Tha*I (Fig. 1B). The latter plasmid was also used to construct plasmids p4, p6, p7 and p8 (Fig. 1C), by fusing it with plasmids pUC118-4, pUC118-6, pUC118-7 and pUC118-8, respectively. The four pUC118-F derivatives carried nested deletions of 358, 613, 674 and 947 nucleotides, respectively, extending from the *Kpn*I site into the *Kpn*I-*Eco*RI fragment of pGB3631. Plasmids pFi and p4i were generated by filling in the *Hin*dIII site (coordinate 1623) located within the 2.2 Kb *Kpn*I-*Eco*RI DNA fragment of plasmids pF and p4, respectively. Such a frameshift introduces a stop codon after the incorporation of another 6 residues.

To construct expression plasmids for the RepR protein the 2.2 kb *Pst*I-*Eco*RI DNA fragment of pUC118-F was inserted into plasmid pT712 (Table 1), thereby generating pBT117. Plasmid pBT119 was constructed by *in vitro* deletion of the small *Bsm*I fragment (nucleotides 1824 to 1914) from pBT117. To construct pBT118, pBT117 was digested with *Bsm*I, subsequently treated with T4 DNA polymerase and ligated. Among the transformants a clone was selected that lacked the *Bsm*I site. Thus, in pBT117 the entire repR open reading frame is present, whereas the open reading frames present in pBT118 and pBT119 code for RepR proteins shortened by 30 and 91 amino acids, respectively.

Plasmid pBT48 was constructed as follows: a *Pvu*II-*Hae*III DNA fragment of plasmid pUB110 which codes for a phleomycin resistance determinant that is expressed in both *E. coli* (5 μg/ml) and *B. subtilis* (0.2 μg/ml) was joined to pUC18 cleaved with *Hin*dII. Plasmid pBT48 does not replicate in *B. subtilis*. For the construction of chimeric *rep* genes various domains were obtained from from *rep* genes of plasmids pUC118-F (pIP501), pDB1021 (pSM19035) and pIL252 (pAMβ1). Chimera 3, for example, was constructed by joint cloning of a 1.6 kb *Kpn*I-HindIII DNA fragment from pDB1021, and a 0.64 kb *Hin*dIII-*Eco*RI fragment from pUC118-F into pBT48 cleaved by *Kpn*I-*Eco*RI. After transformation of *E. coli* JM103 cells the correct plasmids were identified by their size and restriction fragment analysis for the proper DNA segments. Plasmids carrying the chimeric *rep* genes were introduced into *B. subtilis* YB886 cells by transformation and selection for the phleomycin marker.

Analysis of plasmid encoded protein

The BL21DE3 [pT712] host-vector expression system (12) was used to specifically label the RepR proteins produced by pBT117, pBT118 and pBT119 (16). BL21DE3 carrying either pBT117, pBT118, pBT119 or pT712 were grown to OD_{560}= 0.4 in minimal medium and induced with 5 mM IPTG. After 30 min, 200 μg/ml rifampicin were added and the cells were further incubated for 90 min.

Protein labeling was carried out essentially as described by Alonso and Tailor (30). Proteins were separated on denaturing SDS-polyacrylamide gels as described by Laemmli (31).

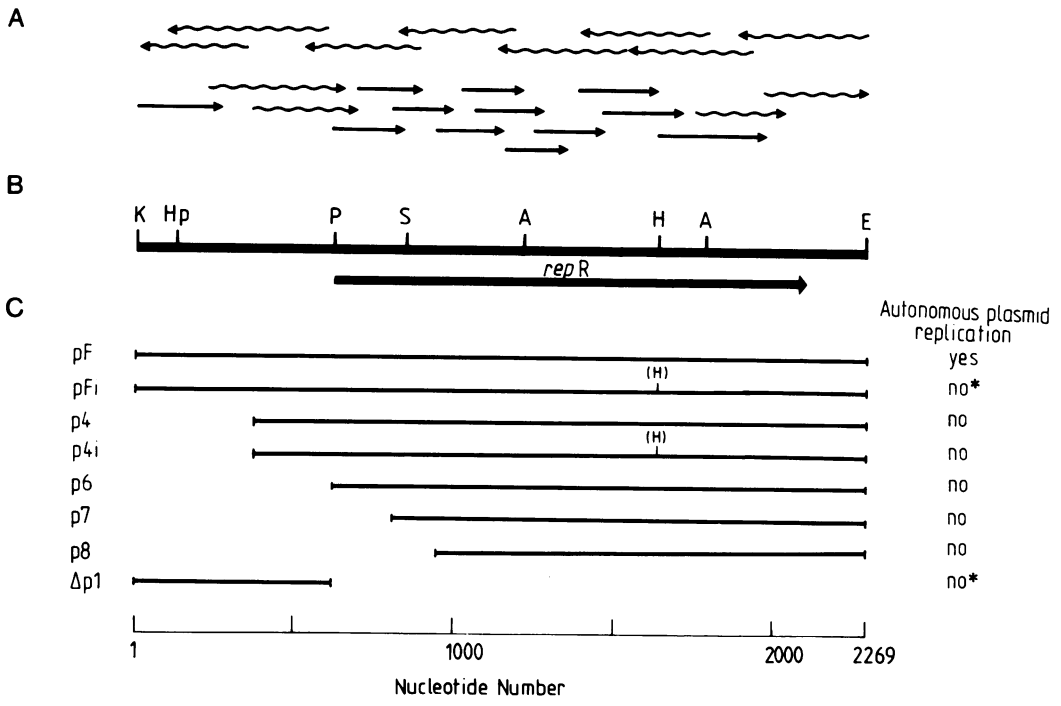

Figure 1: DNA sequencing strategy and physical map of plasmid pF and its derivatives. (A) the arrows represent the direction and extent of the various sequence determinations. The wavy arrows are sequences derived from M13 clones using the sequenase system and synthetic primers. The straight arrows indicate sequences obtained with nonspecific deletion derivatives generated with exonuclease III using the universal primer and DNA polymerase I Klenow fragment. (B) The filled bar denotes the DNA region used in this study. The extent and directionality of the repR gene is denoted by a black arrow. Abbreviations A, AccI; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; K, *Kpn*I; P, *Pvu*II; S, *Sna*BI and (H); denotes that the *Hin*dIII site was filled in. (C) physical map of pF derivatives. Yes and no denotes the ability or inability to replicate *in vivo* autonomously in *B. subtilis*, respectively. The presence of the asterisk indicates that plasmid replication occurs when the RepR product is provided in trans.

Computer analysis

Comparison of primary protein structure, the investigation of DNA and RNA secondary structures, transcription terminators etc. were carried out with the computer software package of the University of Wisconsin Genetic Computer Group on a VAX computer (32, 33).

RESULTS

DNA sequence analysis of the replication region of pIP501

In Fig. 1A is shown the strategy used to sequence the 2.2 Kb *Kpn*I-*Eco*RI DNA fragment of pGB3631 which has been shown to be necessary and sufficient for autonomous plasmid replication (8).

Analysis of the nucleotide sequence (Fig. 2) revealed the presence of only one open reading frame. This open reading frame starts with an ATG at position 620 and extends 1488 nucleotides to a TGA stop codon at nucleotide 2110. A putative RNA polymerase consensus regions at nucleotides 569−574 (−35) and 587−592 (−10), and a ribosomal binding site (RBS) at coordinates 604−610 are located immediately upstream of the coding sequence. The RBS which has a free energy of association calculated by the empirical formula of Tinoco et al. (34) of −16.0 Kcal/mol has a window of 9 bp to the first initiation codon. A structure reminiscent of a transcription terminator is located downstream of the coding sequence, with a hairpin forming region beginning at nucleotide 2120. Translation of this open reading frame results in a polypeptide of 496 amino acids with a predicted molecular weight of 57.4 kDa. As this protein is apparently essential for plasmid replication we designated the open reading frame *rep*R (see below).

Inspection of the primary structure of RepR, deduced from the DNA sequence, revealed between residues 120 to 141 a helix-turn-helix DNA binding motif frequently observed in DNA binding proteins. A glycine residue is present at position 9 [numbering follows the convention of Pabo and Sauer (35)], non-charged residues at position 4 and 15 and residues 2−8 and 15−20 are not occupied by proline (36).

A computer based comparison of RepR with primary protein sequences currently available in the NBRF database (release 22.0) was performed and no significant homologies were detected. As shown in Fig. 3, a high degree of homology (97.4%) was, however, observed with the replication initiation protein (RepS) of the closely related, but non-conjugative plasmid pSM19035 of *Streptococcus pyogenes* (37, 38). The proteins differ in only 11 residues, most of which are located close to the C-terminal. Furthermore, the helix-turn-helix motif is occupied by the same conserved region in RepR and RepS (37, Fig. 3).

Expression of the *rep* gene

To visualize and possibly overproduce the Rep product in *E. coli*, the *rep*R gene was placed under the control of a T7 RNA polymerase promoter on the expression plasmid pT712. The resulting plasmids which carried either the entire open reading frame (pBT117) or lacked 30 (pBT119) and 91 codons (pBT118) within the C-terminal part of the *rep*R gene or the vector itself were introduced into a T7 RNA polymerase expression strain (BL21DE3) bearing plasmid pLysE (12, 16).

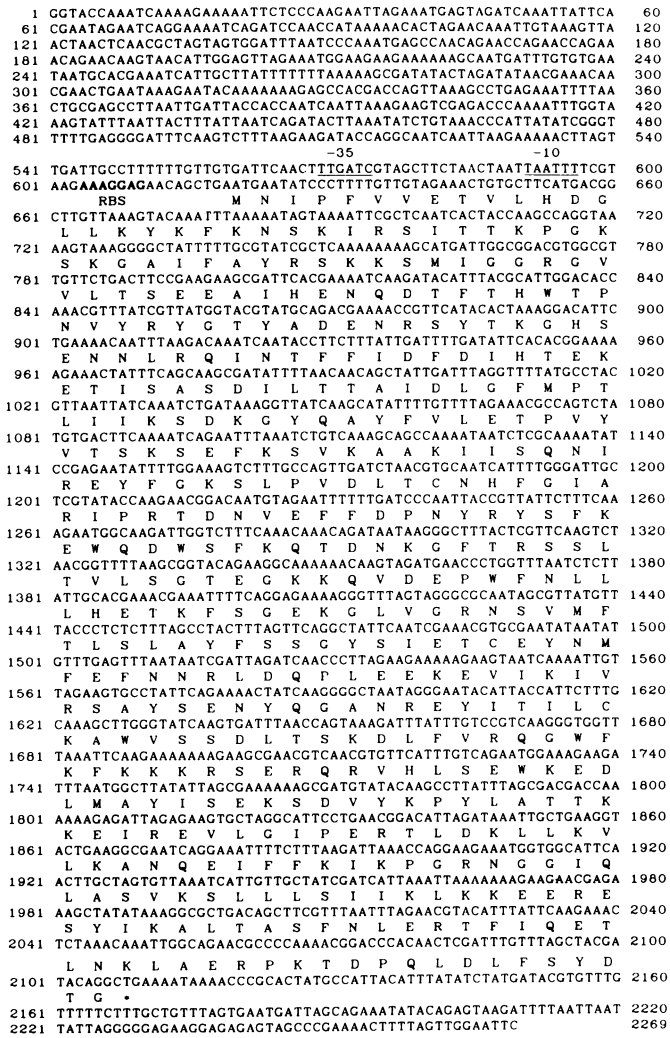

Figure 2: Nucleotide sequence of the minimal replicon of pIP501 and predicted amino acid sequence of the *rep*R gene. The putative promoter (−35 and −10 consensus regions) and the potential ribosomal binding site (RBS) preceding the open reading frame are marked. The translation stop codon is indicated by an asterisk.

SDS-polyacrylamide gel electrophoresis of RepR proteins is presented as an autoradiogram in Fig. 4. A polypeptide of about 57 kDa was detected from extracts of BL21DE3 cells harboring plasmid pBT117 (Fig. 4B). This size corresponded well with that expected for RepR on the basis of the DNA sequence data. When BL21DE3 cells harboring either plasmid pBT118 or pBT119, which contained the truncated *rep* gene were analyzed, the expected smaller proteins (46 and 54 kDa, respectively) instead of the 57 kDa polypeptide were detected (Fig. 4). A small polypeptide of about 14 kDa which was only observed from BL21DE3 cell bearing plasmid pBT117 was detected. We assume that this peptide originates either by abortive termination or corresponds to a degradation product.

The presence of plasmids pBT118 and pBT119, which carried deletion variants of the *rep*R gene permitted normal cell growth, whereas the presence of the pBT117 (wild type *rep* gene) highly impaired cell growth and reduced the efficiency of protein labelling. This observation suggested that the RepR expression is toxic to the cells. Consistent with that are the facts that deletion of the C-terminal part of *rep*R apparently releases *E. coli* cells

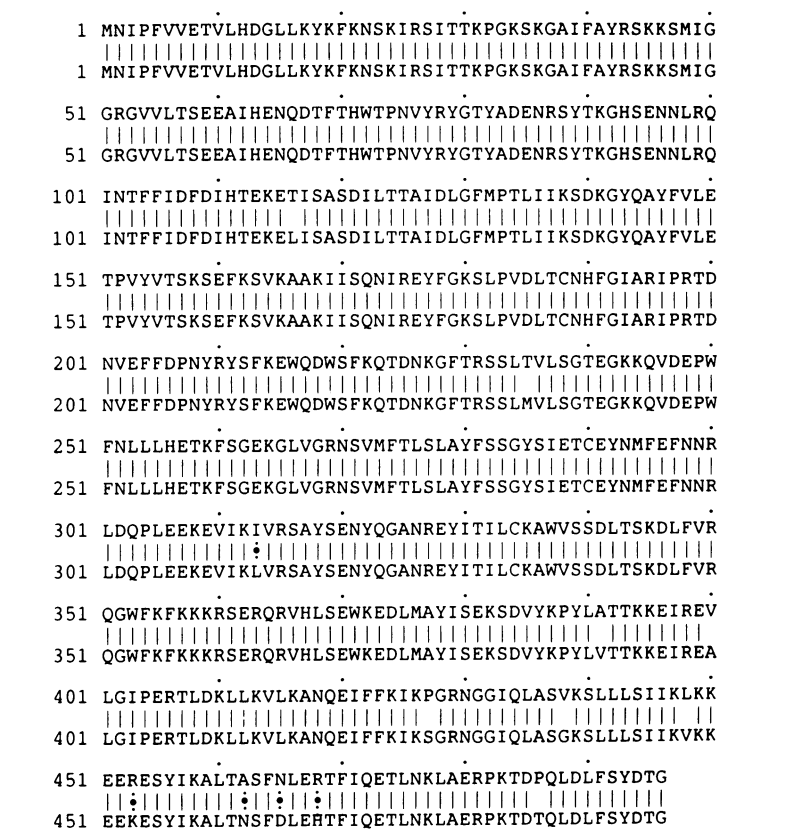

Figure 3: The predicted amino acid sequence of RepR (top) compared to that of RepS of plasmid pSM19035 (bottom). Vertical bars denote identical residues, dots conserved residues, and the absence of either bars or dots indicates a non conserved residue

from this toxic effect and such effect is only detected in *E. coli* cells harboring the gene 1 of bacteriophage T7 (data not shown).

Mapping of the pIP501 replication region

It has previously been shown, that both pIP501-derivatives pGB354 and pGB3631 can autonomously replicate in *B. subtilis* (8). Due to the deletion of the *Kpn*I-HpaI DNA fragment the replication region of pGB356 is 120 bp shorter than that of plasmid pGB3631 (8).

By deletion analysis we attempted to further narrow down the minimal replication region of pIP501-derived plasmids in *B. subtilis*. Since it was unknown whether pGB354 codes for a positive or a negative trans-acting effector essential for plasmid replication, we based our construction on the themosensitive replicon of the Tc^R plasmid pSA0331. To this end, plasmid pF was generated by fusing pUC118-F carrying the 2.2 kb *Kpn*I-*Eco*RI DNA fragment (see Fig. 1B) of plasmid pGB3631, with the thermosensitive replicon pSA0331. *B. subtilis* transformants harboring pF formed Tc^R colonies both at 30°C and after a temperature shift to 47°C. In contrast cells bearing only pSA0331 plated with an efficiency of about 10−4 at 47°C. This suggested that the replication of pF at the higher temperature was directed from the origin located within the *Kpn*I-*Eco*RI fragment of pGB3631.

To correlate the pIP501 replication origin region within this fragment, progressing deletions were introduced from the *Kpn*I site of pF. Plasmids were then analysed for their ability to autonomously replicate (Fig. 1C). The pF-derived plasmids pFi,

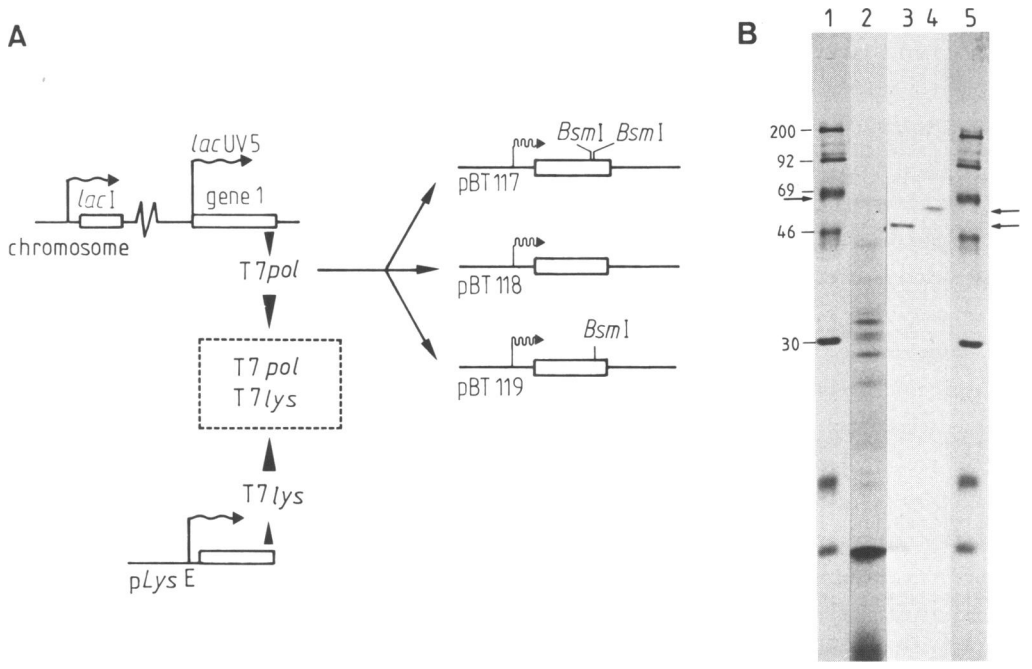

Figure 4: Plasmid encoded polypeptides in *E. coli* BL21DE3 strain. (A) schematic representation of the T7 expression system and the plasmids used. (B) Autoradiogram of [^{14}C]labelled extracts. Lane 2: extract of cells carrying plasmid pBT117, lane 3: pBT118, lane 4: pBT119. Lanes 1 and 5: M_r standard proteins. The arrows mark the RepR product and its truncated forms respectively.

p4, p4i, p6, p7 and p8 could not transform competent cell at 47°C. As expected, Tc^R colonies could be obtained at 30°C, however, the resistance phenotype is lost (< 0.1%) when the cells are grown for about 20 generation at 47°C. Except for pFi, the presence of a helper plasmid (pGB354), introduced by transduction, could not reverse such plasmid loss. Since p4 is only 359 nucleotides smaller than the autonomously replicating plasmid pF, we conclude that the extragenic DNA sequence absent from or truncated in p4 was required in *cis* for DNA replication. In order to further confirm this region the smaller *Kpn*I-PvuII DNA fragment of pGB3631 (615 nucleotides) was joined to the thermosensitive pSA0331 replicon, generating plasmid p1 (Fig. 1C). *B. subtilis* recE4 cells carrying the plasmid pF, pFi, p4, p1 or pSA0331 together with the helper plasmid pGB354 were initially grown at 30°C. The cultures were diluted onto antibiotic-free media, shifted up to 47°C, and allowed to grow for further 12 generations. As summarized in Table 2 strains bearing the plasmid pairs p4 and pGB354 or pSA0331 and pGB354 plated at a very low efficiency on Tc containing plates at 47°C (< 10−4), whereas more than 50% of the cells bearing either p1 and pGB354) or pFi and pGB354 plated with similar efficiency in the presence or absence of Tc at 47°C. Strains carrying only single plasmids, except for those harboring pF, plated at very low efficiency at 47°C.

Since we can distinguish between autonomously replicating plasmids (pGB354 and pF) and those that replicate in the presence of a plasmid helper (p1 and pFi) we conclude that the latter plasmids contained the replication origin (*ori*R) while the helper plasmid provided a positive trans-acting factor essential for replication rather than a repressor of runaway replication. This trans-acting factor was most likely the RepR protein. The *ori*R region, therefore, was localized between coordinates 119−615 (see Fig. 2). This was the only region shared between pGB354 and p1. An inverted repeat and two set of direct repeat DNA

Table 2: Plasmid plating efficiency in the presence or absence of a plasmid helper (pGB354) in the *rec*E4 genetic background.

Plasmid	Temperature °C	Colony forming units/ml
p1	30	1.2×10^8
p1	47	$<1.0\times10^4$
p1[a]	47	8.1×10^7
p4	30	1.1×10^8
p4	47	$<1.0\times10^4$
p4[a]	47	$<1.0\times10^4$
pF	30	1.1×10^8
pF	47	1.6×10^8
pF[a]	47	5.3×10^7
pFi	30	1.1×10^8
pFi	47	$<1.0\times10^4$
pFia	47	4.6×10^7
pSA0331	30	1.0×10^8
pSA0331[a]	47	$<1.0\times10^4$

([a]) in the presence of a plasmid helper pGB354. In the heteroplasmid condition a high segregational instability as well as a poor replication rate was observed.

sequences were identified within this interval (Fig. 5). A tetranucleotide AGAA was present in the direct repeat boxes. No DnaA binding motif was detected at *ori*R. Furthermore, the nucleotides absent in plasmid p4 are shown in bold face cases in Fig. 5.

Plasmid pSM19035 (37, 38) has the same genomic organization as plasmid pIP501. As shown in Fig. 5, a sequence homology of about 96% between *ori*R and its equivalent DNA segment of plasmid pSM19035 was observed.

Construction of chimeric replicons

By bacteriophage SPP1-mediated transduction we were able to show that pSM19035, pAMβ1 and pIP501 derivatives belong

```
119 TAACTAACTCAACGCTAGTAGTGGATTTAATCCCAAATGAGCCAACAGAA 168
    ||||||||||||||||||||||||||||||||||||||||||||||||||
137 TAACTAACTCAACGCTAGTAGTGGATTTAATCCCAAATGAGCCAA..... 181

169 CCAGAACCAGAAACAGAACAAGTAACATTGGAGTTAGAAATGGAAGAAGA 218
    ||||||||||| |||||||||||||||| ||||||||||||||||||||
182 .CAGAACCAGAATCAGAACAAGTAACATTGGATTTAGAAATGGAAGAAGA 230

219 AAAAAGCAATGATTT.GTGTGAATAATGCACGAAATCATTGCTTATTTTT 267
    ||||||||||||||| |||| || |||||||||||||||||||||||||
231 AAAAAGCAATGATTTCGTGTAAAAAATGCACGAAATCATTGCTTATTTTT 280

268 TTAAAAAGCGATATACTAGATATAACGAAACAACGAACTGAATAAAGAAT 317
    || ||||||||||||||||||||||||||||||||||||||||||||||
281 TTTAAAAGCGATATACTAGATATAACGAAACAACGAACTGAATAAAGAAT 330

318 ACAAAAAAAGAGCCACGACCAGTTAA...AGCCTGAGAAATTTTAACTGC 364
    |||||||||||||||||||| | |   | | || | | |||
331 ACAAAAAAAGAGCCACGACCAACTTATTTTCCGCCAAGAAAATAAGTTGC 380

365 GAGCCTTAATTGATTACCACCAATCAATTAAAGAAGTCGAGACCCAAAAT 414
    ||||||||||||||||||||||||||||||||||||||||||||||||
381 GAGCCTTAATTGATTACCACCAATCAATTAAAGAAGTCGAGACCCAAAAT 430

415 TTGGTAAAGTATTTAATTACTTTATTAATCAGATACTTAAATATCTGTAA 464
    ||||||||||||||||||||||||||||||||||||||||||||||||
431 TTGGTAAAGTATTTAATTACTTTATTAATCAGATACTTAAATATCTGTAA 480

465 ACCCATTATATCGGGTTTTTGAGGGGATTTCAAGTCTTTAAGAAGATA.C 513
    ||||||||||||||||||||||||||||||||||||||||||||| |
481 ACCCATTATATCGGGTTTTTGAGGGGATTTCAAGTCTTTAAGAAGATACC 530

514 CAGGCAATCAATTAAGAAAAACTTAGTTGATTGCCTTTTTTGTTGTGATT 563
    ||||||||||| ||||||||||||||||||||||||||||||||||||
531 CAGGCAATCAATTAAGAAAAACTTAGTTGATTGCCTTTTTTGTTGTGATT 580

564 CAACTTTGATCGTAGCTTCTAACTAATTAATTTTCGTAAGAAAGGAGAAC 613
    ||||||||||||||||||||||||||||||||||||||||||||||||
581 CAACTTTGATCGTAGCTTCTAACTAATTAATTTTCGTAAGAAAGGAGAAC 630

614 AG 615
    ||
631 AG 632
```

Figure 5: Nucleotide sequence comparison between the replication origin region of plasmids pGB354, p1 (upper) and its corresponding region of pSM19035 (lower sequence). Dots between or under the nucleotides indicate gaps that have been introduced to maintain an optimal alignment. The presence or absence of vertical bars indicate identical or disimilar nucleotides respectively. A superscript dot was placed every ten nucleotides. Direct repeats sequences are boxed and inverted repeats are indicated by horizontal arrows. Bold characters indicate those which are absent in plasmid p4.

to the same incompatibility group designated *inc*18 (data not shown). To further study the degree of functional similarity betwen plasmids of the *inc*18 group several chimeras were constructed by exchanging fragments covering part of the Rep protein alone or together with the origin region (Fig. 6).

Construction of all chimeras was done in *E. coli* with plasmid pBT48 as the cloning vehicle. Figure 6 summarizes the structure of the different chimeras constructed. Chimeras 1 to 3 were obtained by reciprocal fusions at *Sna*BI or *Hin*dIII sites located at identical positions within the Rep coding region of pSM19035 (pDB1021) and pIP501 (pUC118-F). Chimera 1, carried the pIP501 replication origin region and a *rep* fusion gene reciprocal to that of chimera 3. The Rep protein of chimeras 1 and 2 differ from repR by 8 and 3 amino acids, while the hybrid Rep protein of chimera 2 was identical to RepR (37; see also Fig. 3 and Fig. 5). The DNA sequence between the *Hpa*I and *Pvu*II sites of pSM19035 and pIP501 (Fig. 6) bear the putative origins of replication. These sequences differ by only 31 nucleotides. Most of those differences are located within a 30 nucleotide stretch (coordinates 335−365, Fig. 5). The plasmid chimeras 1 to 3 could be introduced into *B. subtilis* competent cells with equal efficiency. Upon reisolation their structural integrity was confirmed by restriction enzyme analysis. We, therefore, conclude that: i) the N- and C-terminal part of the *inc*18 Rep proteins are interchangeable, and ii) RepR-like proteins can utilize the pSM19035 origin of replication. To further substantiate these conclusions, chimeras 4 and 5 (Fig. 6) were constructed between pIP501 and pAMβ1 replication regions. We assumed that the location of identical restrictions sites within the replication region of pAMβ1 allowed the construction of chimeras without shifting the reading frame of the putative pAMβ1 *rep* gene. Chimeras 4 and 5 were constructed by reciprocal fusions at the *Hin*dIII sites of the pIP501 and pAMβ1 replication regions (Fig. 6). Both chimeras were found to be functional in mediating plasmid replication in *B. subtilis*. Their structural integrity was confirmed by restriction enzyme analysis. These results, therefore, confirmed the above conclusions and extended the

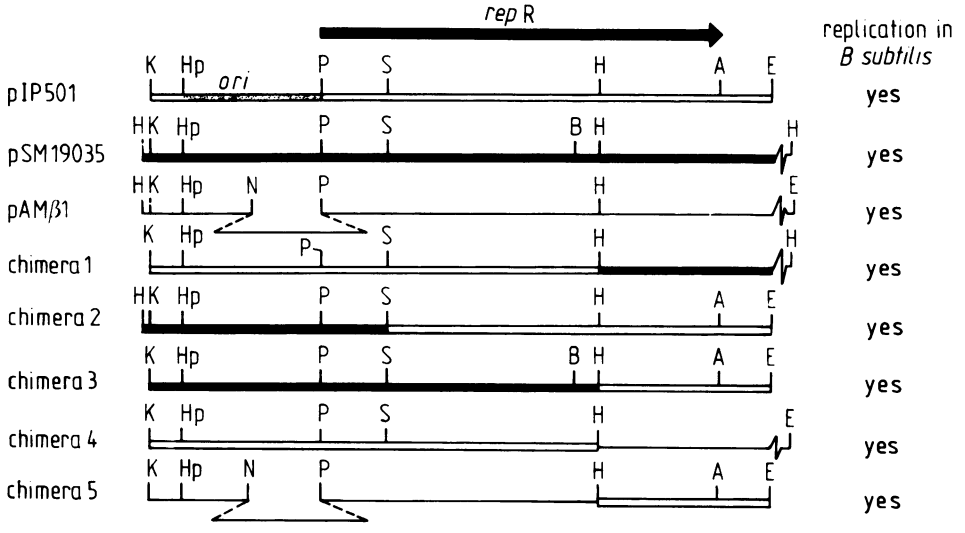

Figure 6: Alignment of the plasmids pIP501, pSM19035 and pAMβ1 replication region and construction scheme for plasmid chimeras. Open bars denote pIP501 DNA, filled bars pSM19035 DNA and the thin line pAMβ1 DNA. The source of those DNAs were plasmids pGB354, pDB1021 and pIL252, respectively. Shaded regions within the open bar denote the plasmid origin (*ori*) region. The extent and directionality of the replication effector (RepR in the case of plasmid pIP501) is indicated by an arrow. Abbreviations A, *Ava*II; B, *Bcl*I; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; K, *Kpn*I; N, *Nsi*I; P, *Pvu*II; and S, *Sna*BI.

interchangeability of Rep domains to pAMβ1. Furthermore it appeared likely, that the extra 300 bp present between the *Hpa*I and *Pvu*II site of pAMβ1 (Fig. 6) are dispensable for pAMβ1 replication.

DISCUSSION

The DNA sequence of the replication region of the conjugative plasmid pIP501 of *S. agalactiae* has been determined. The minimal replication region necessary for propagation in *B. subtilis* wild type cells, was narrowed down to a DNA fragment of 2150 bp. Within this region a gene was identified coding for a positive trans-acting effector, RepR, and a cis-acting segment that provided the plasmid replication origin (*ori*R). *ori*R was localized within a 496 bp DNA segment upstream of the *rep*R gene. Based on circumstantial evidence the location of *ori*R can be further pinpointed to the 327 bp between nucleotide 119 and 359 (Fig. 5). This conclusion is based on the following experimental data: i) a plasmid (p4, Fig. 1) lacking the first 359 bp of the replication region cannot replicate even in the presence of a helper plasmid at 47°C, ii) the DNA sequence between nucleotides 335 and 365 which shows a high degree of variability appeared to be non-essential as the pIP501 and pSM19035 were interchangeable, and iii) a 300 bp DNA fragment of the *inc*18 plasmid pAMβ1 which is dispensable for replication, maps between the *Nsi*I site 260 bp away from the *Hpa*I site. Examination of the DNA sequence between nucleotides 119 and 359 revealed an inverted repeat and two sets of direct repeat sequences. Whether their structure has any functional significance remains to be shown. Preliminary results of deletion analysis of this region, however, support that these inverted repeats are necessary for replication (Brantl and Ceglowski, unpublished).

The nucleotide sequence analysis of the minimal replication region revealed that the presence of only one gene, the *rep*R gene. The putative promoter region, the ribosome binding site and the terminator of the *rep*R gene are all comparable to the consensus sequences or structures. Under denaturating conditions the molecular weight of the protein encoded by the *rep*R gene is 57 kDa, which is consistent with the size of the gene product predicted from the DNA sequence (57.4 kDa). We denote the gene encoding this replication effector as *rep*R. Codon usage within repR did not reveal significant differences from other genes from Gram-positive bacteria (39).

The autonomously replicating plasmid pF differs from the passively replicated plasmid pFi in only 4 nucleotides, within a segment outside the replication origin region. Since, pFi which lacked 135 codons within the C-terminal part of the *rep*R gene does not replicate, in the absence of a helper plasmid, we concluded that RepR is a positive trans-acting effector. RepR contains a putative helix-turn-helix DNA binding motif which is close to the consensus sequence derived from a number of other DNA binding proteins (36). *In vivo* replication depends on this RepR protein. It is likely to interact with sequences at the oriR segment. A comparison at the DNA and amino acid sequence level revealed a high degree (97.3%) of homology between RepR and its counterpart RepS (37, 38) coded for by plasmid pSM19035 of *S. pyogenes*.

Characterization of the minimal replication region as presented in this communication does not exclude the presence of secondary origins of replication on the large parental plasmids pIP501, pSM19035 or pAMβ1.

Plasmids pIP501, pSM19035 and pAMβ1 belong to the same incompatibility group *inc*18. Plasmid pIP501 and the large non-conjugative plasmid pSM19035 share a similar structure at their minimal replication region (37, 38, this report). Futhermore, this similarity of structural organization was a reflection of functional similarity as the construction of the chimeric replicons reported in this communication (chimeras 1, 2 and 3) proved the interchangeability of equivalent DNA segments between the replication of pIP501 and pSM19035. Based on our experiments with chimeras 4 and 5 this conclusion could be extended to the conjugative plasmid pAMβ1 originally isolated from *S. faecalis*.

Comparative analysis of the predicted amino acid sequence of RepR and RepS revealed a divergence within the C-terminal region. Since, DNA fragments encoding those C-terminal parts are exchangeable these residues can not be critical for the recognition of the respective replication origins.

From the construction of chimeras 4 and 5 it is possible to predict that RepR is homologous to the putative replication protein of plasmid pAMβ1. Indeed, the deduced amino acid sequence of the putative initiation replication protein of pAMβ1 (40) is highly homologous to RepR. Hence, those plasmids which belong to the same incompatibility group (*inc*18), all encode similar trans-acting replication proteins which mutually recognize the replication origin of other *inc*18 plasmids. The plasmids from this family have been reported to have a broad host range among the Gram positive bacteria (41).

ACKNOWLEDGEMENT

We are grateful to F.W. Studier for providing us with plasmids and bacterial strains, to A. Sorokin, V. Khazak, P. Ceglowski and N. Minton for providing us with the nucleotide sequence of the minimal replication region of plasmids pSM19035 and pAMβ1 prior to publication. We are grateful to A.C. Stiege and G. Lüder for technical assistance.

REFERENCES

1. Projan,S.J and Novick,R. (1988). Plasmid **19**:203–221.
2. Alonso,J.C. (1989) Microbiologia (SEM) 5,5–12.
3. Novick,R.P. (1989) Annu. Rev. Microbiol. **43**,537–565.
4. Gruss,A. and Erhlich,S.D. (1989) Microbiol. Rev. **53**,231–241.
5. Bron,S., and Luxen,E. (1985) Plasmid **14**,235–244.
6. Rabinovich,P.M., Haykinson,M.Y., Arutyunova,Y.V. and Stepanov, A.I. (1985) In Plasmid in Bacteria. Helinski,D.R., Cohen,S. N., Clewell,D.B., Jackson,D.A. and Hollander,A. eds. Plenum Press. New York and London. pp. 635–656.
7. Horodniceanu,T., Bouanchaud,D.H., Bieth,G., and Chabbert,Y.A. (1976) Antimicrob. Agents Chemother. **10**,795–801.
8. Behnke,D., Gilmore,M.S. and Ferretti,J.J. (1981) Mol. Gen. Genet. **182**,414–421.
9. Clewell,D.B. (1981) Microbiol. Rev. **45**,409–436.
10. Behnke,D., and Gilmore,M.S. (1981) Mol. Gen. Genet. **184**,115–120.
11. Alonso,J.C., Viret,J.-F. and Tailor,R.H. (1987) Mol. Gen. Genet. **208**,349–352.
12. Studier,F.W. and Moffatt,B.A. (1986) J. Mol. Biol. **189**,113–130.
13. Yanisch-Perron,C., Vieira,J., and Messing,J. (1985) Gene **33**,103–119.
14. Alonso,J.C., Lüder,G., and Trautner,T.A. (1986) EMBO J. 5,3723–3728.
15. Vieira,J., and Messing,J. (1987) Methods Enzymol. **153**,3–11.
16. Moffatt,B.A., and Studier,F.W. (1987) Cell **49**,221–227.
17. Iordanescu,S. (1975) J. bacteriol. **124**,597–601.
18. Iordanescu,S., and Surdeanu,M. (1980) Plasmid **4**,1–7
19. Behnke,D., and Gerlach,D. (1987) Mol. Gen. Genet. **210**,528–534.
20. Behnke,D., and Klaus,S. (1983) Zeit. Allgem. Mikrobiol. **23**,539–547.
21. Simon,D., and Chopin,A. (1988) Biochemie **70**,559–566.
22. Rottländer,E., and Trautner, T.A. (1970). Mol. Gen. Genet. **108**,47–60.
23. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982). Molecular Cloning. A laboratory Manual. Cold Spring Harbor Laboratory Press, New York.
24. Deichelbohrer,I., Alonso,J.C., Lüder,G., and Trautner,T.A. (1985) J. Bacteriol. **162**,1238–1243.

25. Alonso,J.C., and Trautner,T.A. (1985) Mol. Gen. Genet. **198**,427–431.
26. Birnboim,H.C., and Doly,J. (1979) Nucl. Acids Res 7,1513–1523.
27. Maciag,I.E., Viret,J.-F. and Alonso,J.C. (1988) Mol. Gen. Genet. **212**,232–240.
28. Henikoff,S. (1984) Gene **28**,351–358.
29. Sanger F., Nicklen S., and Coulson A.R. (1977). Proc. Natl. Acad. Sci. U.S.A. **74**,5463–5467.
30. Alonso,J.C. and Tailor,R.H. (1987) Mol. Gen. Genet. **210**,476–484.
31. Laemmli,U.K. (1970) Nature **227**,680–685.
32. Brendel,V., and Trifanov,E.N. (1984) Nucl Acids Res **12**,4411–4427.
33. Devereux,J., Haeberli,P., and Smithies,O. (1984) Nucleic Acids Res. **12**,387–395.
34. Tinoco,I., Borer,P.N., DengleR,B., Levine,M.D., Uhlenbeck,O.C. Crothers,D.M., and Gralla,J. (1973). Nature New Biol. **246**,40–41.
35. Pabo,C.O., and Sauer,R.T. (1984). Annu. Rev. Biochem. **53**,293–321.
36. Brennan,R.G., and Matthews,B. (1989) J. Biol. Chem. **264**,1903–1906.
37. Sorokin, A., and Khasak, V. (1989) In Genetic transformation and Expression. Eds. Butler,L.O.; Harwood,C. and Moseley,B.E.B. Intercept Ltd. Andover, U.K. pp. 269–281.
38. Brantl,S., Nowak,A., Behnke,D., and Alonso,J.C. (1989) Nucl. Acids Res. **17**,1011.
39. Shields, D.C. and Sharp, P.M. (1987) Nucl. Acids Res. **15**,8023–8040.
40. Swinfield, T.J., Oultram, J.D., Thompson, D.E., Brehm, J.K. and Minton, N.P. (1990) Gene **87**,79–90.
41. Krah, R.E. and Macrina, F.L. (1989) J. Bacteriol. **171**,6005–6012.