university of groningen

University Medical Center Groningen

University of Groningen

Replication and maintenance of plasmids in Bacillus subtilis

Meijer, Wilhelmus Johannes Jozef

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1995

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Meijer, W. J. J. (1995). *Replication and maintenance of plasmids in Bacillus subtilis*. s.n.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 22-09-2023

Chapter VII

Complete nucleotide sequences of and biological functions specified by three Bacillus subtilis (natto) rolling-circle plasmids: pTA1015, pTA1040 and pTA1060

Wilfried J.J. Meijer, G. Bea. A. Wisman, Gerard Venema and Sierd Bron

Manuscript in preparation

SUMMARY

The complete nucleotide sequences of the plasmids pTA1015, pTA1040, and pTA1060 from the Bacillus subtilis (natto) strains IAM1028, IAM1232, and IFO3022, respectively, were determined. These plasmids, which replicate according to the rolling-circle mechanism (RCM), have a structural organization that is similar to most other known RCM plasmids from Gram-positive bacteria. The three pTA-plasmids have highly similar replication functions, both for leading and lagging strand synthesis, and these replicons seem to have a common ancester. Based on homologies of the replication initiation (rep) genes and the putative double-strand origins with corresponding regions in other RCM plasmids all three pTA-plasmids belong to the pC194 group of plasmids. Besides modules that are commonly present on RCM plasmids, such as those inolved in replication, additional modules have been identified on the three pTA-plasmids analyzed. pTA1015 and pTA1060 contain a gene, mob, required for conjugative mobilization. pTA1040 and pTA1060 contain a gene showing homology to the B.subtilis chromosomally-encoded rapA and rapB genes, the products of which delay the onset of sporulation. Previously, we described a module, present on pTA1015 and pTA1040, consisting of two genes, orf1 and sipP, which encodes a putative export protein and a functional type I signal peptidase (Meijer et al., 1995b). Also genes have been identified on these plasmids of which the biological function has not yet been unraveled. The presence of additional modules on the pTA-plasmids may reflect an evolutionary adaptation to the special industrial conditions to which the host cells have been exposed.

INTRODUCTION

As in other bacteria, cloning vectors are important tools for fundamental and applied research in B.subtilis. Frequently, such vectors are based on autonomously replicating plasmids. Although the highly transformable B.subtilis strain 168, which has become the standard strain for most studies, does not harbour endogenous plasmids, the existence of B.subtilis plasmids has been known for quite some time (Tanaka and Koshikawa, 1977; Tanaka et al., 1977; Uozumi et al., 1980; Hara et al., 1983; Yoshimura et al., 1983). Since these plasmids do not confer easily selectable phenotypes to their hosts, most cloning vectors used today for B.subtilis are based on small multicopy plasmids from other Gram-positive bacteria, like staphylococci and streptococci (Ehrlich, 1977; Dubnau, 1983; Lacks et al., 1986; Novick, 1989; Gruss and Ehrlich, 1989; Bron, 1990; Jannière et al., 1993). The latter plasmids can replicate in B.subtilis and express their antibiotic resistance genes in this organism.

So far, the small cryptic plasmids from B.subtilis have attained relatively little attention. Based on size and restriction patterns, Uozumi et al. (1980) have classified most of the cryptic B.subtilis plasmids into six different groups, of which pTA1015, pTA1020, pTA1030, pTA1040, pTA1050, and pTA1060 are representatives. A seventh

group, represented by pFTB14, was reported later (Yoshimura et al., 1983). The few detailed studies on these plasmids, which were mainly focussed on their replication functions, showed that they replicate, like most small plasmids from Gram-positive bacteria, according to the rolling-circle mechanism (RCM) (Bron et al., 1987; Chang et al., 1987; Murai et al., 1987; Devine et al., 1989; Seery and Devine, 1993; Meijer et al., 1995d).

Cryptic plasmids from B.subtilis have raised our interest for the following reasons. First, vectors based on the replication functions of the B.subtilis plasmid pTA1060 are superior to those based on small non-native plasmids with respect to efficiency in cloning, mean insert size, and stability (Bron et al., 1987; Bron, 1990; Haima et al., 1990) (our unpublished results). Second, industrially important traits may be present on several cryptic B.subtilis plasmids. This speculation is based on the fact that many of the endogenous plasmids of B.subtilis known today have been identified in industrial strains. The idea that genes involved in industrial processes may be plasmid-located is not new. Several valuable traits for dairy industry have been identified on plasmids present in lactic acid bacteria (McKay, 1983). Circumstantial support for the presence of additional genes on cryptic B.subtilis plasmids, which are not normally found on other RCM plasmids, is that the former are several kb larger than most other RCM plasmids from Gram-positive bacteria. In addition to these motivations of applied nature, we have a fundamental interest in the organization and evolution of and biological functions specified by cryptic B.subtilis plasmids. Together, these motivations prompted us to characterize a number of cryptic B.subtilis plasmids at the sequence level: pTA1015, pTA1040 and pTA1060 were chosen for this purpose. Each of these plasmids represents a different group in the classification of Uozumi et al. (1980).

Previous studies on parts of these plasmids had already been performed in our group (Bron et al., 1987; Haima et al., 1990; Bron et al., 1991; Meijer et al., 1995d; Meijer et al., 1995b). In the earlier work we analyzed the single-stranded DNA (ssDNA) conversion signals (designated SSO) present on these plasmids (Bron et al., 1987; Meijer et al., 1995d), and we showed that pTA1015 and pTA1040 contain a gene encoding a functional type I signal peptidase (Meijer et al., 1995b).

In the present work the entire sequences of pTA1015, pTA1040 and pTA1060 were determined and computer assisted analysis on the DNA sequences was used to identify putative genes. Several of these genes were analyzed at the functional level. The analyses showed that a high level of homology existed between the genes encoding the essential replication initiation protein (Rep) and the DNA sequences specifying the putative initiation sites for leading-strand synthesis (called double-strand-origins [DSO]). In addition to the rep genes, other (putative) genes were identified. Whereas some of these are present on at least two of the three pTA-plasmids examined, others seem to be unique.

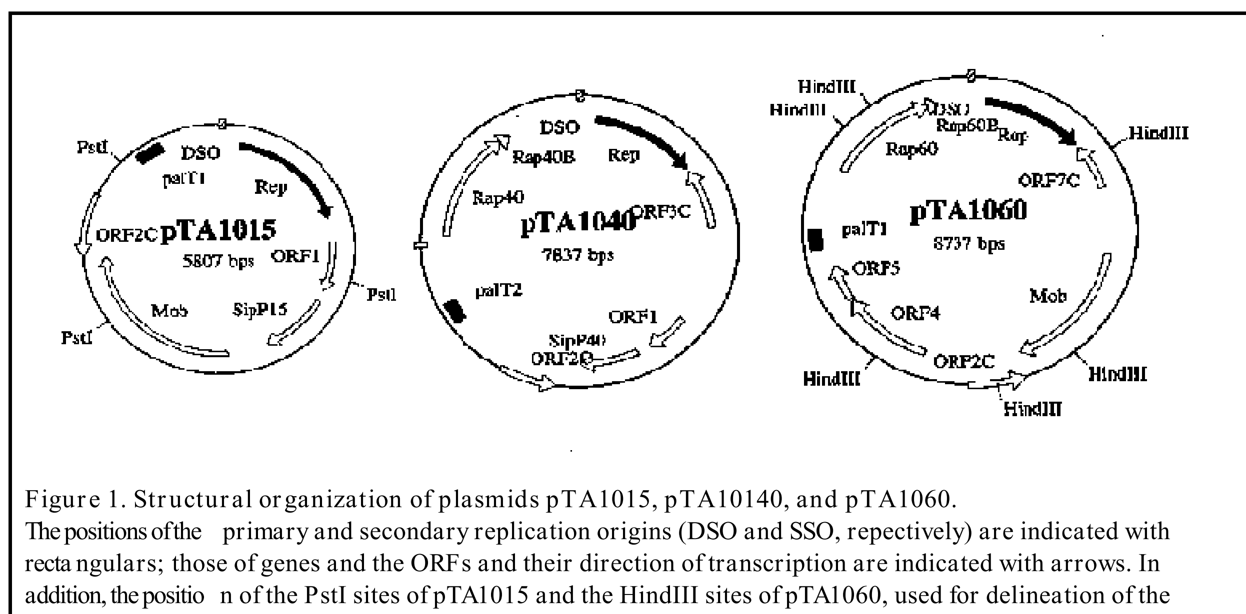


Figure 1. Structural organization of plasmids pTA1015, pTA10140, and pTA1060.
The positions of the primary and secondary replication origins (DSO and SSO, repectively) are indicated with rectangulars; those of genes and the ORFs and their direction of transcription are indicated with arrows. In addition, the position of the PstI sites of pTA1015 and the HindIII sites of pTA1060, used for delineation of the

RESULTS

Structural organization of plasmids pTA1015, pTA1040, and pTA1060

The complete DNA sequence of the plasmids pTA1015, pTA1040, and pTA1060 was determined on both strands using overlapping inserts cloned in pUC18. The size of pTA1015 was shown to be 5807 bp, that of pTA1040 7837 bp, and that of pTA1060 8737 bp. The G+C contents of the plasmids were 40.7% (pTA1015), 36.9% (pTA1040), and 38.1% (pTA1060). The complete nucleotide sequences have been deposited in the EMBL/GenBank/DDBJ nucleotide sequence database and were assigned the accession numbers U32379 (pTA1015), U32378 (pTA1040) and U32380 (pTA1060). The putative nick sites for replication initiation were given coordinate number 3 in each plasmid. The DNA sequences of the three pTA-plasmids with their identified features are presented as an addendum at the end of this thesis.

Generally, RCM plasmids consist of several interchangeable modules, which frequently show considerable homology at the DNA and/or deduced protein level. An essential module comprizes a gene, denoted rep, which encodes the replication initiation protein (Rep), and its cognate DSO. Another module contains a non-transcribed region characterized by a strong dyad symmetry. This region forms the single-strand origin (SSO) and functions as the major initiation site for lagging-strand synthesis. Many RCM plasmids contain, in addition, a module containing a gene, denoted mob or pre, which is involved in conjugative mobilization and site-specific recombination, respectively. Besides these common modules, other plasmid-specific modules can be present.

To examine whether the three pTA-plasmids have a structural organization similar to that of other known RCM plasmids, their DNA sequences and deduced amino acid sequences of identified open reading frames (ORFs) were compared to known sequences in available databases. These comparisons revealed the following putative genes and non-coding functional DNA sequences which are commonly found on RCM plasmids: (i) a rep gene and its cognate DSO, which is located upstream of the rep gene (pTA1015, pTA1040 and

Table 1. Characteristics of the ORFs on pTA-plasmids

ORF/gene	Location [a]	Size (codons)	Mol.mass (kDa) [b]	putative start-codon and RBS [c]	Spacing [d] (bp)
pTA1015					
rep15	166 - 1182	339	39.5	cag aaggag tttttttgttcATG	12
ORF1.15	1401 - 1838	146	15.9	tgct aggagg gaaagtttttATG	12
sipP (pTA1015)	1945 - 2502	186	21.2	atatag aggagg aaatttctTTG	10
mob15	2840 - 4245	482	56.6	ctgaat ggggg gtttttctcATG	11
ORF2C15	4778 - 4302 (C)[e]	159	18.8	tttttac gaggtgat acgttATG	11
pTA1040					
rep40	165 - 1181	339	39.7	gtcag aagggg tttttcactTTG	11
ORF3C40	1829 - 1227 (C)	201	22.2	ttaaggagg atttgaacaatATG	13
ORF1.40	2826 - 3251	142	15.8	gaaaggatgg aagaagaactATG	16
sipP (pTA1040)	3372 - 3926	185	21.5	ccaagcg ggagg aagcgtaaGTG	10
ORF2C40	4587 - 4117 (C)	157	18.3	cat aaagaggtga acccgctATG	12
rap40	5943 - 7067	375	44.3	tgtcg aaggag agagatgtgATG	10
rap40B	7060 - 7176	39	4.1	tc ggagg ggcgagttcttgtATG	15
pTA1060					
rep60	164 - 1183	340	39.5	gctcag aaggag tttttttgTTG	9
ORF7C60	1759 - 1220 (C)	180	20.1	aaagaagg gatgtttttttaaTTG	14
mob60	2438 - 3883	482	56.8	ctgaatc gggggt ttttgtcATG	10
ORF2C60	4387 - 3911 (C)	159	18.8	tttttac gaggtgat acgtgATG	11
ORF4.60	4825 - 5739	305	35.6	catacat aaggag atgtttaATG	8
ORF5.60	5763 - 6158	132	15.1	aagggg atggggg agttattATG	9 or 15
rap60	7199 - 8323	375	44.3	tagg ggagg agttactcggaATG	13
rap60B	8316 - 8426	37	3.9	ttc aaagggg cgatttccgtATG	12

a: The numbers correspond to the sequenceses deposited in the EMBL/Genbank Data Library under the accession numbers U32379 (pTA1015), U32378 (pTA1040) and U32380 (pTA1060). The origin nick sites were taken as positions 3.

Figure 2. Comparison of the replication regions of Bacillus RCM plasmids. ➡
Sequence information was taken from the following sources: pLS11, Hara et al. (1992b); pBS2, Darabi et al. (1989); pUH1, Hara et al. (1991); pBAA1, Devine et al. (1989), and pFTB14, Murai et al. (1987). (A) Alignment of the deduced amino acid sequences of the rep genes from pTA1015 (REP15), pTA1060 (REP60), pTA1040 (REP40), pLS11 (REPLS11), pBS2 (REPBS2), pUH1 (REPUH1), pBAA1 (REPBAA1), and pFTB14 (REPFTB14). Conserved amino acid residues are boxed. The three amino acid residues shown to have a catalytic role in replication initiation of pC194 (Gros et al., 1994) are indicated with arrows. (B) Alignment of putative promoters of the rep genes. The consensus B.subtilis F^A promoter sequence is shown in the top line of this figure. Matches with the consensus sequence are indicated with asterisks. In addition, the number of nucleotides separating the putative promoter from the startcodon (spacing) and the startcodons themselves are shown (these features for pLS11, pBS2, pUH1, pBAA1, and pFTB14 are according to their published data). (C) Alignment of DSO regions. Conserved nucleotides are boxed. Inverted repeated sequences are indicated with arrows. The putative origin nick-site is indicated with an arrow. The boxed region of 8 nucleotides containing the nick-site is conserved in almost all pC194-type plasmids.

pTA1060); (ii) a mob gene (pTA1015 and pTA1060); and (iii) an SSO (pTA1015, pTA1040 and pTA1060). Previous work showed that the SSOs of these plasmids belong to the palT1 (pTA1015 and pTA1060) or palT2 (pTA1040) type of lagging strand initiation sites, both of which are highly efficient in B.subtilis (Meijer et al., 1995d). The positions of the SSOs relative to their cognate rep genes have now been established (see Fig. 1). In addition to these characteristic RCM modules, other putative genes were identified. Some of these showed significant homology to previously identified genes and these plasmid-located genes were given names using the nomenclature of their homologous equivalents. The directions of transcription and the localization of the different genes, as well as the positions of the DSOs and the SSOs, are shown in Fig. 1. Characteristics of the identified genes are shown in Table 1. The data show that the three pTA-plasmids have a structural organization similar to that of other RCM plasmids from Gram-positive bacteria. In the following sections comparisons and experimental analyses of the different genes and modules will be discussed.

The replication module

All RCM plasmids studied so far contain a gene, rep, encoding a protein that is essential for the initiation of replication. Through their DNA-binding and nicking/closing activity the Rep proteins introduce a strand- and site-specific nick in the plasmid at the DSO. The 3'-OH end of the nick-site is subsequently used for the initiation of leading strand synthesis. The rep genes of pTA1015, pTA1040, and pTA1060 were identified by their sequence homology with the rep genes from other RCM plasmids. The requirement of this gene for replication was demonstrated for pTA1015 and pTA1060 (see below).

Based on sequence similarity of their DSOs and homology of the replication proteins, the RCM plasmids from Gram-positive bacteria can be grouped into five classes, which are represented by pT181, pC194, pE194, pSN2 and pTX14-3. The replication regions of the cryptic B.subtilis plasmids pBAA1 (Devine et al., 1989), pBS2 (Darabi et al., 1989), pLS11 (Hara et al., 1992b), and pUH1 (Hara et al., 1991), and that of the Bacillus amyloliquefaciens plasmid pFTB14 (Murai et al., 1987) have been sequenced. All these Bacillus plasmids belong to the pC194 class of RCM plasmids

(A)

```
REP15     M-YSSENDYIILEDKTATGKKRDWKGKKRRTNLMAEHYEALQSKTGIPYYGKKAEKLQSCAEQLSFKRDPETGKLKLYQAQFCKV  84
REP60     MIYSSENDYSILEDKTATGKKRDWKGKKRRTNLMAEHYEALESKIGAPYYGKKAEKLISCAEYLSFKRDPETGKLKLYQAHFCKV  85
REP40     M-YSSESNYSIFQDKTATGKKRDWKGKKVRSLLMAMHYEGLEKRTGAPYYGKKAEKVCDCAEQLAFKRDLETGRLRLYQAYFCKV  84
REPLS11        MIISILKIKPQQVKSGIGRGKKDGRTSWRTVEALESKIGAPYYGKKAEKLISCAEYLSFKRDRRPGKLKLYQAHFCKV  78
REPBS2                                   MAEHYEALESKIGAPYYGKKAEKLISCAEYLSFKRDPETGKLKLYQAHFCKV  52
REPUH1         MIISSLRTRPQQVKSGIGRGKKDGKILWLSTMKLYRYKLVYLTHAKKLPNCAVVRKVFRLNSTRRRKLKLYQAQFCRV  78
REPBAA1                                  MAEHYEALESKIGAPYYGKKAEKLISCAEYLSFKRDPETGKLKLYQAHFCKV  52
REPFTB14  M-YSSENDYSILEDKTATGKKRDWRGKKRRANLMAEHYEALERRIGAPYYGKKAERLSECAEELSFKRDPETGRLKLYQAHFCKV  84

REP15     RLCPMCAWRRSLKIAYHNKLIVEEANRQYGCGWIFLTLTVRNVEGDSLKPMIADHMKGWNRLFQYKRVKVADLGYFRALEITKNH  169
REP60     RLCPMCAWRRSLKIAYHNKLIVEEANRQYGCGWIFLTLTIRNVKGERLKPQISAMMEGFRKLFQYKKVKTSVLGFFRALEITKNH  170
REP40     RLCPMCAWRRSLKIAYHNKLIVEEANRQYKPAWIFLTLTVKNVQGDRLKQTISDMMKGFMKLMKYKKVKDSAVRGYFRALEITKNH  169
REPLS11   RLCPMCRWRRSLKIAYHNKLIVEEANRQYCCGWIFLTLTIRNVKGERLKPQISEMMEGFRKLFQYKKVKTSVLGFFRALGIAKNH  163
REPBS2    RLCPMCAWRRSLKIAYHNKLIVEEANRQYGCGWIFLTLTVRNVKGERLKPQISEMMEGFRKLFQYKKVKTSVLGFFRALEITKNH  137
REPUH1    RLCPMCAWRRSLKIAYHNKLIVEEANRQYGCGWIFLTLTVRNVEGDGLKPMIADHMKGHNRLFAYKRVKVADLGYFRALEITKNH  163
REPBAA1   RLCPMCAWRRSLKIAYHNKLIVEEANRQYGCGWIFLTLTVRNVKGERLKPQISEMMEGFRKLFQYKKVKTSVLGFFRALEITKNH  137
REPFTB14  RLCPMCAWRRSLKIAYHNKLIIEEANRQYGCGWIFLTLTVRNVKGERLKPQISEMMEGFRKLFQYKKVKTSVLGFFRALEITKNE  169

REP15     EEDTYHPHFHVLLPVKKSYFTHNYIKQSEWTSLWKRAMKLDYTPIVDIRRVKGRAKIDAEQIASDVREAMMEQKAVLEISKYPVK  254
REP60     EEDTYHPHFHVLIPVRKNYFGKNYIKQAEWTSLWKKAMKLDYTPIVDIRRVKGKAKIDAEQIANDVRKAMMEQKAVLEISKYPVK  255
REP40     EEQTYHPHFHVLIPVRTGYFGKNYIKQKEWTSFWKRAMKMDYTPIVHVQRVKGKKGIDAEAIEQEVRKEMREQKAILEISKYPVK  254
REPLS11   EEDTYHPHFHVLIPVRKNYFGKNYIKQAEWTSLWKKAMKLDYTPIVDIRRVKGEAVLDAEQIENDVRKAMMEQKAVLEISKYPVK  248
REPBS2    EEDTYHPHFHVLIPVRKNYFGKNYIKQAEWTSLWKKAMKLDYTPIVDIRRVKGEAKIDAEQIESDVREAMMEQKAVLEISKYPVK  222
REPUH1    EEDTYHPHFHVLLPVKKSYFTHNYIKQSEWTSLWKRAMKLDYTPIVDIRRVKGKAKIDAELIENDVREAMMEQKAVLEISKYPVK  248
REPBAA1   EEDTYHPHFHVLIPVRKNYFGKNYIKQAEWTSLWKKAMKLDYTPIVDIRRVKGEAKIDAEQIESDVREAMMEQKAVLEISKYPVK  222
REPFTB14  EEDTYHPHFHLLLPVKKNYFGKNYIKQAEWTSLWKKAMKLDYTPIVDIRRVKGRVKIDAEQIESDVREAMKEQKAVLEISKYPVK  254

REP15     DTDVVRGNKVTSDNLNTVFYLDDALSARRLIGYGGILKEIHKELNLGDAEDGDLVKIEEEDDEVANEAFEVMAYWHPGIKNYIIR  339
REP60     DTDVVRGNKVTSDNLNTVLYLDDALAARRLIGYGGILKEIHKELNLGDAEDGDLVKIEEEDDEVANGAFEVMAYWHPGIKNYIIK  340
REP40     DTDVIRGNEVTEENLQTVYYLDDALSARRLIGYGGILKEIHKELNLTDAEDGDLIRIEEDDDVANETFEVMAHWHVGIKNYIIT  339
REPLS11   DTDVVRGNKVTEDNLNTVLYLDDALAARRLMGYGGMLKEIHKELNLGDAEDGDLVKIEEEDEVQNGAFEVMAYWHPGIKNYIIR  333
REPBS2    DTDVVRGNKVTEDNLNTVLYLDDALAARRLIGYGGILKEIHKELNL                                         268
REPUH1    DTDVVRGNKVTEDNLNTVFYLDDALSRRRLIGYGGILKEIHKELNLGDAEDGDLVKIEEEDDEVANEAFEVMAYWHPGIKNYIIR  333
REPBAA1   DTDVVRGNKVTEDNLNTVLYLDDALAARRLIGYGGILKEIHKELNLGDAEDGDLVKIEEEDEVANGAFEVMAYWHPGIKNYIIK  307
REPFTB14  DTDVVRGSKVTDDNLNTVFYLDDALSARRLIGYGGILKEIHKELNLGDAEGGDLVKIEEEDEVANGAFEVMAYWHPGIKNYILK  335
```

(B)

```
                                                          spacing   startcodon
consensus  TTGACA ------------------- TATAAT

pTA1015    TTGACA TGTAGAAACTCCTTCCGC TATTAT - - -  61 - - ATG
pTA1060    TTGACA GGTAAAAACTCCTTCTGC TATTAT - - -  59 - - TTG
pTA1040    TTGACA GTGAAAAACCCCATCTGC TATTAT - - -  61 - - TTG
pLS11      TTGACA GGTAAAAACTCCTTCTGC TATTAT - - -  76 - - ATG
pBS2       TTGACA GGTAAAAACTCCTTCTGC TATTAT - - -158 - - ATG
pUH1       TTGACA TGTAGAAACTCCTTCCGC TATTAT - - -  78 - - ATG
pBAA1      TTGACA GGTAAAAACTCCTTCTGC TATTAT - - -158 - - ATG
pFTB14     TTGACA GGTAAAAACTCCTTCTGC TATTAT - - -  65 - - ATG
           ******     * **** ** ** ** ******
```

(C)

pTA1015
pTA1040
pTA1060
pBAA1
pBS2
pFTB14
pLS11
pUH1

(reviewed by Seery et al., 1993). An alignment of the deduced protein sequences of the three pTA-plasmids with those of the other Bacillus plasmids mentioned is shown in Fig. 2A. The alignment shows that also the three pTA-plasmids belong to the pC194 group. Apart from differences in the N-terminus, these Rep proteins are highly homologous to each other. The percentages of identity of the various proteins when compared pairwise are shown in Table 2. Mutation analysis of RepA, the replication protein of pC194, has shown that at least three of its amino acid residues have a catalytic role (Gros et al., 1994). The corresponding amino acids, one tyrosine and two glutamate residues (indicated in Fig. 2A) are, except for one glutamate in pLS11, conserved in all plasmids included in this comparison.

Murai et al. (1987) have delineated the promoter region of the pFTB14 rep gene to about 95 bp. Within this region, sequences were identified resembling the -35 and -10 consensus sequences of B.subtilis σ^A promoters. Therefore, these sequences are likely to constitute the promoter of the rep gene of pFTB14. Highly homologous sequences are present upstream of the rep genes of the three pTA-plasmids, as well as in the other listed plasmids (Fig. 2B), suggesting that the expression of all these Rep proteins is driven by highly homologous promoters.

Characteristically, DSOs are present in a plasmid region that has the potential to form secondary structures. The sequence 5'-TCTTGATA-3' is found at the origin nick-site of most members of the pC194 group of plasmids (Gruss and Ehrlich, 1989; Seery et al., 1993). As shown in Fig. 2C, this consensus nick-site sequence is also conserved in the three pTA-plasmids and, as in other members of the pC194 class of plasmids, the DSO is located within a region of dyad symmetry upstream of its cognate rep gene.

The plasmid region containing the rep gene and its cognate DSO are sufficient to drive replication of plasmids pUH1 (Hara et al., 1991), pFTB14 (Murai et al., 1987), pLS11 (Hara et al., 1992b), and pBAA1 (Devine et al., 1989), in B.subtilis. The replication functions of pTA1060 are

Table 2. Homologies between Rep proteins from Bacillus plasmids

	pTA1015	pTA1040	pTA1060	pLS11	pBS2	pUH1	pBAA1	pFTB14
pTA1015	-	77	89	76	87	79	82	87
pTA1040	77	-	77	67	78	62	72	76
pTA1060	89	77	-	88	99	71	92	92
pLS11	76	67	88	-	93	74	87	83
pBS2	87	78	99	93	-	74	100	93
pUH1	81	62	71	74	74	-	70	72
pBAA1	82	73	92	87	100	71	-	86
pFTB14	87	76	92	83	93	71	86	-

Identities are shown according to the sequence data obtained in these studies and the protein sequences of published data

present on the 2.2 kb HindIII fragment (Haima et al., 1987), which contains the rep gene and its DSO (Fig. 1). Also the 2.4 kb PstI region of pTA1015, containing its rep gene and DSO (Fig. 1), are sufficient to drive autonomous replication in B.subtilis (our unpublished results).

The SSO of replication

The RCM of replication is characterized by the production of single-stranded (ss) DNA replication intermediates. Efficient conversion of ssDNA into duplex plasmid DNA is initiated from specific, non-coding plasmid regions, the SSOs, which have a high potential to form secondary structures. Previously, we have reported the cloning and sequencing of the SSOs of pTA1015, pTA1040, and pTA1060 (Meijer et al., 1995d). The results showed that, whereas the SSOs of pTA1015 and pTA1060, designated palT1, are almost identical to each other, the SSO of pTA1040, designated palT2, is less homologous (77%). Furthermore, the results led to the idea that probably all known B.subtilis RCM plasmids contain either a palT1- or a palT2-type of SSO. Moreover, we showed that both types of SSO are highly efficient ssDNA conversion signals in B.subtilis (Meijer et al., 1995d). Whereas in pTA1015, pTA1020, and pUH1, the SSO and the DSO sequences flank each other, these primary and secondary replication origins are separated in pTA1040 and pTA1060 by several kb (Fig. 1). We showed that, in its natural setting, the conversion of ssDNA occurred with high efficiency in pTA1015, pTA1040 and pTA1060 (Meijer et al., 1995d). This corroborates the general view that the position of SSOs relative to their DSOs usually does not affect their functionality.

pTA1015 and pTA1060 contain functional mob genes but no **γ**-gtp genes

γ-Glutamyl transpeptidase (γ-GTP) is an enzyme which catalyzes the hydrolysis of glutathione to glutamic acid and the transfer of the γ-glutamyl group of glutathione to an amino acid or peptide (Ryoichi et al., 1991). The enzyme isuseful as a dough conditioner for increasing the specific volume and for suppression of retrogradation of bread. γ-GTP is widely present in various organisms, from bacteria to higher animals (Meister and Tate, 1976). Hara et al. (1983; 1993) reported that some B.subtilis strains produce active γ-GTP. The γ-gtp gene from the B.subtilis chromosome of strain SJ138 has been cloned and sequenced (Ryoichi et al., 1991). Hara et al. (1992a) claim to have identified a γ-gtp gene on plasmid pUH1. Remarkably, the pUH1-located γ-gtp gene has no significant homology with the B.subtilis SJ138 chromosomal γ-gtp gene, nor with that of E.coli (Suzuki et al., 1989). Plasmids pTA1015 and pTA1060, but not pTA1040, contain an ORF, designated Mob in Fig. 1, encoding a putative protein with a deduced amino acid sequence that is almost identical to that of the postulated γ-GTP protein of pUH1 (Fig. 3). In addition, the deduced protein sequences show significant homology to those from mob genes. The latter genes, which are present on various RCM plasmids, are essential for conjugative mobilization (Priebe and Lacks, 1989; Selinger et al., 1990; van der Lelie et al., 1990; Oskam et al., 1991).

We analyzed the possible involvement of the corresponding genes on pTA1015 and pTA1060 in plasmid mobilization and/or γ-GTP activity. For this purpose, the pTA1015 derivatives pTAB11A, pTAB11B, pTAB13 and pTAB31 were constructed (Fig. 4, and Materials and Methods). In pTAB11A and pTAB11B, the kanamycine resistance

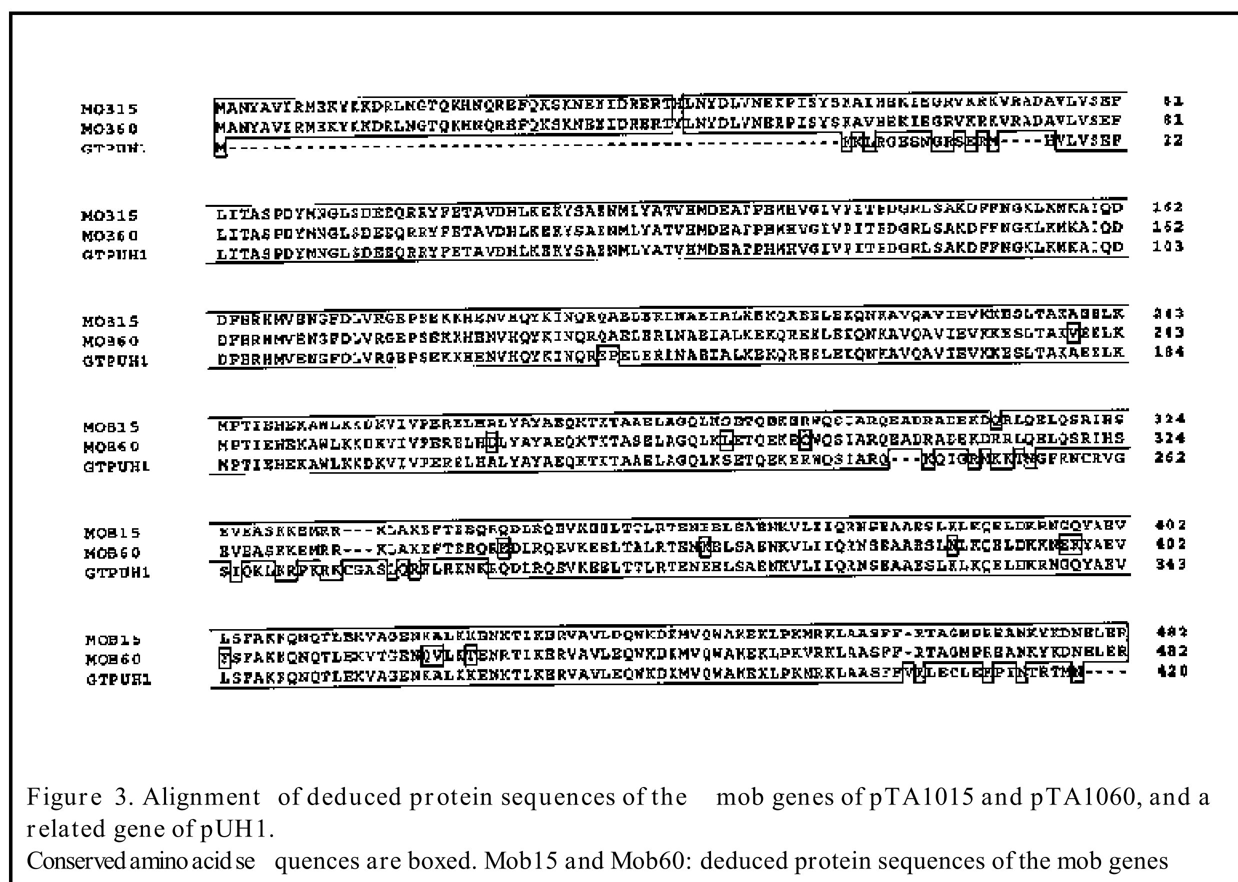


Figure 3. Alignment of deduced protein sequences of the mob genes of pTA1015 and pTA1060, and a related gene of pUH1.
Conserved amino acid sequences are boxed. Mob15 and Mob60: deduced protein sequences of the mob genes

(Km^R) gene originating from the streptococcal plasmid pJH1 was cloned into the ORF2C coding sequence. pTAB13 is a derivative of pTAB11A in which a 500 bp Eco47^{III} fragment, located within the mob coding sequences, was deleted. This deletion also resulted in a frameshift of the mob gene at the Eco47^{III} fusion site (the first 115 codons are still intact and a stopcodon is present after 170 codons). In plasmid pTAB31, the Km^R gene is present in the ORF1 coding sequence (in the same transcriptional orientation as the rep gene), leaving ORF2C intact. These derivatives of pTA1015 were introduced into B.subtilis strain BD630. Next, γ-GTP activities were measured in the culture media of growing cells. Different media were tested: TY medium, minimal medium, sporulation medium and DGYP medium. Samples were collected from cultures at different phases of growth (exponential -, transition -, stationary, late stationary [8-10 hours after transition] and very late stationary phase [20-30 hours after transition]). Culture supernatants were collected and assayed for γ-GTP activity after removing the cells by centrifugation. In addition, from each sample γ-GTP activities were measured in crude cell-lysates of the collected cells. As a positive control, commercially available γ-GTP was used. Whereas the level of γ-GTP activity observed with pure γ-GTP corresponded with the expected level of activity depending on the amounts of enzyme used, no γ-GTP activity was measured in either of the B.subtilis samples tested (results not shown). These data show that pTA1015 does not produce functional γ-GTP, or that the amounts of γ-GTP were too low to be detected, when present in B.subtilis BD630 cells.

Next, we tested whether the

pTA1015-located ORF, designated Mob, encodes a functional mobilization protein. It was shown before that the conjugative plasmid pLS20 enables coresident RCM plasmids like pUB110, pTB913, pBC16 or pLS19 to be transferred to other cells in interspecific matings (Koehler and Thorne, 1987; Oskam et al., 1991). The plasmids pTAB11A, pTAB11B, pTAB13, pTAB31 and pUB110 (positive control) were used to transform B.subtilis 3335 UM4 protoplasts carrying pLS20, and the strains obtained served as donors in matings with B.subtilis 168 strain 1012Cm. The results of the matings (Table 3) demonstrated that pTA1015 derivatives can be mobilized and that the intact ORF, indicated as Mob, is required for this function. A derivative of pTA1060, pBB2, in which the homologous ORF is intact, was also shown to be mobilizable (results not shown). Accordingly, the corresponding genes of pTA1015 and pTA1060 are designated mob15 and mob60, respectively.

The results in Table 3 also show that the frequency of mobilization was significantly higher with pTAB11A (interrupted ORF2C), as compared to pTAB31 (intact ORF2C). Several explanations for the observed difference in mobilization frequencies between pTAB11A and pTAB31 were considered. The first is that readthrough transcription activity from the promoter of ORF2C in pTAB31 interferes with the expression of the convergently transcribed mob gene. This possible transcriptional interference may be relieved in pTAB11A due to the orientation of the cloned Km^R gene. However, the observation that pTAB11A and pTAB11B, which only differ by the

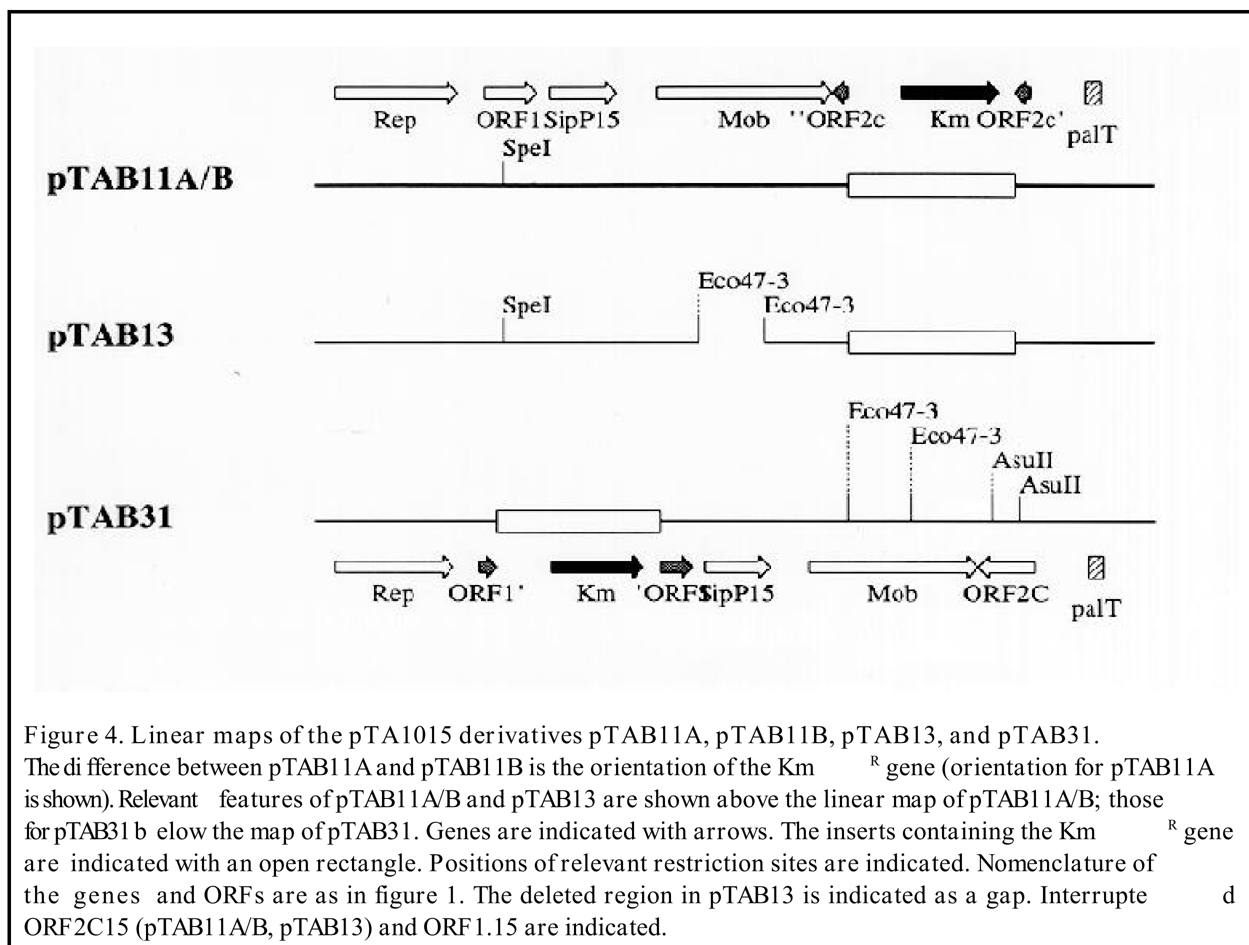


Figure 4. Linear maps of the pTA1015 derivatives pTAB11A, pTAB11B, pTAB13, and pTAB31. The difference between pTAB11A and pTAB11B is the orientation of the Km^R gene (orientation for pTAB11A is shown). Relevant features of pTAB11A/B and pTAB13 are shown above the linear map of pTAB11A/B; those for pTAB31 below the map of pTAB31. Genes are indicated with arrows. The inserts containing the Km^R gene are indicated with an open rectangle. Positions of relevant restriction sites are indicated. Nomenclature of the genes and ORFs are as in figure 1. The deleted region in pTAB13 is indicated as a gap. Interrupted ORF2C15 (pTAB11A/B, pTAB13) and ORF1.15 are indicated.

Table 3. Mobilization frequencies

Plasmid	Genotype	Frequency[a]
pTAB11A and pTAB11B	ORF2C15$^-$, mob15$^+$	8.4 * 10^{-5}
pTAB13	ORF2C15$^-$, mob15$^-$	< 1 * 10^{-9}
pTAB31	ORF1.15$^-$, mob15$^+$	9.0 * 10^{-6}
pUB110	mob$^+$	2.9 * 10^{-4}

a: frequencies are calculated as the number of transconjugants per donor cell. The frequencies are given as the mean frequencies of five independent experiments. (-) denotes the interruption of th e ORF by the Km^R gene (ORF2c, ORF1) or deletion of an internal fragment of the ORF DNA sequence (mob).

orientation of the cloned Km^R gene, show similar mobilization frequencies makes this explanation unlikely. An alternative explanation is that interruption of either the ORF2C or ORF1 affects the plasmid copy number, or the viability of the host cells. To assay for possible differences in plasmid copy numbers, equal amounts of total DNA extracts from B.subtilis cells harbouring pTAB11A, pTAB11B or pTAB31 were separated by agarose gel electrophoresis and stained with ethidium bromide. No clear differences in plasmid copy numbers between these plasmids were observed (results not shown). Also, no obvious differences in the growth rates or the number of colony-forming units per OD600 were observed between the various plasmid-containing B.subtilis cultures, indicating that the observed differences in mobilization frequencies were neither caused by differences in cellular growth rates or viability. A possible alternative explanation, which implies a direct effect of the gene product of ORF2C on mobilization frequencies, is described in the following section.

The RCM plasmids pT181 and pE194 contain a gene specifying a protein showing a high level of homology with Mob proteins. Attempts to mobilize pE194 with pLS20 (Koehler and Thorne, 1987), and pE194 or pT181 with pGO1 (Projan and Archer, 1989) have been unsuccessful. These genes were shown to be involved in site-specific, recA independent, recombination events which result in plasmid cointegrate formation (Gennaro et al., 1987) and were named pre (for: plasmid recombination enzyme). The Pre proteins act at a specific target site, called Recombination Site A (RS_A), the position of which overlaps with the -10 promoter sequence of these pre genes (Gennaro et al., 1987). The DNA sequences of the RS_A sites show dyad symmetry. Sequences homologous to the RS_A sites are also located upstream of the mob genes present on plasmids pLAB1000 (Josson et al., 1990), pMV158 (van der Lelie et al., 1989), pUB110 (van der Lelie et al., 1989), pTB913 (van der Lelie et al., 1989), and pGI2 (Mahillon and Seurinck, 1988). Selinger et al. (1990) showed that a plasmid containing a region encompassing the RS_A site of pUB110 can be mobilized if the pUB110 Mob protein is provided in trans. This indicates that these RS_A-like sequences are the target sites for the Mob protein and, therefore, these sites have been called oriT (Selinger et al., 1990).

Analysis of pTA1015 and pTA1060 revealed sequences that are highly homologous to oriT/RS_A-sites at a position 57 bp upstream of the potential startcodons of the mob genes. These sequences, which are fully conserved between the two plasmids, are, therefore, likely to constitute

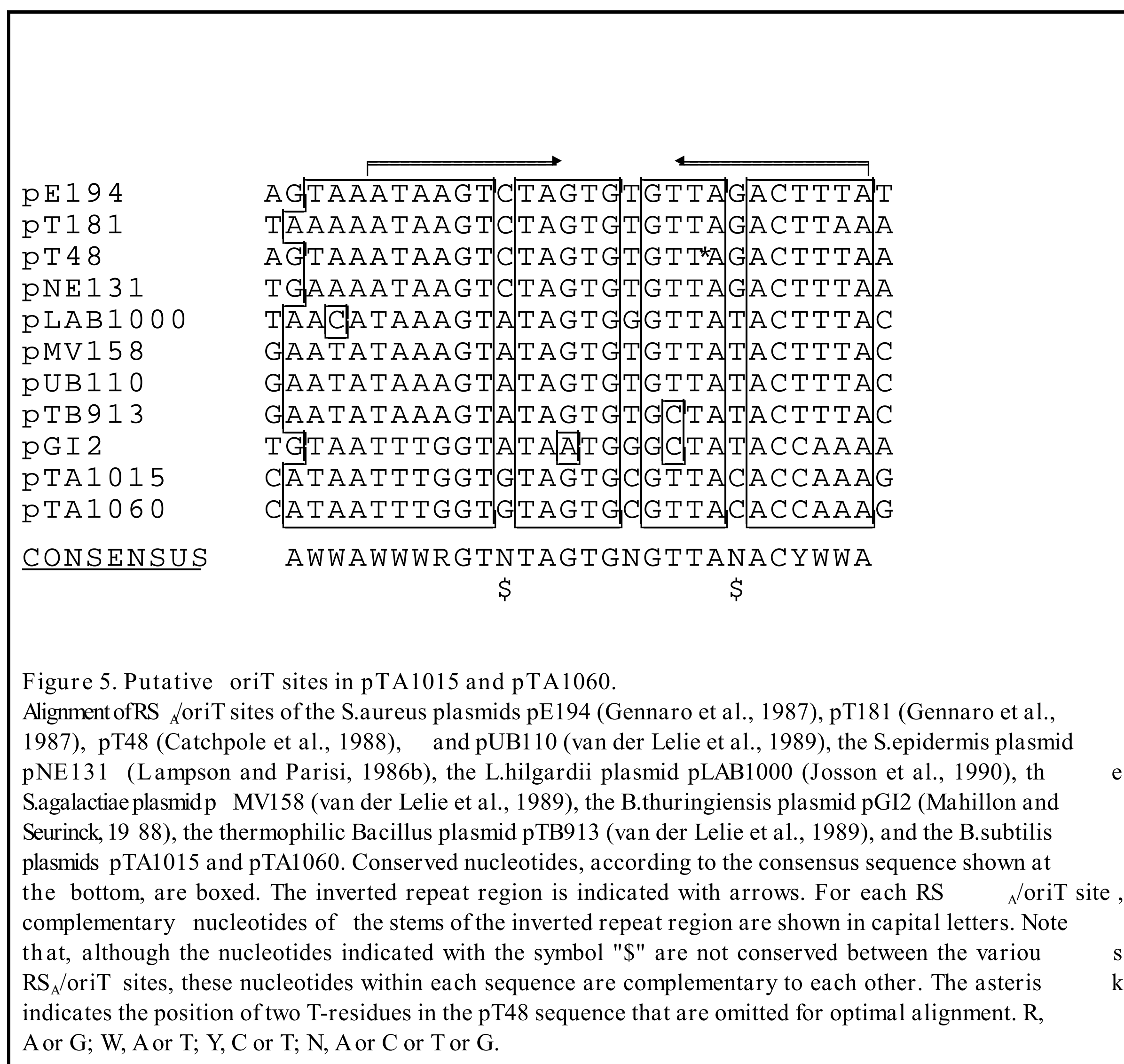


Figure 5. Putative oriT sites in pTA1015 and pTA1060.
Alignment of RS_A/oriT sites of the S.aureus plasmids pE194 (Gennaro et al., 1987), pT181 (Gennaro et al., 1987), pT48 (Catchpole et al., 1988), and pUB110 (van der Lelie et al., 1989), the S.epidermis plasmid pNE131 (Lampson and Parisi, 1986b), the L.hilgardii plasmid pLAB1000 (Josson et al., 1990), the S.agalactiae plasmid pMV158 (van der Lelie et al., 1989), the B.thuringiensis plasmid pGI2 (Mahillon and Seurinck, 1988), the thermophilic Bacillus plasmid pTB913 (van der Lelie et al., 1989), and the B.subtilis plasmids pTA1015 and pTA1060. Conserved nucleotides, according to the consensus sequence shown at the bottom, are boxed. The inverted repeat region is indicated with arrows. For each RS_A/oriT site, complementary nucleotides of the stems of the inverted repeat region are shown in capital letters. Note that, although the nucleotides indicated with the symbol "$" are not conserved between the various RS_A/oriT sites, these nucleotides within each sequence are complementary to each other. The asterisk indicates the position of two T-residues in the pT48 sequence that are omitted for optimal alignment. R, A or G; W, A or T; Y, C or T; N, A or C or T or G.

the target sites for the pTA1015- and pTA1060-encoded Mob proteins. An alignment of these RS_A/oriT sites is shown in Fig. 5. Since sequences highly homologous to RS_A/oriT sites are also present on the plasmids pT48 (Catchpole et al., 1988) and pNE131 (Lampson and Parisi, 1986a), which do not contain mob or pre genes, we screened pTA1040 (which, likewise, does not contain a mob gene) for the presence of sequences homologous to RS_A/oriT. However, no such sequences were identified on pTA1040.

The high level of homology between the deduced protein sequences of the mob15 and mob40 products is also reflected at the DNA level (95.9% identity over 1350 bp). Remarkably, the high level of homology on the DNA level is not limited to the Mob-encoding regions and the putative oriT areas but extends to approximately 150 bp further upstream of the mob genes. These homologous sequences have the potential to form secondary structures. A possible function of these dyad symmetries is unknown.

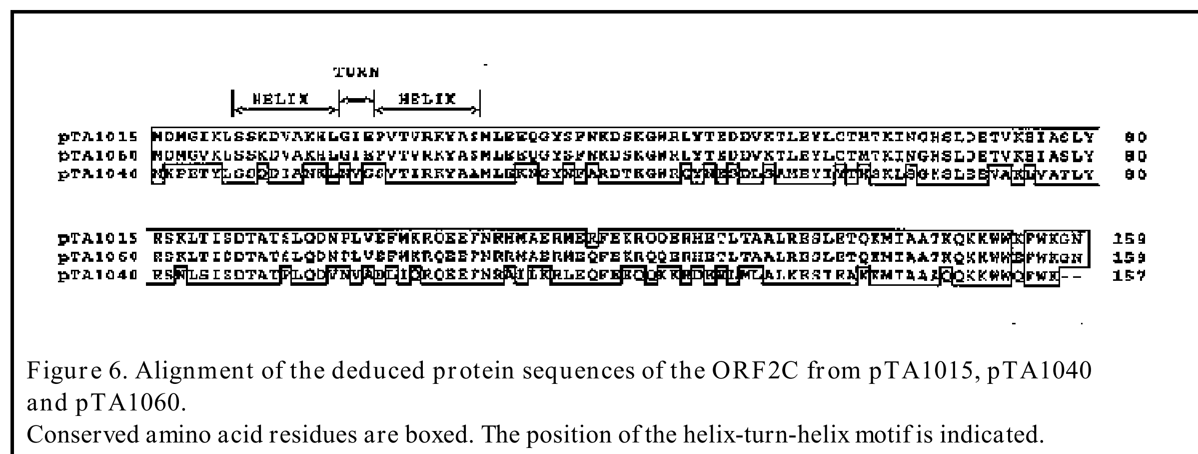


Figure 6. Alignment of the deduced protein sequences of the ORF2C from pTA1015, pTA1040 and pTA1060.
Conserved amino acid residues are boxed. The position of the helix-turn-helix motif is indicated.

The ORF2C module

All three pTA-plasmids contain a highly homologous ORF2C with an orientation of transcription opposite to that of most other ORFs/genes on these plasmids (Fig. 1). In pTA1015 and pTA1060, the stopcodons of the convergently oriented mob gene and ORF2C are separated by only 10 and 21 basepairs, respectively. These short intergenic regions are part of sequences that have the potential to form secondary structures. An alignment of the deduced ORF2C protein sequences is shown in Fig. 6. A potential DNA binding helix-turn-helix motif is present in the N-terminal portion of these proteins. Such motifs are often present in bacterial transcription regulator proteins. Since the ORF2C of pTA1015 and pTA1060 are located adjacent to the mob genes, and interruption of ORF2C of pTA1015 resulted in increased mobilization frequencies, we conceived that the ORF2C products might be negative regulators of the mob genes. Support for this idea was not obtained, however, when the effects of intact and interrupted ORF2C on ß-galactosidase expression driven by the promoter of the mob gene of pTA1015 were studied (results not shown). Interruption of ORF2C neither had a clear effect on plasmid copy number or plasmid maintenance (results not shown). Therefore, the biological function of the ORF2C remains obscure.

The orf1/sipP modules on pTA1015 and pTA1040

We have identified before a module on pTA1015 and pTA1040 which contains an ORF (ORF1) and a sipP gene (Meijer et al., 1995b). ORF1 encodes a putative export protein, the function of which is unknown; and sipP encodes a functional type I signal peptidase (SPase). We assume that ORF1 and sipP constitute one functional module. This assumption is based on the following observations: (i) on both plasmids the genes have a similar structural organization; (ii) a high level of DNA homology was observed between both ORF1 and sipP genes and their short intergenic region; and (iii) single transcription products have been identified from which both the Orf1 product and SipP can be translated, indicating that they are organized in an operon (our unpublished results).

The rap module on pTA1040 and pTA1060

The plasmids pTA1040 and pTA1060 both contain a large ORF, the deduced protein sequence of which shows significant homology to the B.subtilis chromosomally-located rapA and rapB genes (Perego et al., 1994). In addition, we have identified a gene homologous to rapA/rapB, designated orfA, on yet another

cryptic plasmid from B.subtilis. The latter plasmid, pLS20, replicates, unlike pTA1015 and pTA1040, according to the theta mechanism (Meijer et al., 1995a). An alignment of the deduced protein sequences of the five rap-like genes is shown in Fig. 7. The percentages of similarities between the various proteins when compared pairwise are shown in Table 4. The corresponding genes from pTA1040 and pTA1060 have been named rap40 and rap60, respectively.

RapA was formerly identified as a glucose-starvation-induced protein and was named accordingly gsiA (Mueller et al., 1992). Expression of this gene was shown to cause a delay in the onset of sporulation (Mueller and Sonenshein, 1992). In vitro studies showed that the rapA gene product is a protein-aspartate phosphatase acting on SpoOF~P, a response regulator component of the phosphorelay system (Perego et al., 1994). As a consequence, the cellular level of SpoOA~P, which is the determining factor in the induction of sporulation (Hoch, 1993), is lowered, thereby preventing the onset of sporulation (Perego et al., 1994). The homologous rapB gene was identified as part of the European B.subtilis genome sequencing project (Perego et al., 1994). Also this gene product is a protein-aspartatate phosphatase active on SpoOF~P and, like RapA, expression of RapB results in delaying the onset of sporulation (Perego et al., 1994).

RapA is under the control of the ComP-ComA two-component signal transduction system (Mueller et al., 1992). RapB, however, is not induced by the ComP-ComA system, but seems to be

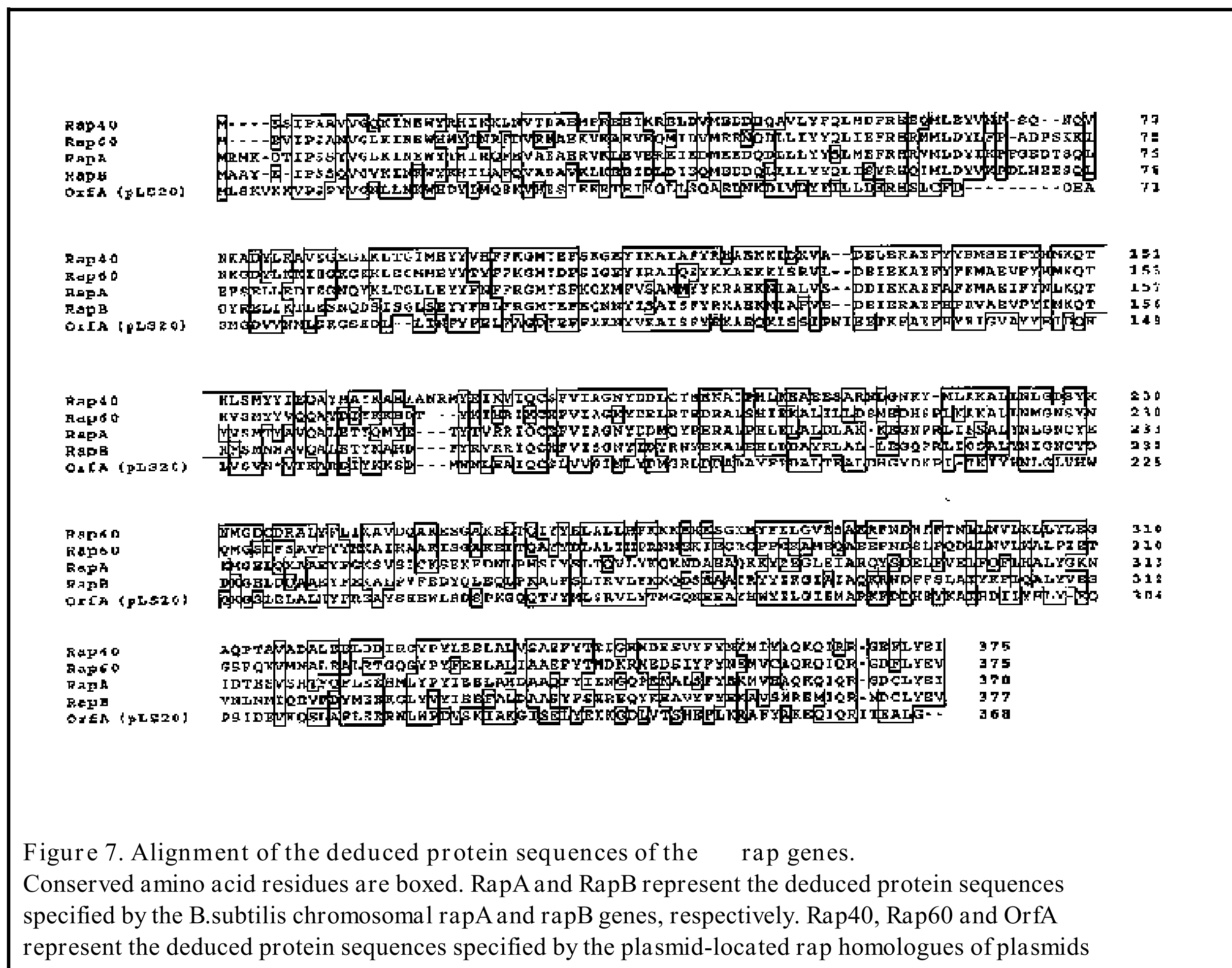


Figure 7. Alignment of the deduced protein sequences of the rap genes.
Conserved amino acid residues are boxed. RapA and RapB represent the deduced protein sequences specified by the B.subtilis chromosomal rapA and rapB genes, respectively. Rap40, Rap60 and OrfA represent the deduced protein sequences specified by the plasmid-located rap homologues of plasmids

regulated by the AbrB transition state regulator (Perego et al., 1994), which plays a role in several stationary phase processes (Strauch and Hoch, 1992). Thus, initiation of sporulation is prevented by RapA or RapB, the synthesis of these proteins being activated by different physiological conditions. The identification of genes on plasmids which show homology to rapA and rapB makes it tempting to believe that all the products encoded by these genes belong to an extended family of phosphatases, the members of which may be expressed under special physiological conditions.

Table 4. Similarities between homologous rap gene products of B.subtilis

	RapA	RapB	Rap40	Rap60	OrfA
RapA	-	66	51	55	38
RapB	65	-	57	54	34
Rap40	52	56	-	70	30
Rap60	56	53	69	-	31
OrfA	38	33	31	32	-

Similarities are shown as % of identical + similar amino acid residues

With respect to the regulation of rap40 and rap60, it may be relevant to note that upstream of these genes DNA sequences are present that show homology to the consensus ComA~P recognition site (Fig. 8), suggesting that these genes are under the control of the ComA-ComP system. In addition, the consensus SpoOA~P recognition sequence, 5'-TGNCGAA-3', is present just upstream of the putative RBS of rap40. This so-called "OA box" is not present upstream of rap60.

A small ORF, which could encode 44 amino acids, is located at the 3' end of the rapA gene in such a way that the coding regions of these genes overlap for 8 nucleotides. This small gene was called gsiAB by Mueller et al. (1992). Upstream of its putative ATG startcodon, a potential RBS (5'-AGGAG-3'; spacing of 10 nt), and downstream of the gsiAB gene a putative rho-independent transcriptional terminator, is present. This structural organization indicates that rapA and gsiAB may be transcribed as an operon (Mueller et al., 1992). Interestingly, small putative genes are also present at the 3' ends of the other four rapA homologues, and all of them have a structural organization similar to that of rapA/gsiAB (for potential ribosomal binding sites, see Table 1). The observation that all these rap homologues are followed by a small putative gene suggests that the latter have a biological function which is associated with the rap-gene function. We propose to rename gsiAB as rapAB. Similarly, we will designate the putative small genes at the 3' ends of rapB, rap40, rap60 and orfA as rapBB, rap40B, rap60B and orfAB, respectively. The deduced amino acid sequences encoded by these small genes are shown in Fig. 9. Analysis of these sequences revealed that, except for RapBB, their N-terminal parts have features typical for type I signal peptides: (i) a positively charged NH_2 terminus, followed by (ii) a stretch of hydrophobic residues, and (iii) a more polar C-region with potential SPase cleavage sites (von Heijne,

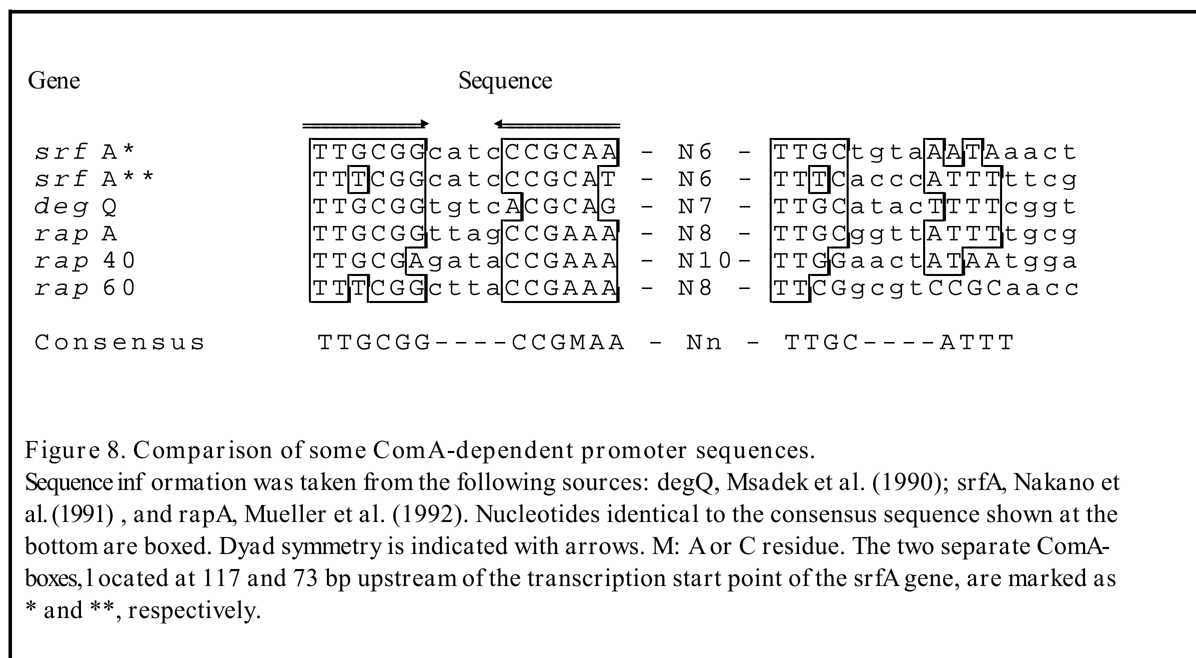


Figure 8. Comparison of some ComA-dependent promoter sequences.
Sequence information was taken from the following sources: degQ, Msadek et al. (1990); srfA, Nakano et al. (1991) , and rapA, Mueller et al. (1992). Nucleotides identical to the consensus sequence shown at the bottom are boxed. Dyad symmetry is indicated with arrows. M: A or C residue. The two separate ComA-boxes, located at 117 and 73 bp upstream of the transcription start point of the srfA gene, are marked as * and **, respectively.

1986; Simonen and Palva, 1993). This

suggests that, except for RapBB, the small gene products are secretory proteins.

Unique modules on pTA1015, pTA1040 and pTA1060

The genes and modules described above all are present on at least two of the three pTA-plasmids examined. In addition, unique modules were identified on pTA1040 and pTA1060.

Downstream of the rep gene of pTA1040 an ORF, designated ORF3C40, is located which could encode a protein of 201 amino acids. The rep gene and ORF3C40 are convergently transcribed and their are separated by 39 bp. The intergenic region contains dyad symmetry. The putative ATG startcodon of ORF3C40 is preceded by a potential RBS (AAGGAGG) with a spacing of 13 bp. No significant homology was found with other proteins when the deduced protein sequence of ORF3C40 was compared with available sequences in the

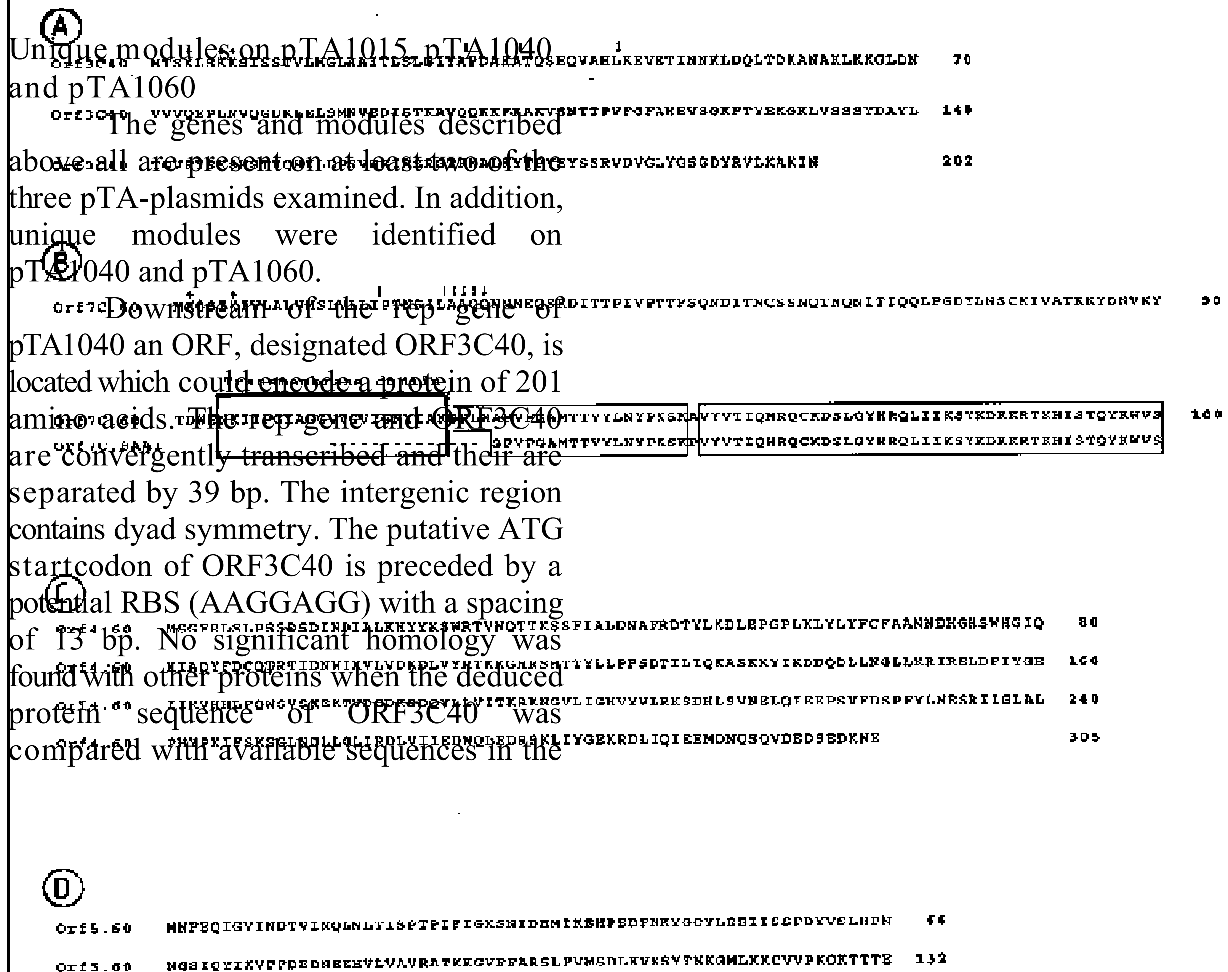
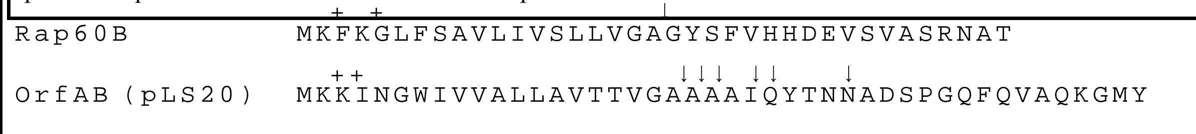


Figure 10. Deduced protein sequences specified by unique genes on pTA10140 and pTA1060.
(A) Deduced protein sequence of the putative ORF3C product of pTA1040. The positively charged amino acid residues at the NH_2 terminus (+) and potential SPase cleavage sites (9) are indicated above the sequence. (B) Deduced protein sequence of the ORF7C product of pTA1060, which is aligned with the partly sequenced 3' part of a homologous gene present on pBAA1 (conserved amino acid residues are boxed; ---: not sequenced). The positively charged amino acid residues at the NH_2 terminus (+) and the potential SPase cleavage sites (9) are indicated above the sequence. The potential transmembrane domain is boxed (double lined). (C) and (D) Deduced protein sequences of ORF4.60 and ORF5.60 of pTA1060.

RapAB MKSKWMSGLLLVAVGFSFTQVMVHAGETANTEGKTFHTAARNQT

RapBB MKYFTKKDSAGRIFNFMQKKGSIGTFTISGENRKNSPYSTVF

Rap40B MKFNALLLLIVCASLLIVSGSSFVIQQDSNVSVASRKAT



Rap60B MKFKGLFSAVLIVSLLVGAGYSFVHHDEVSVASRNAT

OrfAB (pLS20) MKKINGWIVVALLAVTTVGAAAAIQYTNNADSPGQFQVAQKGMY

Figure 9. Deduced protein sequences of the small genes that are translationally coupled to the rap genes.
Positively charged amino acid residues at the NH_2 terminus (+) and potential SPase cleavage sites (9) are indicated above the sequence.

databases. Analysis of this putative protein revealed typical features of a signal peptide at its N-terminus, implying that the ORF3C40 product may be secreted into the medium. The deduced protein sequence of the ORF3C40 product, together with the potential SPase cleavage sites, are shown in Fig. 10A.

Also in pTA1060, an ORF, ORF7C60, is located downstream of the rep gene. This ORF could encode a protein of 180 amino acids. Several similarities exist between ORF3C40 of pTA1040 and ORF7C60 of pTA1060. Both are convergently transcribed with respect to their rep gene and, like pTA1040, the intergenic region between their stopcodons is very small (30 bp), and this region also has dyad symmetry. Also the deduced protein sequence of the ORF7C60 product contains a putative signal peptide at its N-terminus. Despite these similarities, the deduced protein sequences of ORF3C40 and ORF7C60 show little homology. Unlike the putative ORF3C40 product, that of ORF7C60 has a potential transmembrane spanning segment (amino acids 95 to 115). According to the positive-inside-out-rule of Sipos and von Heyne (1993), the topology of the product of ORF7C60 is such that the N-terminal part of the protein is located at the outer side of the cytoplasmic membrane and the C-terminal part in the cytoplasm. No significant homology was found between the deduced protein sequence of the ORF7C60 product and protein sequences in the databases. However, an ORF is present downstream of the rep gene of pBAA1, the deduced amino acid sequence of which is almost identical to that of the C-terminal part of the deduced ORF7C60 product of pTA1060. In pBAA1, the directions of transcription of the ORF homologous to ORF7C60 and the rep gene are, as in pTA1060, convergent. Moreover, also in pBAA1, the short intergenic region contains dyad symmetry. The available sequence data of pBAA1 extend to only 216 bp downstream of the rep gene. As a consequence, only the 61 3'-terminal codons of the pBAA1-located ORF are known. An alignment of the deduced amino acid sequences of the two genes is shown in Fig. 10B, in which also the potential signal peptide and the region constituting the transmembrane helix are indicated. So far, the biological function of ORF7C60 is unknown.

pTA1060 contains, in addition, unique sequences between the palT-type SSO and ORF2C. This region contains two ORFs, designated ORF4.60 and ORF5.60. ORF4.60 could encode a protein of 305 amino acids. Its ATG startcodon is preceded by a potential RBS (AAGGAG) with a spacing of 9 bp. ORF5.60 could encode a protein of 132 amino acids. Its putative ATG startcodon is preceded by two potential RBSs; one (GGAG) with a spacing of 6 bp and the other (AAAGGGG) with a spacing of 15 bp. The coding regions of ORF4.60 and ORF5.60, which are both transcribed in the same direction as the rep gene, are separated by only 20 bp on which the two possible RBSs of orf5.60 are located. This structural organization suggests that the two ORFs are transcribed as an operon. A potential σ^A-type promoter is located 29 bp upstream of the startcodon of ORF4.60 (TTGAGT-- 17 nt --TATGAT). Neither the deduced protein sequence of ORF4.60, nor that of ORF5.60 shows significant homology with sequences available in the databases. Until now, the biological function of these putative genes is unknown.

DISCUSSION

In the present work we have determined the complete DNA sequences of three cryptic B.subtilis plasmids, pTA1015, pTA1040 and pTA1060, and tried to analyze the function of several of their

genes. A major conclusion from these studies is that these plasmids, in addition to modules that are involved in replication and conjugative mobilization and which are typical for RCM plasmids, contain modules that have not been identified before on other RCM plasmids. Whereas one of these modules, ORF2C, is shared by all three pTA-plasmids, modules were also detected that are present on two of the plasmids analyzed and others that are unique for one plasmid.

By comparing a large number of B.subtilis isolates, Uozumi et al. (1980) identified twenty strains that harboured a relatively small plasmid. Based on the size and the restriction profiles, these plasmids could be classified into six groups, represented by pTA1015, pTA1040, pTA1060, pTA1020, pTA1030, and pTA1030. Also plasmids isolated in independent studies by Tanaka et al. (1977; 1977) and Hara et al. (1983) can be classified into one of these six groups. A seventh group, represented by pFTB14, was later added by Yoshimura et al. (1983). pFTB14 was isolated from the B. amyloliquefaciens strain S294. B.subtilis and B.amyloliquefaciens are highly related. In fact, based on DNA homology studies the B.subtilis strain IFO3022 harbouring pTA1060 is now considered to be a strain of B.amyloliquefaciens (Seki et al., 1975). Thus, many, and perhaps all, B.subtilis and B.amyloliquefaciens RCM plasmids belong to one of the seven known plasmid groups. To determine whether the B.subtilis RCM plasmids pBS2 (Darabi et al., 1989), pLS11 (Hara et al., 1992b), pBAA1 (Devine et al., 1989), and pUH1 (Hara et al., 1991), from which the replication regions have been sequenced, also belong to one of these groups, their published size and restriction profiles were compared with members of the known groups. This approach showed that pBAA1 belongs to the pTA1020 group, pUH1 to the pTA1015 group, and pBS2 and pLS11 both belong to the pTA1060 group of plasmids. Plasmids within one group are expected to be highly homologous to each other, and may even be (nearly) identical. A very high level of homology was indeed observed when the corresponding sequenced regions of pTA1060, pBS2 and pLS11, all belonging to the pTA1060 group, were compared (pTA1060 and pBS2 are 97.5% identical over 2279 bp; pTA1060 and pLS11 are 96.9% identical over 1606 bp; and pBS2 and pLS11 are 93% identical over 1606 bp). Most of the differences between the corresponding plasmid regions are located in sequences that have the potential to form secondary structures and, therefore, these differences may result from sequencing errors. This would imply that the plasmids within one group could be identical. In agreement with this assumption is the observation that the observed homology between pBS2 and pTA1060 extends beyond the replication region: the deduced amino acid sequence of an ORF identified upstream of the rep gene of pBS2 (Darabi et al., 1989) is identical to part of the deduced protein sequence specified by the pTA1060-located rapA60 gene.

Based on its size and restriction pattern, pUH1 belongs to the pTA1015 class of plasmids. The complete sequence of pUH1, although not not present in the databases, has been published (Hara et al., 1992c). Comparison of the sequences of these two plasmids showed an identity of 99.3% over 5807 bp. Also in this case, the few differences in the DNA sequence between the two plasmids are predominantly observed in plasmid regions that have the potential to form secondary structures. These data indicate that pUH1 and pTA1015 are (nearly) identical. The same may apply to other plasmids which have been classified in one and the same group.

The alignment of different Rep proteins specified by Bacillus RCM plasmids (Fig. 2A), shows that the C-

terminus of the Rep protein of pBS2 is 39 amino acids shorter than that of the other Rep proteins examined. As described above, we consider pBS2 and pTA1060 to be identical. When the DNA sequence of the 3' region of the rep gene of pTA1060 is compared with the corresponding region of pBS2, only three differences are observed [one additional C at position 1943, an A/T substitution at position 1985, and a deletion of one T at position 2011, in the sequence of pBS2 (Darabi et al., 1989)]. Considerable variation seems to be present at the N-termini of the various Rep proteins. The differences between pUH1 and pLS11 can be traced back to three differences on DNA level, which may be due to sequencing errors. We consider it likely that the Rep proteins of pUH1 and pLS11 are in fact identical to those of pTA1015 and pTA1060. The reading frames for the Rep proteins of pBAA1 and pBS2 are open upstream of their published putative startcodon. When the deduced amino acid sequences of these upstream regions are aligned with the N-terminal sequences of the Rep proteins of the pTA-plasmids and of pFTB14, they are all identical. Since the DNA sequence of pBS2 is highly identical to that of pTA1060, this was to be expected. The observed conservation between the rep genes, including that of pBAA1, suggests that the Rep protein encoded on pBAA1, like those encoded on the other plasmids discussed here, contains 339 amino acids (only the Rep protein of pTA1060 contains 340 amino acids). In agreement with this idea is the observation that a potential RBS (AAGGAG) is present upstream of the proposed new startcodon of the pBAA1 rep gene.

pBAA1
pTA1060
pFTB14
pTA1015
pTA1040

Figure 11. Phylogenetic tree representation of the evolutionary distance of rep genes of plasmids pTA1015, pTA1040, pTA1060, pBAA1 and pFTB14.

The DNA sequences of the replication regions of five plasmids representing members of the seven known groups of RCM plasmids in B.subtilis and B.amyloliquefaciens are now known. The levels of homology between the replication regions of these plasmids are indicative for their relatedness. A dendrogram illustrating the relatedness of Rep proteins encoded by these plasmids is shown in Fig. 11. This comparison shows that pBAA1 and pTA1060 are the most closely related plasmids. In agreement with this conclusion is the observation that these plasmids contain an almost identical gene downstream of their rep gene. Among the five plasmids compared, pTA1040 has the most deviating sequence of the replication region. Of these plasmids, only pTA1040 contains a palT2-type SSO. The other plasmids (the SSO of pFTB14 is not known), contain a palT1-type SSO (Meijer et al., 1995d). This observation supports our idea that the other four plasmids are more related to each other than to pTA1040.

The results presented in this article showed that pTA1015 and pTA1060 contain functional mob genes which are preceded by putative oriT target sites. Hara et al. (1992a) claimed that a region of pUH1 (99.3% identical to pTA1015 over 5807 bp) encodes γ-GTP. However, the present results showed that this region is almost identical to the mob region of pTA1015 and is not correlated with γ-GTP production, at least not in the B.subtilis strain BD630. Koehler and Thorne (1987) have reported that also pLS19, another plasmid belonging to the group represented by pTA1015, did not confer γ-GTP activity to its host. These observations make it highly unlikely that the region of pUH1 analyzed by Hara et al. (1992a) encodes a γ-GTP enzyme. At present, it can not be ruled out, however,

that plasmid-encoded genes activate a chromosomally-encoded γ-gtp gene. We tried to test this possibility by studying the effects of pTA1015 derivatives on the expression of lacZ placed under the control of the promoter of the chromosomally-located γ-gtp gene. For this purpose, we attempted to amplify the promoter region of γ-gtp genes of different B.subtilis strains (BD630, 8G5, UM4) by PCR techniques using a set of primers based on known sequences of the B.subtilis SJ138 γ-gtp gene. These attempts failed to amplify fragments of the expected size, indicating that the strains tested do not contain a γ-gtp gene, or that the degree of diversity on the DNA level was too high.

These studies have revealed the presence of several genes which, so far, had not been identified on RCM plasmids. One of these genes, with related but non-identical representatives on pTA1015 and pTA1040, encodes a functional type I signal peptidase (Meijer et al., 1995b). Two other related genes, rap40 and rap60, present on pTA1040 and pTA1060, respectively, are likely to be involved in the regulation of sporulation. Homologues of the plasmid-encoded sip and rap genes are also present on the chromosome of B.subtilis (van Dijl et al., 1992; Mueller et al., 1992; Perego et al., 1994). An intriguing question concerns the possible advantage for the cell of having plasmids containing homologues of chromosomally-encoded genes. Several answers to this question are conceivable. First, although the plasmid- and chromosomally-located genes are related, they are not identical. Therefore, each of these genes may have its own specific function. The presence of multiple homologues of a gene may enable the host cell to respond optimally to changing conditions. Second, the presence of genes on multi-copy plasmids can lead to increased amounts of the product. This may be advantagous under conditions in which the gene product is a limiting factor. We have already shown before that the chromosomally-encoded Signal peptidase (SipS) can be a limiting factor with certain secretory proteins (van Dijl et al., 1992). Conceivably, the presence of plasmid-encoded rap genes may lead to higher levels of Rap proteins and/or synthesis of Rap proteins under a wider range of environmental conditions, which in turn will lead to an extended delay in the onset of sporulation than is obtained with the chromosomally encoded single copy genes alone. Taken into account that these plasmids are all isolated from industrial Bacillus strains, this may suggest that selection for increased protein production has been the driving force for the plasmid location of these genes.

Future research concerning the sip and rap genes will be needed for a better understanding of the precise function of these genes and why they are plasmid-located. Furthermore, it will be challenging to determine the biological function of the small genes that are translationally coupled to the rap genes and to unravel the function of the additional plasmid-located genes identified in this work.

MATERIALS AND METHODS

Bacterial strains, plasmids and media. Bacterial strains and plasmids used are listed in Table 5. TY medium, used for culturing Escherichia coli and B.subtilis, contained Bacto tryptone (1%), Bacto yeast extract (0.5%) and NaCl (1%). TY plates contained in addition 2% agar. Minimal media for B.subtilis were as described by van Sinderen et al. (1994). DGYP medium was as described by Ryoichi et al. (1991). Sporulation medium for B.subtilis was as described by Schaeffer (1965). Kanamycin and ampicillin were added to final concentrations of 10 and 50 µg/ml, respectively. When regenerating protoplasts were selected for resistance to kanamycin, the concentration of this antibiotic was increased to 150 µg/ml.

DNA techniques. DNA manipulations were carried out according to Sambrook et al. (1989). Restriction enzymes were obtained commercially and used as indicated by the suppliers. Plasmid DNA was isolated by the alkaline lysis method (Sambrook et al., 1989). DNA fragments were isolated from gels using the Qiaex Gel Extraction Kit (Qiagen Inc., Chatsworth, USA). Total DNA lysates were prepared as described before (Meijer et al., 1995c). Southern transfers to Gene Screen plus membranes were carried out as described (Meijer et al., 1995d). Probe labelling, DNA hybridization conditions and washing steps were performed using the enhanced chemiluminescence DNA labelling and detection system (Amersham International, Amersham, UK). DNA sequences were determined by the dideoxy chain termination method (Sanger et al., 1977) using either the T7 DNA polymerase sequencing kit (Pharmacia, Uppsala, Sweden), or the automated A.L.F. sequencer (Pharmacia, Uppsala, Sweden). In the latter method, Fitc-labeled primers or Fitc-labeled dATP were used in the reactions using the standard protocol described by Zimmermann et al. (1990). DNA sequences and deduced amino acid sequences were analyzed using version 6.8 of the PCGene Analysis Program (Intelligenetics Inc., Mountain View, CA, USA). The FASTA algorithm of Lipman and Pearson (1985) was used for protein comparisons in the Swiss protein and genomic DNA databank sequences (release January 1995; MIPS, Martinsried, FRG) and the EMBL Nucleotide Sequence Database. The RDF2 program was used to evaluate sequence similarities (Pearson, 1990). To calculate z values, the KTUP value was set at 2, and 500 random shuffles of the test sequences were performed. Alignments with z values greater than 6 were considered significant; alignments with z values below 3 were considered insignificant.

Transformation of B.subtilis and E.coli. Competent cells and protoplasts of B.subtilis were prepared and transformed as described (Bron, 1990; Chang and Cohen, 1979). $CaCl_2$-treated E.coli cells were transformed as described by Sambrook et al. (1989).

γ-GTP activity assay. Measurements of γ-GTP acitivity were performed as described before (Ryoichi et al., 1991). Purified Bovine-kidney γ-GTP was obtained from Sigma (Axel, The Netherlands).

Mobilization experiments. Donor and recipient cells were grown until late logarithmic phase. Appropriate dilutions of the cultures were plated and incubated overnight at 37 °C to determine the numbers of viable donor and recipient cells. Next, equal volumes (1 ml) of cultures containing donor and recipient cells were mixed, pelleted, resuspended in 0.5 ml of TY medium, and plated onto non-selective agar. After overnight mating at 37 °C, the cell layer was collected and resuspended in TY broth. Appropriate dilutions were plated on selective agars. The numbers of transconjugants were scored after overnight

Table 5. Strains and plasmids used

Bacterial strain	Relevant properties	Reference
E.coli		
JM83	thi, Δ(lac-proAB) [F proAB lacI^q lacZΔM15]	Messing, 1979
B.subtilis		
8G5	trpC2, his, met, tyr-1, ade, nic, ura, rib	Bron and Venema, 1972
PSL1	leuA8, arg15, thrA, recA4, r^-_M m^-_M	Ostroff and Pène, 1983
BD630	hisA1, leu-8, metB5	Albano et al., 1989
3335 UM4	containing pLS20	Koehler and Thorne, 1987
1012Cm	leuA8, metB5, r^-_M m^+_M, Cm^R	Oskam et al., 1991
Plasmids	**Relevant properties**	**Reference**
pTA1015	Cryptic plasmid from B.subtilis IAM1028, 5.8 kb	Uozumi et al., 1980
pTA1040	Cryptic plasmid from B.subtilis IAM1232, 7.8 kb	Uozumi et al., 1980
pTA1060	Cryptic plasmid from B.subtilis IFO3022, 8.7 kb	Uozumi et al., 1980
pTAB11A	pTA1015 containing the KmR gene of pKM1, ORF2C15$^-$, 7.0 kb	This study
pTAB11B	As pTAB11A, reversed orientation of the Km^R gene	This study
pTAB13	Deletion derivative of pTAB11A, mob15$^-$, 6.5 kb	This study
pTAB31	pTA1015 containing the Km^R gene of pKM2, ORF1.15$^-$; 7.2 kb	This study
pKM1	pUC7 containing Km^R gene from pJH1	Kiel et al., 1987
pKM2	pMTL25 derivative containing Km^R gene from pKM1, Ap^R, Km^R	Meijer et al., 1995a
pUC18	High-copy number E.coli vector containing MCS, Ap^R	Yanish-Perron et al., 1985
pBB2	pTA1060 containing Cm^R and Km^R gene, mob60$^+$	Bron, 1990
pUB110	Natural plasmid from S.aureus, mob$^+$, Km^R, Bleo^R, 4.5 kb	Oskam et al., 1991

Abbreviations: Km^R, Bleo^R, Ap^R: antibiotic markers conferring resistance to kanamycin, bleomycin, chloramphenicol and ampicillin, respectively. (-) denotes interruption of the gene/ORF by insertion of the Km^R gene or deletion of internal sequences.

incubation at 37 °C.

Plasmid maintenance assay. Plasmid maintenance was determined as described before (Meijer et al., 1995c). Briefly, overnight cultures grown in selective media were diluted 100,000-fold in non-selective media, after which the percentage of plasmid-containing cells (that is: the percentage of antibiotic-resistant cells) was determined as a function of growth-time. Growth was normally followed for 60 generations.

PCR techniques. PCR was carried out essentially as described by Innis and Gelfand (1990). The proofreading-proficient Vent DNA polymerase (New England Biolabs, Beverly, MA, USA) was used throughout. Template DNA was denatured for 1 min at 94 °C. Next, primers were used to amplify DNA fragments in 30 cycles of denaturation (30 sec; 94°C), primer annealing (1 min 50°C), and DNA synthesis (3 min 73°C).

Construction of plasmids. Plasmids pTAB11A and pTAB11B were constructed by replacing the 0.24 kb Csp45I fragment of pTA1015 by the 1.4 kb AccI fragment of pKM1. The latter fragment contains the Km^R gene originating from the Streptococcus faecalis plasmid pJH1. The only difference between pTAB11A and pTAB11B is the orientation of the cloned insert (see Fig. 4). pTAB13 was obtained by deleting the 0.5 kb $Eco47^{III}$ fragment, which is located internally in the mob15 gene, of pTAB11A. pTAB31 was constructed by cloning the 1.4 XbaI fragment of pKM2, containing the Km^R gene, into the unique SpeI site of pTA1015. In the latter construct the direction of transcription of the Km^R gene is in the same orientation as that of the rep gene.

ACKNOWLEDGEMENTS

We thank Dr. P.Terpstra for his contribution in DNA sequencing and Henk Mulder for preparing the figures. We are indebted to Leendert Hamoen, Bert-Jan Haijema, Rob Meima and Steven de Jong for helpful discussions. Funding for the project, of which this work is a part, was provided by STW (Stichting Technische Wetenschappen, The Netherlands) and Gist-Brocades B.V. (Delft, The Netherlands). Part of this work was supported by a CEC grant to S. Bron (BRIDGE Program BIOT-CT910268)

REFERENCES

Albano, M., Breitling, R., and Dubnau, D.A. (1989). Nucleotide sequence and genetic organization of the Bacillus subtilis comG operon. J. Bacteriol. 171: 5386-5404.

Bron, S., Bosma, P., Van Belkum, M.J., and Luxen, E. (1987). Stability function in the Bacillus subtilis plasmid pTA1060. Plasmid 18: 8-15.

Bron, S. (1990). Plasmids. In Molecular Biological Methods for Bacillus. Harwood, C.R. and Cutting, S.M. (eds.) Chichester, UK: John Wiley & Sons Ltd. pp. 75-174.

Bron, S., Holsappel, S., Venema, G., and Peeters, B.P.H. (1991). Plasmid deletion formation between short direct repeats in Bacillus subtilis is stimulated by single-stranded rolling- circle replication intermediates. Mol. Gen. Genet. 226: 88-96.

Bron, S. and Venema, G. (1972). Ultraviolet inactivation and excision repair in Bacillus subtilis. I. Construction and characterization of a transformable eightfold auxotrophic strain and two ultraviolet-sensitive derivatives. Mutat. Res 15: 1-10.

Catchpole, I., Thomas, C., Davies, A., and Dyke, K.G.H. (1988). The nucleotide sequence of Staphylococcus aureus plasmid pT48 conferring inducible macrolide -lincosamide -streptogramin B resistance

and comparison with similar plasmids expressing constitutive resistance. J. Gen. Microbiol. 134: 697-709.

Chang, S., Chang, S., and Gray, O. (1987). Structural and genetic analysis of a par locus that regulates plasmid partition in Bacillus subtilis. J. Bacteriol. 169: 3952-3962.

Chang, S. and Cohen, S.N. (1979). High frequency transformation of Bacillus subtilis protoplasts by plasmid DNA. Mol. Gen. Genet. 168: 111-115.

Darabi, A., Forough, R., Bhardwaj, G., Watabe, M., Goodarzi, G., Gross, S.C., and Watabe, K. (1989). Identification and nucleotide sequence of the minimal replicon of the low copy number plasmid pBS2. Plasmid 22: 281-286.

Devine, K.M., Hogan, S.T., Higgins, D.G., and McConnell, D.J. (1989). Replication and segregational stability of the Bacillus plasmid pBAA1. J. Bacteriol. 171: 1166-1172.

Dubnau, D. (1983). Molecular cloning in Bacillus subtilis. In Experimental manipulation of gene expression. Inouye, E. (ed.) New York: Academic Press Inc. pp. 33-51.

Ehrlich, S.D. (1977). Replication and expression of plasmids from Staphylococcus aureus in Bacillus subtilis. Proc. Natl. Acad. Sci. USA 74: 1680-1682.

Gennaro, M.L., Kornblum, J., and Novick, R.P. (1987). A site-specific recombination function in Staphylococcus aureus plasmids. J. Bacteriol. 169: 2601-2610.

Gros, M.F., Bidnenko, V., and Ehrlich, S.D. (1994). Active site of the replication protein of the rolling circle plasmid pC194. EMBO J. 13: 4412-4420.

Gruss, A. and Ehrlich, S.D. (1989). The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. 53: 231-241.

Haima, P., Bron, S., and Venema, G. (1987). The effect of restriction on shotgun cloning and plasmid stability in Bacillus subtilis Marburg. Mol. Gen. Genet. 209: 335-342.

Haima, P., van Sinderen, D., Bron, S., and Venema, G. (1990). An improved β-galactosidase α-complementation system for molecular cloning in Bacillus subtilis. Gene 93: 41-47.

Hara, T., Zhang, J., and Ueda, S. (1983). Identification of plasmids linked with polyglutamate production in Bacillus subtilis (natto). J. Gen. Appl. Microbiol. 29: 345-354.

Hara, T., Nagatomo, S., Ogata, S., and Ueda, S. (1991). Molecular structure of the replication origin of a Bacillus subtilis (natto) plasmid, pUH1. Appl. Environ. Microbiol. 57: 1838-1841.

Hara, T., Nagatomo, S., Ogata, S., and Ueda, S. (1992a). The DNA sequence of gamma-glutamyltranspeptidase gene of Bacillus subtilis (natto) plasmid pUH1. Appl. Environ. Microbiol. 37: 211-215.

Hara, T., Nakajima, K., Saito, H., Ishizaki, A., Ogata, S., and Ueda, S. (1992b). Sequence analysis of replication origin of plasmid pLS11 of Bacillus subtilis IFO 3022. Biosci. Biotech. Biochem. 56: 223-227.

Hara, T., Yasuda, S., Ogata, S., Soeda, E., and Ueda, S. (1992c). Complete nucleotide sequence of Bacillus subtilis (natto) plasmid responsible for gamma-polyglutamate synthesis. J. Fac. Agr. Kyushu univ. 36: 209-218.

Hara, T., Ogata, S., and Ueda, S. (1993). Plasmid distribution in gamma-polyglutamate-producing Bacillus strains isolated from "Dan-Douchi", a "natto"-like non-salty fermented soybean food in China. J. Gen. Appl. Microbiol. 39: 75-82.

Hoch, J.A. (1993). Regulation of the Phosphorelay and the Initiation of Sporulation in Bacillus-subtilis. Annu. Rev. Microbiol. 47: 441-465.

Innis, M.A. and Gelfand, D.H. (1990). Optimization of PCRs. In PCR protocols: a guide to methods and applications. Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds.) Academic Press.

Jannière, L., Gruss, A., and Ehrlich, S.D. (1993). Plasmids. In Bacillus subtilis and other Gram-positive bacteria; Biochemistry, physiology and molecular genetics. Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds.) Washington: American Society for Microbiology, pp. 625-644.

Josson, K., Soetaert, P., Michiels, F., Joos, H., and Mahillon, J. (1990). Lactobacillus hilgardii plasmid pLAB1000 consists of two functional cassettes commonly found in other Gram-positive organisms. J. Bacteriol. 172: 3089-3099.

Kiel, J.A.K.W., van der Vossen, J.P.M.J., and

Venema, G. (1987). A general method for the construction of Escherichia coli mutants by homologous recombination and plasmid segregation. Mol. Gen. Genet. 207: 294-301.

Koehler, T.M. and Thorne, C.B. (1987). Bacillus subtilis (natto) plasmid pLS20 mediates interspecies plasmid transfer. J. Bacteriol. 169: 5271-5278.

Lacks, S.A., Lopez, P., Greenberg, B., and Espinosa, M. (1986). Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol. 192: 753-765.

Lampson, B.C. and Parisi, J.T. (1986a). Nucleotide sequence of the constitutive macrolide-lincosamide-streptogramin B resistance plasmid pNE131 from Staphylococcus epidermidis and homologies with Staphylococcus aureus plasmids pE194 and pSN2. J. Bacteriol. 167: 888-892.

Lampson, B.C. and Parisi, J.T. (1986b). Naturally occurring Staphylococcus epidermidis plasmid expressing constitutive mocrolide-lincosamide-streptogramin B resistance contains a deleted attenuator. J. Bacteriol. 166: 479-483.

Lipman, D.J. and Pearson, W.R. (1985). Rapid and sensitive protein similarity searches. Science 227: 1435-1441.

Mahillon, J. and Seurinck, J. (1988). Complete nucleotide sequence of pGI2, a Bacillus thuringiensis plasmid containing Tn4430. Nucleic Acids Res. 16: 11827-11828.

McKay, L.L. (1983). Functional properties of plasmids in lactic streptococci. Antonie van Leeuwenhoek 49: 259-274.

Meijer, W.J.J., de Boer, A., van Tongeren, S., Venema, G., and Bron, S. (1995a). Characterization of the replication region of the Bacillus subtilis plasmid pLS20: a novel type of replicon. Nucleic Acids Res. In press

Meijer, W.J.J., de Jong, A., Wisman, G.B.A., Tjalsma, H., Venema, G., Bron, S., and van Dijl, J.M. (1995b). The endogenous Bacillus subtilis (natto) plasmids pTA1015 and pTA1040 contain signal peptidase-encoding genes: identification of a new structural module on cryptic plasmids. Mol. Microbiol. In press

Meijer, W.J.J., van der Lelie, D., Venema, G., and Bron, S. (1995c). Effects of the generation of single-stranded DNA on the maintenance of plasmid pMV158 and derivatives in Lactococcus lactis. Plasmid 33: 91-99.

Meijer, W.J.J., Venema, G., and Bron, S. (1995d). Characterization of single strand origins of cryptic rolling-circle plasmids from Bacillus subtilis. Nucleic Acids Res. 23: 612-619.

Meister, A. and Tate, S. (1976). Glutathione and related γ-glutamyl compounds: Biosynthesis and utilization. Annu. Rev. Biochem. 45: 559-604

Messing, J. (1979). A multipurpose cloning system based on the single stranded DNA bacteriophage M13. Recombinant DNA technical Bullitin, NIH publication 79-99.2: 43-48.

Moran, C.P., Lang, N., LeGrice, S.F.J., Lee, G., Stephans, M., Sonenshein, A.L., Pero, J., and Losick, R. (1982). Nucleotide sequence that signal the initiation of transcription and translation in Bacillus subtilis. Mol. Gen. Genet. 186: 339-346.

Msadek, T., Kunst, F., Klier, A., Rapoport, G., and Dedonder, R. (1990). Signal transduction pathway controlling synthesis of a class of degradative enzymes in Bacillus subtilis: expression of the regulatory genes and analysis of mutations in degS and degU. J. Bacteriol. 172: 824-834.

Mueller, J.P., Bukusoglu, G., and Sonenshein, A.L. (1992). Transcriptional regulation of Bacillus subtilis glucose starvation-inducible genes: control of gsiA by the ComP-ComA signal transduction system. J. Bacteriol. 174: 4361-4373.

Mueller, J.P. and Sonenshein, A.L. (1992). Role of the Bacillus subtilis gsiA gene in regulation of early sporulation gene expression. J. Bacteriol. 174: 4374-4383.

Murai, M., Miyashita, H., Araki, H., Seki, T., and Oshima, Y. (1987). Molecular structure of the replication origin of a Bacillus amyloliquefaciens plasmid pFTB14. Mol. Gen. Genet. 210: 92-100.

Nakano, M.M., Magnuson, R., Myers, A., Curry, J., Grossman, A.D., and Zuber, P. (1991). srfA is an operon required for surfactin production, competence development, and efficient sporulation in

Bacillus subtilis. J. Bacteriol. 173: 1770-1778.

Novick, R.P. (1989). Staphylococcal plasmids and their replication. Annu. Rev. Microbiol. 43: 537-565.

Oskam, L., Venema, G., and Bron, S. (1991). The large Bacillus plasmid pTB19 contains two integrated rolling-circle plasmids carrying mobilization functions. Plasmid 26: 30-39.

Ostroff, G.R. and Pène, J.J. (1983). Molecular cloning with bifunctional plasmid vectors in Bacillus subtilis. III. Isolation of a spontaneous mutant of B.subtilis with enhanced transformability for Escherichia coli propagated chimaeric plasmids. J. Bacteriol. 156: 934-936.

Pearson, W.R. (1990). Methods Enzymology. Doolittle, R.F. (ed.) pp. 63-98.

Perego, M., Hanstein, C., Welsh, K.M., Djavakhishvili, T., Glaser, P., and Hoch, J.A. (1994). Multiple protein aspartate phosphatases provide a mechanism for the integration of diverse signals in the control of development in Bacillus subtilis. Cell 79: 1047-1055.

Priebe, S.D. and Lacks, S.A. (1989). Region of the streptococcal plasmid pMV158 required for conjugative mobilization. J. Bacteriol. 171: 4778-4784.

Projan, S.J. and Archer, G.L. (1989). Mobilization of the relaxable Staphylococcus aureus plasmid pC221 by the conjugative plasmid pGO1 involves three pC221 loci. J. Bacteriol. 171: 1841-1845.

Ryoichi, K., Toru, M., Shigenori, O., Moriyuki, S., and Kazuo, Y. (1991). Process for the production of gamma-glutamyl transpeptidase. Patent Nr. 91101620.2: 1-22.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Schaeffer, P., Millet, I., and Aubert, J. (1965). Catabolite repression of bacterial sporulation. Proc. Natl. Acad. Sci. USA 54: 704-711.

Seery, L. and Devine, K.M. (1993). Analysis of features contributing to activity of the single-stranded origin of Bacillus plasmid pBAA1. J. Bacteriol. 175: 1988-1994.

Seery, L.T., Nolan, N.C., Sharp, P.M., and Devine, K.M. (1993). Comparative analysis of the pC194 group of rolling circle plasmids. Plasmid 30: 185-196.

Seki, T., Oshima, T., and Oshima, Y. (1975). Taxonomic study of Bacillus by deoxyribonucleic acid-deoxyribonucleic acid hybridization and interspecific transformation. Int. J. Syst. Bacteriol. 25: 258-270.

Selinger, L.B., McGregor, N.F., Khachatourians, G.G., and Hynes, M.F. (1990). Mobilization of closely related plasmids pUB110 and pBC16 by Bacillus plasmid pXO503 requires trans-acting open reading frame β. J. Bacteriol. 172: 3290-3297.

Simonen, M. and Palva, I. (1993). Protein secretion in Bacillus species. Microbiol. Rev. 57: 109-137.

Sipos, X. and von Heijne, X. (1993). ??? Eur. J. Biochem. 213: 1313-1320?

Strauch, M.A. and Hoch, J.A. (1992). Transition state regulators: sentinels of Bacillus subtilis post-exponential gene expression. Mol. Microbiol. 7: 337-342.

Suzuki, H., Kumagai, H., Echigo, T., and Tochikura, T. (1989). DNA sequence of the Escherichia coli K-12 gamma-glutamyltranspeptidase gene, ggt. J. Bacteriol. 171: 5169-5172.

Tanaka, T., Kuroda, M., and Sakaguchi, K. (1977). Isolation and characterization of four plasmids from Bacillus subtilis. J. Bacteriol. 129: 1487-1494.

Tanaka, T. and Koshikawa, T. (1977). Isolation and characterization of four types of plasmids from Bacillus subtilis (natto). J. Bacteriol. 131: 699-701.

Uozumi, T., Ozaki, A., Beppu, T., and Arima, K. (1980). New cryptic plasmid of Bacillus subtilis and restriction analysis of other plasmids found by general screening. J. Bacteriol. 142: 315-318.

van der Lelie, D., Bron, S., Venema, G., and Oskam, L. (1989). Similarity of minus origins of replication and flanking open reading frames of plasmids pUB110, pTB913 and pMV158. Nucleic Acids Res. 17: 7283-7294.

van der Lelie, D., Wösten, H.A.B., Bron, S., Oskam, L., and Venema, G. (1990).

Conjugal mobilization of streptococcal plasmid pMV158 between strains of Lactococcus lactis subsp. lactis. J. Bacteriol. 172: 47-52.

van Dijl, J.M., de Jong, A., Vehmaanperä, J., Venema, G., and Bron, S. (1992). Signal peptidase I of Bacillus subtilis: patterns of conserved amino acids in prokaryotic and eukaryotic type I signal peptidases. EMBO J. 11: 2819-2828.

van Sinderen, D., Ten Berge, A., Haijema, B.J., Hamoen, L., and Venema, G. (1994). Molecular cloning and sequence of comK, a gene required for genetic competence in Bacillus subtilis. Mol. Microbiol. 11: 695-703.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14: 4683-4690.

Yanish-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33: 103-119.

Yoshimura, K., Yamamoto, O., Seki, T., and Oshima, Y. (1983). Distribution of heterogeneous and homologous plasmids in Bacillus spp. Appl. Environ. Microbiol. 46: 1268-1275.

Zimmermann, J., Voss, H., Schwager, C., Stegemann, J., Erfle, H., Stucky, K., Kristensen, T., and Ansorge, W. (1990). A simplified portocol for fast plasmid DNA sequencing. Nucleic Acids Res. 18: 1067