AN INVESTIGATION INTO THE PARTITION FUNCTIONS OF THE BROAD-HOST-RANGE PLASMID pTF-FC2

Anthony S. G. Smith

A dissertation submitted in fulfilment of the requirements of the Degree of Doctor of Philosophy in the Faculty of Science, University of Cape Town

Cape Town, August, 1997

The University of Cape Town has been given the right to reproduce this thesis in whole or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Scientific discovery is voyage not a destination

Contents

Acknowledgements

I wish to thank my supervisor, Professor Doug Rawlings who has guided me through this long and at times perplexing project. He was always ready to help and advise no matter how busy he was and for this I am very grateful. It is his enthusiasm that kept me going even when partition systems looked like replication systems and everything looked totally inexplicable.

No scientist can work in isolation. It is only through being able to discuss and bounce ideas off your colleagues that your models and theories can be tested and refined. I have been very lucky to work in the *Thiobacillus* Research Unit with a team of people who are motivated, dedicated and smart. My particular thanks to Dr Ros Chapman and Dr Shelly Deane for answering my perpetual questioning and responding to my theorising with patient good humour. To Dr Cliff Dominy and Dr Will Bourn with whom I have spent many a late night discussing this work and other aspects of science, thanks. Thanks to Di James whose boundless enthusiasm and energy assisted in making sequencing an entertainment. I would also like to thank all the other members past and present of the Microbiology Department for creating an enjoyable work environment.

My deepest gratitude goes to Jaana-Maria Ball for putting up with me during this long project.

I acknowledge the financial support of GENCOR and the FRD without whose financial assistance none of this work would have been possible. Thanks also to Kenn Gerdes and Rasmus Bugge Jensen for the kind gifts of pOU82 and pKG339, especially the former which has saved me many hours at the bench. I am also extremely grateful to Susan Gottesman for the kind gift of the *clpX* and *lon* mutants.

Abbreviations

A	adenine
Ap	ampicillin
ATP	adenosine triphosphate
bp	base pair(s)
C	cytosine
Cb	carbenecillin
Cm	chloramphenicol
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylene-diaminetetra-acetic acid
FBI	fold back inhibition
EtBr	ethidium bromide
G	guanine
IHF	integration host factor
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase pair(s)
kDa	kiloDalton
Km	kanamycin
LA	Luria Bertani Agar
LB	Luria Bertani Broth
Mr	relative molecular mass
nt(s)	nucleotide(s)
OD_{600}	optical density at 600nm
OD_{550}	optical density at 550nm
OD_{420}	optical density at 420nm
ONPG	o-nitrophenyl-β-D-galactopyranoside
ORF(s)	open reading frame(s)
*Ori*C	chromosomal origin of replication
*Ori*V	vegetative origin of replication
PAGE	polyacrylamide gel electrophoresis

PCR	polymerase chain reaction
Pol I	DNA polymerase I
RBS	ribosome binding site
RNA	ribonucleic acid
MCS	multiple cloning site
mM	millimolar
mRNA	messenger ribonucleic acid
RNAse	ribonuclease
SDS	sodium Dodecyl Sulphate
Sm	streptomycin
Sp	species
T	thymine
TB	terrific broth
TIR	translational initiation region
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

List of Figures

List of Tables

Abstract

The broad-host-range *Thiobacillus ferrooxidans* plasmid pTF-FC2 is stably inherited over many generations despite a low copy number. The *pas* genes which lie between the *repB* primase and the *repA* helicase encode a proteic plasmid stabilisation system and are capable of stabilising the unstable heterologous R1 replicon present on pOU82. The deletion of the *pas* genes has been shown not to change the copy number of mutant plasmids. This suggested that the *pas* genes are not involved in replication and function as a stabilisation cassette. The *pasA* gene encodes an antidote, the *pasB* gene a toxin which exerts a bacteriocidal effect in an *E. coli* host and the *pasC* gene a protein which moderates the toxic effect of PasB. The PasC is unique in proteic plasmid stabilisation systems and reduces PasB toxicity only in the presence of PasA. PasA is able to repress the *pas* promoter and the addition of PasB increases this repression. PasC has been shown not to effect the *pas* promoter by itself. In the presence of PasA and PasB, PasC reduces the ability of PasA and PasB to repress the *pas* promoter. PasC is thought to stabilise the interaction of PasA and PasB and in doing so reduces their ability to function as repressors of the *pas* operon. The IncQ plasmid RSF1010 which has similarity to pTF-FC2 has two genes in a position analogous to the *pasA*, *pasB* and *pasC* genes. These genes have been found to be unable to function as a plasmid stabilisation system.

The ability of the *pas* system to stabilise an unstable heterologous replicon varies greatly in different *E. coli* host strains. In some strains the *pas* genes were unable to stabilise a heterologous replicon while in others a 100-fold increase in stabilisation was achieved. While PasB was highly toxic in all strains tested, the ability of PasA and PasC to moderate these toxic effects varied. Strains in which the *pas* system confers the greatest degree of plasmid stabilisation are those that were most sensitive to the PasB toxin even in the presence of both PasA and PasC. A model is proposed in which the degree of plasmid stabilisation achieved in a particular strain, appeared to rely on an equilibrium in PasA and PasB interaction. In strains in which the proteins bind too strongly the toxin may not be released and thus be less able to kill plasmid free cells. In strains in which the proteins interact too weakly, the plasmid containing cells may be subject to inappropriate toxicity even in the presence of PasA and the plasmid will be lost from a population.

Chapter 1

General Introduction

The definition of a plasmid includes the requirement that it is dispensable to the host. It is only in specialised environments that plasmids are beneficial to the host by providing advantageous functions which include antibiotic resistance and degradative metabolic pathways. Under non-selective growth conditions plasmids are a metabolic burden on the cell and it is beneficial for a cell to lose plasmids. This, however, does not always happen. Despite the inability of certain plasmids to confer a selective advantage on cells, they are stably maintained in a population. Assuming that plasmids are randomly segregated at cell division, the probability that a daughter cell inherits a plasmid is expressed by the relationship $P_0 = 2^{1-n}$ where P_0 is the probability of a plasmid free cell arising per division and n is the number of copies of the plasmid at division (Williams and Thomas, 1992). In cells with a copy number of ≥ 10 this gives a probability below 1 in 10^6 which is approaching the level at which mutations affecting plasmid maintenance occur (Nordström and Austin, 1989). Plasmids with copy numbers of greater that 10 are theoretically stable but in practice factors such as concatenation and uneven plasmid distribution within the cell cause worse than random segregation and stability is only achieved at higher copy numbers. Experimental evidence, however, shows that even plasmids with copy numbers below 10 are stably maintained. This implies that these plasmids must have mechanisms which ensure stable inheritance.

Stability can be achieved by increasing the plasmid copy number or interfering with the replication mechanism of the plasmids. For increased stability to be attributed to a stability mechanism rather than disruption of normal plasmid regulation, two requirements must be met. The system must not alter the copy number of the plasmid and it must be dispensable for normal functioning of the plasmid. Stable inheritance mechanisms function either by active partition systems which ensure each daughter cell inherits at least one plasmid or systems that increase the number of segregation units available for random distribution or systems that ensure plasmid free progeny are killed or retarded. Active partition systems such as the *sop* of the F plasmid (Ogura and Hiraga, 1983a) control a dynamic process where plasmids are

physically removed to separate intra-cellular locations which, once the cells have divided, will become part of different daughter cells. Computer simulation of R1 partition has indicated that it is unlikely that all plasmids are evenly distributed between both daughter cells (Rosenfeld and Grover, 1993). Either the two plasmids produced by the most recent round of replication are partitioned to each daughter cell or single randomly chosen plasmids are actively partitioned to each daughter cell. In both cases the remaining plasmids are randomly distributed. Multimer resolution systems such as *loxP*/Cre of P1 (Austin *et al.*, 1981) ensure that plasmids are physically separate and thus able to be randomly distributed in proportion to their copy number. Systems that kill or retard plasmid free segregants like the *ccd* system of F (Ogura and Hiraga, 1983b) are not true stability systems in that they do not ensure random or better than random distribution of plasmids. By killing plasmid free cells they maintain the ratio of plasmid free to plasmid containing cells which is the measure of stability.

Plasmids can be viewed as ancillary genetic elements that only provide benefit to their host. Partition systems, however, give plasmids the ability to act as parasites. By ensuring their own selection irrespective of the benefit they confer, plasmids containing stabilisation systems switch between mutualism and parasitism. In environments where the plasmid provides benefit to the cell, the genetic load of the plasmid is balanced by the benefit. When the environment does not select for the plasmid or when the plasmid has no beneficial genes, plasmids containing stabilisation systems exert a genetic load and act as parasites. Without stabilisation systems the plasmids conferring no benefit would be lost as cells containing them would be at a selective disadvantage. Thus stabilisation systems in their many forms exert influence far beyond being mere dispensable genetic elements.

1.1 Multimer resolution systems

Multimer resolution systems have been shown to stabilise various plasmids including P1, ColE1 and RK2 (Austin *et al.*, 1981; Summers and Sherratt, 1984; Roberts *et al.*, 1990). These systems typically consist of a *cis* acting site and a *trans* acting site-specific resolvase. Multimers are cut and rejoined at the *cis* acting site by the resolvase thus separating molecules interlinked by either homologous recombination or by replication. The role of these systems is to ensure that the number of units available for segregation is maximised and approaches the copy number of the plasmid as closely as possible. P1 encodes the Cre recombinase which acts

on the *loxP* site to resolve dimers produced by *recA*-mediated recombination (Sternberg and Hamilton, 1981). Austin *et al.*, (1981) were able to show that not only was the *loxP* site able to prevent the formation of stable composite structures, it was also needed for maintenance fidelity. These properties require not only the presence of *loxP* but also the Cre recombinase.

ColE1 utilises the *E. coli* chromosomally encoded XerC recombinase to act on the *cer* site which is located on a 380 bp *Hpa* II fragment (Summers and Sherratt, 1984; Colloms *et al.*, 1990). Chromosomally encoded genes such as *argR* and *pepA* are also essential for *cer*-mediated recombination. Their exact role in multimer resolution, however, has not been determined (Stirling *et al.*, 1988; Stirling *et al.*, 1989). A variant of the *cer* site which is only 50bp in size has been shown not to require *argA* and *pepA* for multimer resolution, implying that these genes serve an ancillary role (Summers, 1989). Summers and Sherratt (1984), demonstrated that the XerC/*cer* system could stabilise ColE1, however, Roberts and Helinski (1992) were unable to show increased stabilisation of mini RK2 plasmids by the *cer* site even though it did resolve multimers. More recently Patient and Summers (1993) have shown that the 70 nucleotide Rcd transcript which is expressed from the P_{cer} promoter located at the *cer* site acts to block cell division and is essential for *cer*-mediated plasmid stability but not for multimer resolution. It appears that the Rcd transcript blocks protein synthesis and hence cell division so that multimers may be resolved by the XerC resolvase. Once the multimers have been resolved transcription ceases from the P_{cer} promoter and normal growth is allowed to continue.

RK2 which is indistinguishable from RP4, RP1, R18 and R68 (Burkardt *et al.*, 1979; Stokes *et al.*, 1981) was shown to be partially stabilised by a *cis* acting *mrs* site and the *parA* gene (Roberts *et al.*, 1990). The *parA* gene which has considerable homology to the resolvases of the *Tn3* family (Gerlitz *et al.*, 1990) is part of the *parCBA* operon. Two forms of ParA are produced, the 24.2 kDa and the 22.7 kDa polypeptides, which are each capable of site-specific recombination at similar rates (Eberl *et al.*, 1994). The *mrs* site spans 110 bp between two divergent promoters one of which drives the *parCBA* operon and the other the *parDE* operon. ParA binds to three sites of dyad symmetry in the *mrs* region (Eberl *et al.*, 1994) and inhibits its own production 2.8-fold (Eberl *et al.*, 1992; Davis *et al.*, 1992a). Repression of *parCBA* is presumably due to bound ParA preventing RNA polymerase binding to the promoter region of the *parCBA* operon. Based on homology to *Tn3* type *res* sites the crossover point occurs in the centre of the 10bp palindromic spacer region between the invert repeats which form the ParA binding site 1. Site 1 overlaps the -10 region and RBS of *parCBA* (Eberl *et al.*, 1994). Recent studies by Sobecky *et al.* (1996) indicate that RK2 stabilisation relies on at least two other mechanisms which are encoded in the same region as the

multimer resolution system, with the multimer resolution system contributing insignificantly to plasmid stability (Sobecky *et al.*, 1996).

Transposons such as *Tn21*, *Tn1000* and *Tn3* encode site-specific recombination systems and as these are frequently found on plasmids it is probable that they may also stabilise plasmids by resolving multimers (Nordström and Austin, 1989). The degree to which plasmids are stabilised by resolution of multimers is small and usually supplemental to other systems such as active partition or post-segregational killing. The different results obtained by various researchers only serves to stress how poor a method of plasmid stabilisation multimer resolution is. It is interesting to note that few studies prove multimer resolution stabilises plasmids, more often there is the assumption that because the system is capable of multimer resolution it must stabilise the plasmid.

1.2 Active partition systems

Unlike multimer resolution systems which can at best achieve exactly random segregation, active partition systems achieve a better than random distribution. These systems encode two proteins and a *cis* acting site which is thought to act in a manner similar to a eukaryotic centromere (reviewed by Williams and Thomas, 1992).

1.2.1 *sopABC* of F

The *sop* locus of F encodes the *trans* acting SopA and SopB proteins which are 44 and 38 kDa, respectively, and the *cis* acting *sopC* (*incD*) site (Ogura and Hiraga, 1983a; Austin and Wierzbecki, 1983). The *sopC* region has 12 direct repeats of a 43 bp sequence and each of the repeats contains a pair of 7 bp inverted repeats (Helsberg and Eichenlaub, 1986). Lane *et al.* (1987) found that mini-F plasmids can still be stably maintained with as few as 2 of the repeats. These results are in disagreement with those of Ogura and Hiraga (1983a) and Mori *et al.* (1986) who found that stability was proportional to the number of repeats present. No adequate reason can be found for this discrepancy. When these repeats are carried on a high copy

number plasmid they give rise to IncD type incompatibility (Mori *et al.*, 1986). The SopA protein binds to the repeat 5′-CTTTGC-3′ which occurs four times in the promoter region of the *sopABC* operon. This binding is enhanced by SopB (Mori *et al.*, 1989). The binding of a SopA and SopB complex to the promoter of *sopABC* effectively autoregulates the operon (Hiraga, 1992; Mori *et al.*, 1989). Figure 1.1 shows key components of a model for *sopABC*-mediated partition.

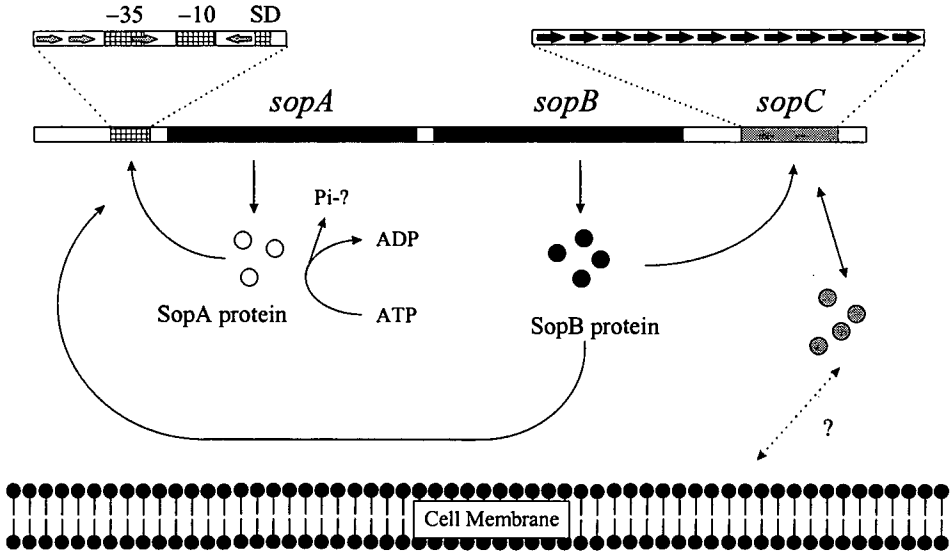


Figure 1.1 Overview of Sop gene function adapted from Hiraga, 1992. Black boxes indicate genes, grey boxes cis-acting regions and hatched boxes promoter regions. SopA protein is indicated by ○ and Sop B protein ●, unidentified host proteins ◍, 43bp repeats ➡ and SopA binding site 5′-CTTTGC-3 by ⇨.

The MinD which has been shown to inhibit septation shows homology to SopA, indicating that SopA may have a role in addition to autoregulation (Motallebi-Veshareh *et al.*, 1990; de Boer *et al.*, 1989). SopA shows Mg^{2+} dependent ATPase activity which is stimulated by the presence of DNA and further stimulated by SopB. SopB does not stimulate the ATPase activity in the absence of DNA (Watanabe *et al.*, 1992). The target of phosphorylation by SopA is as yet unknown. It does, however, not appear to be SopB, as a phosphorylated form of SopB has not been detected (Watanabe *et al.*, 1992). SopB binds as a dimer to the 7bp inverted repeats of *sopC* utilising mainly its C-terminal domain (Mori *et al.*, 1989; Hanai *et al.*, 1996). High

intracellular levels of SopB cause repression of genes adjacent to the *sopC* element and gives rise to the IncG phenotype. This appears to be due to the genes being sequestered in a nucleoprotein complex which limits accessibility to DNA gyrase and DNA adenosine methylase (Lynch and Wang, 1995; Kusukawa *et al.*, 1987). The complex formation requires unidentified host factors, as SopB by itself was incapable of complex formation and did not bind outside the *sopC* region even at high concentrations (Hanai *et al.*, 1996). Deletions of greater than 35 amino acids from the N-terminus of SopB prevents this gene silencing effect, and probably its partition function (Hanai *et al.*, 1996). It is unlikely that SopB is involved in attachment to the chromosome as *mukB* mutants which are deficient in chromosomal partition are capable of partitioning the F plasmid (Ezaki *et al.*, 1991). A more probable mechanism would involve the binding of *sopC*-bound SopB either directly or indirectly to a membrane associated host encoded protein which is involved in allocation of chromosomes to each daughter cell. SopB has been found by Watanabe *et al.* (1989) to sediment with membrane fractions implying a role in membrane attachment. However, the validity of these results is questionable as sedimentation profiles did not exactly match that of the cell membrane (Hiraga, 1992). The plasmid QpH1 from *Coxiella burnetii* encodes two polypeptides with considerable homology to the SopA and SopB proteins of F (Lin and Mallavia, 1994). These proteins are required for plasmid stabilisation. Tn*phoA* mutagenesis has indicated that the SopA analog QsopA is membrane associated. Further studies are needed to clarify the exact cellular location of SopA and SopB and their association with host proteins.

1.2.2 *parABS* of P1

The prophage of bacteriophage P1 is stably maintained in *E. coli* as a plasmid with a copy number of one or two (Prentki *et al.*, 1977). Like the Sop region of the F plasmid the 2.1 kb *par* region consists of two proteins, ParA (44 kDa) and ParB (38 kDa) and a cis-acting *incB/parS* region which is analogous to the *sopC* of F (Austin and Abeles, 1985; Abeles *et al.*, 1985). The *parS* site consists of an 84 bp AT-rich region which includes a 13 bp palindrome as well as a IHF binding site (Funnell, 1988a; Hiraga, 1992). The minimal partition proficient *parS* site can be reduced to the 22bp to the right of the IHF site. This, however, creates a new incompatibility

group, most likely due to a change in conformation induced by the loss of the IHF site (Martin *et al.*, 1987). ParB binds to a complicated non-symmetrical region to the right of the IHF site and at a far lower affinity to a partial copy of the sequence 5′-TCGCCA-3′ on the left (Funnell and Gagnier, 1993). Binding of ParB, and IHF which requires a supercoiled substrate, is co-operative, with IHF stimulating ParB binding by 10000-fold (Funnell, 1991). The same face of the DNA on either side of the IHF site is recognised (Funnell and Gagnier, 1993). Binding motifs have been identified but it is unclear if ParB recognises a linear DNA sequence or tertiary structure (Funnell and Gagnier, 1993; Funnell and Gagnier, 1994). ParB, which occurs as a dimer in dilute solutions, binds to the stronger site on the right. Binding by IHF then bends the DNA, allowing the ParB to come into contact with the lower affinity site on the left and thus wrapping the *parS* DNA around the ParB and IHF core (Funnell, 1991; Funnell and Gagnier, 1993). This protein-DNA complex shows no ParA involvement (Funnell, 1988b). Like SopA, ParA negatively autoregulates its own production 5-fold, and ParA and ParB together cause repression to increase to 50-fold (Abeles *et al.*, 1985; Funnell, 1988b). While ParA binding to its own promoter has been detected, no evidence has been found for ParB binding to the promoter (Davis *et al.*, 1992b). It is probable that ParB binds to ParA resulting in increased affinity of the complex or its ParA component for the promoter of the *parAB* operon. The site-specific binding of ParA is ATP dependent. This is in contrast to SopA binding to *sopC* which does not require ATP (Davis *et al.*, 1992b; Mori *et al.*, 1989). ParA ATPase activity is strongly stimulated by ParB. ParB is, however, not the target for phosphorylation (Davis *et al.*, 1992b). It is possible that the ATPase activity of ParA is also required for partition steps other than *parAB* autoregulation. Identification of the target of ParA-mediated phosphorylation will assist in determining at which other sites ParA acts. The steps involved in partition that occur after ParB-IHF-*parS* complex formation, are relatively unstudied but because of the similarity of par to the *sop* locus of F, it is probable that a substantially similar set of events occur.

1.2.3 *parA* of R1

The *parA* locus of R1 and *stb* locus of closely related plasmid NR1 (R100), although basically similar to *sop* of F and *parS* of P1, have a different layout of components. The *cis*-acting *parC* site of R1 lies within the promoter region of the two *trans*-acting proteins ParM(R1)/StbA(NR1) (36 kDa) and ParR/StbB (13 kDa) (Gerdes and Molin, 1986c; Dam and

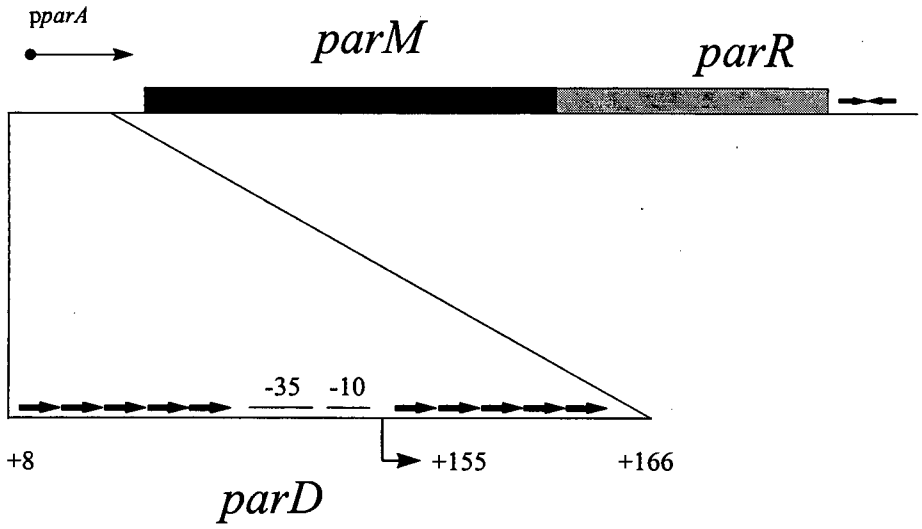


Figure 1.2 Genetic organisation of the *parA* locus. The *parM* and *parR* genes are transcribed from the *parA* promoter, which is located in within the *parC* site. The transcriptional terminator is represented by inverted arrows. The detailed map of the *parC* region shows the relative position of potential -35 and -10 sequences of the *parA* promoter, and the 10 11 bp direct repeats. (adapted from Dam and Gerdes, 1994)

Gerdes, 1994). The *parC* site covers 160 bp and consists of 10 direct repeats with the consensus sequence 5′-AAA(A/C)(A/C)(A/C)A(A/C)(A/C)CC-3′, the repeats are evenly divided with five upstream of the -35 region and five downstream of the -10 region (Dam and Gerdes, 1994). At high copy number *parC* exerts incompatibility which is dependent on the number of repeats present. Unlike *sopC* of F and *parS* of P1, *parC* exerts incompatibility only at artificially high copy numbers (Dam and Gerdes, 1994). ParR binds to *parC* and represses its own and ParM production greater than 100-fold. ParM shows no regulatory effects.

A region of at least 108 bp situated just downstream of the *parMR* promoter acts as a strong transcriptional activator of *parMR* (Jensen *et al.*, 1994). Strong expression from the *parMR* promoter coupled with equally strong ParR repression of the promoter ensures that the *parA* locus is repressed under normal conditions. Tight regulation is required as overproduction of ParR causes instability. This parallels the *incG* instability in P1 caused by overproduction of ParB. The observation that constructs deficient in *parR* yet are $parC^+$, $parM^+$ are not viable (Dam and Gerdes, 1994; Min *et al.*, 1988), taken together with the autoregulatory role of ParR, implies that over-expression of ParM is lethal. Dam and Gerdes (1994), however, also found

that overproduction of ParM causes a reduction in growth rate of the host cells but does not seem to be lethal. The discrepancy may be due to the presence of the transcriptional enhancer element without its cognate promoter in the fragments used by Dam and Gerdes to test ParM overproduction. ParM shows sequence similarity to a group of ATPases which include a number of cell cycle proteins and actin (Bork *et al.*, 1992) and it is possible that this may reflect an involvement in intracellular plasmid movement. Attachment to the outer membrane has been detected in par^+ but not par^- R1 derivatives affirming the role of membrane binding in active partition (Gustafsson *et al.*, 1983). Differences in results between studies on the NR1 *stb* locus and the R1 *par* locus have been observed (Jensen *et al.*, 1994; Dam and Gerdes, 1994). However, as pointed out by Jensen *et al.* (1994), these can be ascribed to differences in experimental design and do not reflect differences in the partition systems.

1.2.4 *parCBA* of RK2

The *par/mrs* locus of RK2/RP4 forms a hybrid system that is part multimer resolution system (reviewed in section 1.1) and part active partition system. The *cis*-acting site lies between the divergent promoters for the *parCBA* operon and the *parDE* operon and consists of approximately 140 bp with two direct repeats of six and seven bases and two invert repeats of six and eight bases, respectively (Roberts *et al.*, 1990). Unlike a typical active partition system *par/mrs* has 3 genes. The *parC* gene encodes a 10 kDa protein, the parB an 18 kDa protein, and parA a 24 kDa protein (Gerlitz *et al.*, 1990). The role of the ParA site-specific recombinase has been described in section 1.1. Gerlitz *et al.* (1990) were able to show that plasmids which are recombination proficient but lack parts of parC and parB were not fully stabilised. The *parB* gene which is essential for full stability has considerable homology to an extra-cellular nuclease of *Staphlococcus aureus* (Shortle, 1983) and is able to exhibit nuclease activity (Eberl *et al.*, 1994). No role has been identified for *parC*.

1.2.5 Active partition models

All active partition models incorporate three main stages, (a) plasmid-protein interaction, (b) plasmid pairing and attachment to a 'movement' site and (c) finally separation and relocation within the cell. Plasmid coded proteins bind to their *cis*-acting site which in the case of R1

(NR1) and RK2 (RP4) is also the site of autoregulation. For F and P1, sites of autoregulation and partition are separate. Protein binding can occur either to the replicating plasmid as soon as a new par site is synthesised or to free plasmid (Figure 1.3). The free plasmid now either binds to another plasmid and subsequently to a movement site or it can bind to the movement site independently. When proteins bind to the partition sites of the newly synthesised daughter plasmids during replication, the daughter plasmids bind to each other immediately and form a linked pair (Austin and Nordström, 1990; Martin *et al.*, 1987; Nordström and Austin, 1989).

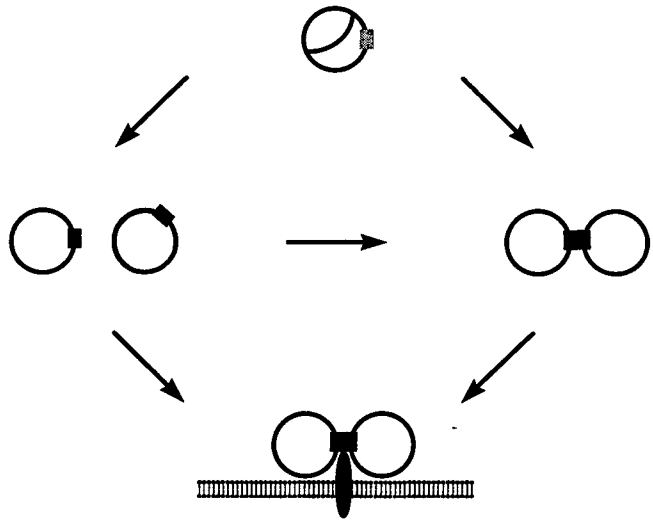


Figure 1.3 Model for plasmid binding to a movement site. The grey box indicates vacant par binding site, black boxes indicate occupied par binding sites and the black oval, the movement site. Adapted from Nordström and Austin, 1990

Computer simulations by Rosenfeld and Grover (1993) suggest that in R1 at least one randomly chosen pair of plasmids are actively partitioned with the remainder of the population being randomly segregated. This model, however, requires a single unique receptor, a situation which is experimentally difficult to prove and hitherto unheard of. Equipartition (plasmids equally partitioned to both daughter cells) is another option and implies that the number of receptors is not limiting. As the incompatibility caused by the *cis*-acting site of R1, when present at high copy number, cannot be relieved by over-expression of the partition proteins ParR and ParM, it is probable that movement sites are limiting in this case at least (Jensen *et al.*, 1994). Experimental evidence suggests that a protein-DNA complex aided by host factors binds to a membrane site(s) (Gustafsson *et al.*, 1983, Lin and Mallavia, 1994). This would

require exact targeting to the plane of septum formation and orientation around the plane so that each plasmid of the pair would end up on opposite sides of the septum.

Hemi-methylation has been implicated in the timing of both chromosomal and plasmid replication (Russell and Zinder 1987; Campbell and Kleckner 1990; Malki *et al.*, 1992; Abeles *et al.*, 1993). Newly synthesised chromosomal DNA will contain hemi-methylated DNA. The DNA is sequestered to the membrane where a slow process of methylation takes place till the DNA is fully methylated, at which point it is released and is again available for replication. This process limits the rate of initiation of replication as well as ensuring that the chromosome is membrane bound, which could allow for chromosome segregation by a membrane associated mechanism. The role of hemi-methylation in plasmid segregation and membrane attachment deserves a considerable amount of research as there is a significant probability that hemi-methylation may serve a similar role in some plasmid systems. The ParA protein of F has been found to have some similarity to the MinD genes which together with MinC can inhibit septation (de Boer *et al.*, 1989). This similarity may indicate that they both act at the plane of septation. Another interesting model proposed by Williams and Thomas (1992) as well as

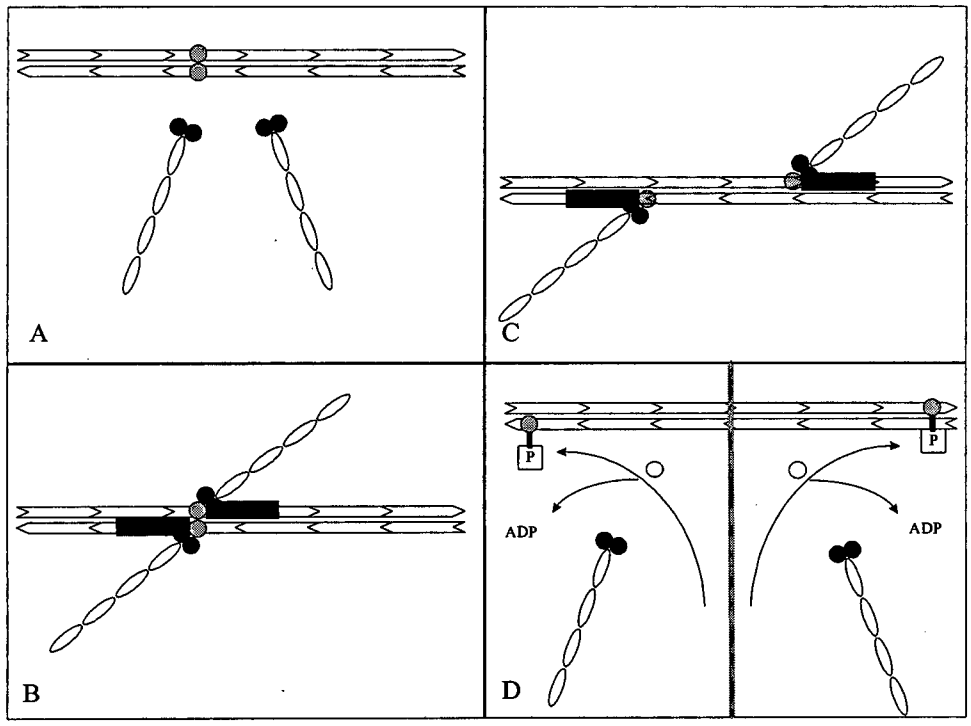


Figure 1.4 Anti-parallel microfilament model of partition adapted from Williams and Thomas (1992). The symbols are, ◍ filament binding protein, ● a par binding protein, ▬ a motor complex and ⊂⊃ a unit of the microfilaments. The grey vertical line indicates plane of septum formation and ○○○○ plasmid DNA.

Hiraga *et al.* (1992) is that plasmids bind to pairs of anti-parallel microfilaments. Once bound to the filaments, the pairs separate and move to opposite poles of the cell. The initial binding can either be as a pair or as individuals with binding of the second plasmid triggering the movement (illustrated in Figure 1.4 A, B and C). The discovery that *E. coli mukB* (mutants which are unable to segregate their chromosomes) are deficient in what appears to be a mechanochemical enzyme raises the possibility that at similar type of enzyme system may move the plasmid along the microfilaments (reviewed in Hiraga, 1992; Niki *et al.*, 1991). The recent discovery of such filaments along the long axis of cells, either under or in the cell envelope, greatly strengthens this model (Hiraga *et al.*, 1996). The ATPase activity of the A proteins of P1 (ParA) and F (SopA) may be involved in either the unpairing of the plasmids so that they can move apart or in the release of separated plasmids from the filaments (as shown in Figure 1.4 D). While much is known about the early stages of partition most of the studies on the later stages of partition have been limited to speculation based on similarities between proteins required for partition and proteins of known function.

1.3 Post-segregational killer systems

Unlike multimer resolution and active partition systems which try to ensure random or better than random distribution, respectively, post-segregational killer systems do not alter the distribution of plasmids to daughter cells. By killing or retarding the growth of plasmid free cells they ensure that plasmid containing cells predominate in a population. The inhibition of killing is either effected by RNA:RNA or protein:protein interactions and activation of the poison relies on different rates of processing or decay for the poison protein or poison mRNA and the antidote protein or antisense RNA.

1.3.1 RNA-based poison - antidote systems

RNA-based post-segregational killing relies on the poison mRNA folding back on itself to prevent its own translation. Processing of the mRNA produces an mRNA which does not fold back and is a translationally active, in plasmid containing cells this active mRNA binds to an antidote in the form of antisense RNA and is rapidly degraded. This ensures that there is no active mRNA capable of being translated into poison present in plasmid containing cells. In

plasmid free cells no new antisense RNA is produced and rapid degradation of the existing antisense RNA ensures that poison mRNA is translated into the poison and the cell is subsequently killed. The subject of RNA-based poison - antidote systems has been reviewed by Gerdes *et al.*, 1990a.

1.3.1.1 *hok/sok* of R1

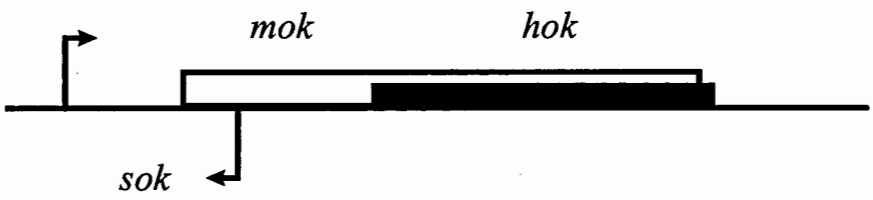


Figure 1.5 Genetic organisation of the *parB* locus of plasmid R1, from Gerdes *et al.*, (1990a). Transcriptional initiation signals are indicated by arrows, the *mok* ORF as an open box and the *hok* ORF as a closed box.

The *parB* locus of R1 encodes 2 overlapping ORFs, *mok* and *hok*, as well as an antisense RNA, *sok* (Figure 1.5), and is capable of stabilising a wide range of plasmids (Gerdes *et al.*, 1985; Gerdes *et al.*, 1996a; reviewed by Gerdes *et al.*, 1990a). The *mok* and *hok* genes are transcribed to give 2 transcripts, one of 441 nt and one of 398 nt (Gerdes *et al.*, 1990b; Thisted *et al.*, 1994a). The 3′ ends of these transcripts fold back and pair with the translational initiation region (TIR) of *mok* to produce inactive mRNA, this process is called fold back inhibition (FBI) (Thisted *et al.*, 1995). These transcripts are very stable (t½ > 15 minutes) due the protection of the 3′ end by FBI and the 5′-end by the formation of a small stem loop (Thisted *et al.*, 1995). The *sok* RNA is also unable to bind to the *mok* TIR as it is sequestered by the FBI structure. The 398 nt mRNA is slowly processed at its 3′ end by either ribonuclease II or polynucleotide phosphorylase to produce a stable 361 nt active transcript (Thisted *et al.*, 1995). The *sok* gene produces an unstable antisense transcript of 67 nt which overlaps the *mok/hok* TIR (Thisted *et al.*, 1994b). In plasmid containing cells the *sok* RNA binds to the TIR of the 361 nt *mok* mRNA to form a RNA:RNA hybrid. This inactive hybrid and hybrids formed with full length *mok/sok* mRNAs are rapidly degraded by RNAse III (Gerdes *et al.*, 1992). The processing of the inactive transcripts to active transcripts is the rate limiting step in cell killing.

In plasmid containing cells there is also a slow rate of *sok* binding to the full length *mok/hok* mRNA. This binding also competes with the 3′-end of *mok/sok* mRNA for binding to the TIR of *mok* (Thisted *et al.*, 1994b). Both the slow rate of processing and slow *sok* hybridisation to full length mRNAs ensure that there is not a wasteful high turnover of RNA which would occur if active mRNA was transcribed directly (this process is shown in Figure 1.6).

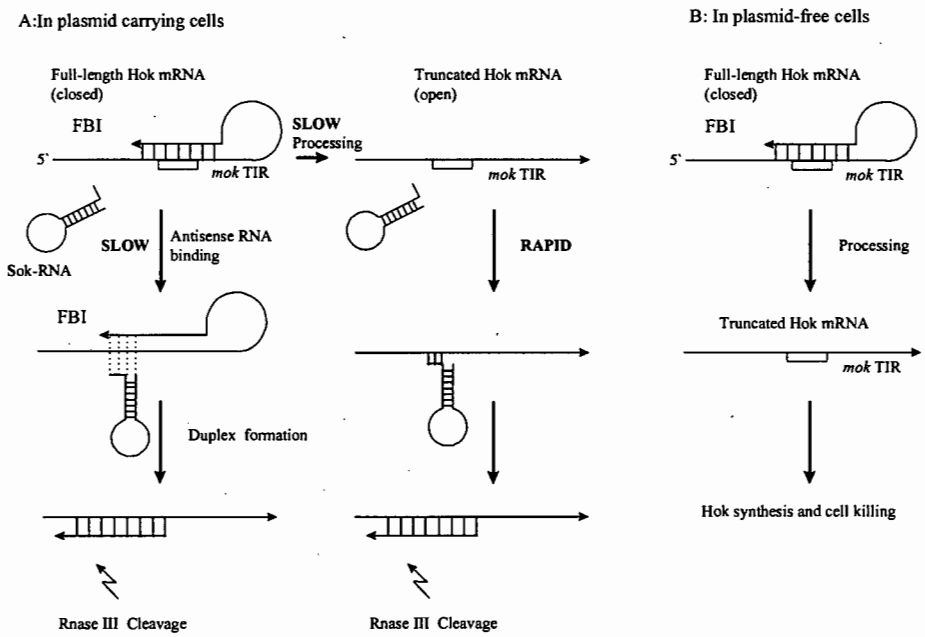


Figure 1.6 Schematic model explaining activation of translation of Hok mRNA. (from Thisted *et al.*, 1994a)

In plasmid-free cells *sok* RNA rapidly degrades (t½ = 30 seconds) and is not replenished (Gerdes *et al.*, 1990b), the stable *mok/sok* mRNA is processed to the active form and translated into a 52 amino acid protein Hok (Gerdes *et al.*, 1986b). Translation of *hok* is coupled to the translation of *mok*, probably due to the disruption of a stem-loop structure which sequesters the TIR of *hok* (Thisted *et al.*, 1994a). The role, if any, of the 70 amino acid (Gerdes *et al.*, 1990a) Mok protein is as yet unknown. It is, however, probable that it serves some other role than merely regulating *hok* expression. The Hok protein which associates with the cell membrane has strong similarity to the *relF* gene of the *E. coli* chromosome and both cause the disruption of trans-membrane potential and cell death resulting in the formation of 'ghost' cells (Gerdes *et al.*, 1986b). Despite the 40% homology between the RelF and Hok, the *relB* operon of which *relF* is part was unable to stabilise an unstable plasmid (Gerdes *et al.*, 1986b). The *relF* gene

lacks a clearly defined overlapping gene analogous to the mok of R1 as well as an antisense promoter. The lack of a regulatory mechanism may explain inability of the *relB* operon to stabilise an OriC plasmid (section 1.3.1.2 discusses the *relB* operon). A side effect of the regulatory mechanism of the *parB* locus is that it also serves to limit the extent and severity of phage infections. When T4 phage infects a cell it halts all host transcription within 3 to 4 minutes (Rabussay, 1983). Upon T4 infection, cells containing the *parB* locus will no longer be able to synthesise new *sok* RNA and the existing pool of *sok* will rapidly decay. With no inhibitor, active *hok* mRNA will be translated and the Hok protein will cause cell death. This killing of infected cells reduces burst size by 40% as well as increasing the eclipse time and latent period (Pecota and Wood, 1996). The benefits of phage induced death are felt only at a population level as are the benefits of all forms of post-segregational killing. The *relB* operon may moderate the affects of phage infection, and this may explain the prevalence of genes homologous to *hok* on bacterial chromosomes.

1.3.1.2 *flm* and *srnB* of F, *pnd* of R483 and R16, *relB* and *gef* of the *E. coli* chromosome

The *hok* killer gene family is a group of plasmid and chromosomally encoded genes who share a common regulatory mechanism and near identical layout and encode similar proteins. All members of the *hok* gene family encode toxic peptides of 49 to 52 amino acids with 20 amino acids being conserved in all peptides (Gerdes *et al.*, 1990a).

Table 1.1 Comparison of *hok* killer gene systems. Adapted from Gerdes *et al.* (1990a).

Location	Locus	Toxin	Antisense	Coupled peptide
R1	*hok/sok*	Hok (52 aa)	*sok*	Mok
F	*flm*	FlmA (52 aa)	*flmB*	FlmC
F	*snrB*	SrnB' (49aa)	*srnC*	SrnB
R16	*pndA*	PndA (50 aa)	*pndB*	PndC
R483	*pndA*	PndA (50 aa)	*pndB*	PndC
E.coli chromosome	*gef*	Gef (50 aa)	*sof*	ORF69
E.coli chromosome	*relB* operon	RelF (51 aa)	?	?

The *flm* system of F detected by its ability to stabilise unstable plasmids is nearly identical to the *hok/sok* system with the FlmA protein differing from Hok at only one position (Loh *et al.*,

1986; Gerdes *et al.*, 1990a). The Gef and RelF toxins vary at only 10 amino acids while the PndA toxins vary by only 3 amino acids (Gerdes *et al.*, 1990a). In addition to the sequence similarity shown by these systems they have all been shown to be capable of causing cell death (Gerdes *et al.*, 1985; Loh *et al.*, 1986; Ohnishi and Schlessinger, 1972; Ohnishi and Akimoto, 1980; Ono *et al.*, 1987; Gerdes *et al.*, 1986b; Poulsen *et al.*, 1989). The mechanisms of regulation of toxin expression have been studied in detail for *hok/sok*, *snrB* and *pnd* loci and all are regulated by antisense binding and FBI (Thisted *et al.*, 1994a; Gerdes *et al.*, 1992). The *flmA* gene is predicted to be more stable than *flmB* and appears to have its expression blocked by *flmB,* and in all probability shares a common regulatory mechanism to *hok/sok*, *srnB* and *pnd* (Loh *et al.*, 1988). Translation of the *gef* gene is coupled to the upstream ORF69 which is regulated by *sof* mRNA of at least 61 nucleotides (Poulsen *et al.*, 1991). The regulatory mechanism of the *relB* locus is less clear. While ORF60 could be transcriptionally coupled to *relB* it lacks a recognisable RBS and an ATG codon, it is thus more probable that the *relE* gene fulfils this function (Poulsen *et al.*, 1991). Plasmid stabilisation may not be the sole role of *hok* type genes. A large number of conjugative plasmids carry *flm* and *hok* homologous loci on their transfer leading regions and it is possible that these regions may serve to ensure effective conjugation. As the leading strand enters a cell both sense and anti-sense are transcribed and should the plasmid manage to establish itself, continued transcription will ensure that no toxin is produced. If the plasmid is destroyed by the recipient's restriction-modification system, the antisense RNA will rapidly decay and slow processing of the toxin mRNA will give rise to active toxin which will rapidly kill the recipient. It seems probable that *hok* type killer systems may also ensure effective spread of promiscuous plasmids in a population. This mechanism would be effective only for a population not as is the case with the anti-restriction systems such as the *ardA* gene family which act to ensure survival of the individual plasmid (Read *et al.*, 1992). The roles of plasmid stability and enhancement of promiscuous spread do not account for the presence of *gef* and *relB* on the *E. coli* chromosome. The *relB* operon was unable to stabilise an unstable *oriC* mini-chromosome and the questionable need for a killer type chromosomal maintenance system, since chromosome free cells will not survive, mitigates against a role in chromosome maintenance (Gerdes *et al.*, 1986b). The containment of phage infections as shown by the *hok/sok* locus (see section 1.3.1.1) is also applicable to host encoded genes but is limited to a small number of phages which disrupt host transcription (Pecota and Wood, 1996). At present no role for chromosomally located killer genes is readily apparent.

1.3.2 Protein-based poison - antidote systems

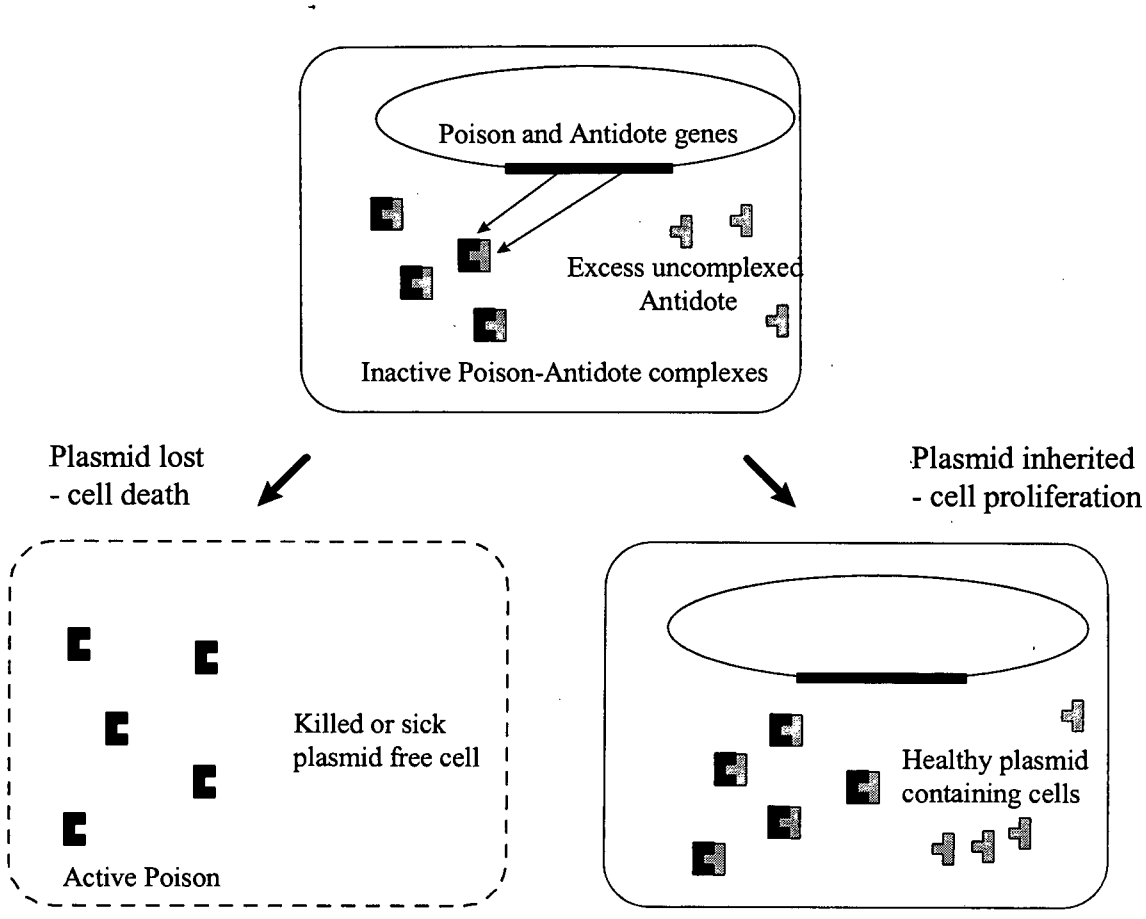


Figure 1.7 Overview of proteic killer systems activation.

Protein-based poison - antidote systems which are also known as proteic plasmid stabilisation systems appear to be simpler than RNA-based systems in their regulation. The toxic protein is inhibited by a less stable antidote. In plasmid free cells this antidote cannot be replaced when it decays and the active poison is then able to kill the cell. The cartoon in Figure 1.7 illustrates this principal. The subject of proteic stabilisation systems has been reviewed in Jensen and Gerdes, 1995a. Proteic killer systems show a wide range of stabilisation abilities and the degree of stabilisation by each system is effected by both growth medium and host strain (Jensen *et al.*, 1995; Roberts *et al.*, 1993; Roberts and Helinski, 1992; Tsuchimoto and Ohtsubo, 1989; personal observations). These differences suggest that despite an apparently simple mechanism, proteic killer systems are perhaps more complex than at first appears to be the case.

1.3.2.1 *ccd* of F

The *ccd* stability system was not initially thought to involve post-segregational killing, but the coupling of cell division to plasmid copy number (Ogura and Hiraga, 1983b). Later studies suggested that *ccd* was responsible for post-segregational killing and was able to increase the stability of an unstable mini-R1 plasmid 10-fold (Jaffé *et al.*, 1985; Jensen *et al.*, 1995). The *ccd* operon which is also known as the *lynA* locus consists of the *ccdA/LetA/H* (8.7 kDa), *ccdB/LetD/G* (11.7 kDa) and *repD* (27 kDa) genes (Bailone *et al.*, 1984; Bex *et al.*, 1983; Lane *et al.*, 1986; Miki *et al.*, 1984a; Miki *et al.*, 1984b). The operon is negatively regulated by a 69 kDa complex of CcdA and CcdB while the RepD protein which is a site-specific-resolvase has no discernible regulatory effect (Tam and Kline, 1989a; Tam and Kline, 1989b). Neither the CcdA protein nor the CcdB protein by itself has any affect on the control of transcription from the *ccd* operon (Salmon *et al.*, 1994). The complex which is a 4:3 combination of CcdA:CcdB binds at several sites over a region of 113 bp. This region is 5′ to and overlaps the start of the operon. Only the C-terminal 41 amino acids of CcdA are required for antidote activity, the N-terminal portion carries the DNA binding ability (Bernard and Couturier, 1991). Despite being co-regulated, far less CcdB is produced that CcdA, which ensures that CcdB is adequately neutralised (Tam and Kline, 1989a). In daughter cells that inherit a plasmid, the autoregulated promoter will ensure that sufficient poison:antidote complex is maintained in the cells so that there is a large enough intracellular pool of inactive poison to ensure that daughter cells can be killed should they not inherit a copy of the plasmid. The unbound CcdA has a half life of 30 minutes while bound to CcdB it has a half life of 1 hour. CcdB has a half life of 2 hours either bound or unbound (van Melderen *et al.*, 1994). The Lon protease is responsible for degradation of CcdA and in Lon-deficient mutants degradation of CcdA and CcdB occurs at the same rate and consequently the ccd system is unable to stabilise unstable plasmids in these mutants (van Melderen *et al.*, 1994). As CcdA (both bound and unbound) decays, the CcdA:CcdB complex breaks down and is no longer able to repress the *ccd* operon. This in turn causes the production of more CcdA and CcdB which replenishes the intracellular pool and shuts off production of the proteins. In plasmid free cells the decaying proteins are not replaced and CcdB becomes active and binds to the GyrA subunit of DNA gyrase in a 1:1 ratio, which inactivates the gyrase (Bernard and Couturier, 1992; Maki *et al.*, 1996; Miki *et al.*, 1992). The inactivation of gyrase is reversible by the addition of CcdA which has a higher affinity for CcdB than does GyrA (Maki *et al.*, 1996). This ensures that the cell is able to survive minor

regulatory imbalances that occur when the CcdA level drops and some unbound CcdB is able to transiently bind GyrA. The inactivation of gyrase causes a decrease in supercoiling and induction of the SOS response which leads to cell filamentation (Maki *et al.*, 1992). Induction of the SOS response is a side-effect of CcdB action and not the cause of cell killing (Jaffé *et al.*, 1985). Cell killing is most probably due to the formation of an inactive gyrase/DNA complex which blocks replication forks. This is similar to the mechanism of action of quinolone drugs (Willmott *et al.*, 1994). Blocking of replication forks leads to an accumulation of regions of single stranded DNA which triggers the SOS response. A further 'knock-on' effect is that the RecA protease, once activated, can cleave the Lambda cI repressor causing phage induction (Ruiz-Echevarría *et al.*, 1991b). Chromosomal mutants resistant to CcdB fall into 2 main classes. The first are mutants in the GyrA subunit which are thought to prevent CcdB binding. The second are mutants in heat shock proteins which are thought to act as chaperonins ensuring that CcdB is correctly folded (Bernard and Couturier, 1992; Miki *et al.*, 1988; Miki *et al.*, 1992; van Dyk *et al.*, 1989). Recently it has been shown that a mutation in the *zfiA* gene (*csrA*) can reverse the effects of mutations in the *tldD* (*pmbA*), *tldE* and *groE* genes (Murayama *et al.*, 1996). The *tldD*, *tldE* and *groE* gene products have been proposed to stimulate interaction between CcdB and the GyrA subunit. The *zfiA* gene product which has also been implicated in carbon storage regulation is thought to inhibit the interaction of CcdB with the GyrA subunit while the *tldD*, *tldE* and *groE* gene products counteract the inhibition. This mechanism of host protein involvement is at present speculative.

1.3.2.2 *parDE* of RK2

Of the well studied proteic plasmid stabilisation systems the parDE system of RK2 is the most effective, stabilising a R1 replicon between 100 times in rich medium and greater than 500 times in supplemented minimal medium (Jensen *et al.*, 1995). The *parDE* genes are located on a 790 bp fragment adjacent to the *parCBA* operon described in section 1.2.4 but are transcribed in the opposite direction (Roberts and Helinski, 1992). While the *parCBA* operon seems to provide efficient stabilisation of RK2 at higher copy numbers, the *parDE* operon appears to play a more important role at lower copy numbers. However, no cross-talk has been detected between these systems (Sobecky *et al.*, 1996). The ParD protein (9 kDa) which acts as the antidote is produced in 30-fold excess to the ParE (11.7 kDa) which acts as the poison This is probably as a result of the TGG start codon of ParE (Johnson *et al.*, 1996; Roberts and

Helinski, 1992; Roberts *et al.*, 1994). The *parDE* promoter is twice as strong as the promoter of *parCBA* and is repressed 5-fold by ParD (Davis *et al.*, 1992a). Unlike the *ccd* system, only the antidote ParD is required for autoregulation (Roberts *et al.*, 1993). The ParD protein exists as a dimer and binds through its C-terminal domain to a 33 bp region overlapping the promoter and transcriptional start of the *parDE* operon (Roberts and Helinski, 1992; Roberts *et al.*, 1993). The operator region contains 2 pairs of inverted repeats which are thought to act as binding sites for ParD (Roberts *et al.*, 1993). ParE shows no regulatory effects on the *parDE* operon (Roberts *et al.*, 1993; Johnson *et al.*, 1996). ParD dimers bind to ParE dimers to form D^2E^2 inactive tetramers. This binding does not change the regulatory effect exerted by ParD or alter the ParD footprint (Johnson *et al.*, 1996). The ParE toxin causes cell filamentation but unlike CcdB this does not appear to be due to inhibition of DNA gyrase since stability is not lost in CcdB insensitive DNA gyrase mutants (Roberts *et al.*, 1994). Lon mutants are also able to stably maintain *parDE* plasmids indicating that Lon is not involved in the differential decay of poison and antidote (Roberts *et al.*, 1996). Although the later steps in cell killing are not well studied, it seems that a mechanism similar to that seen in the *ccd* system will act to ensure that only plasmid free cells are killed.

1.3.2.3 *parD* of R1

The *parD* locus of R1 and the *pem* locus of R100 encode identical proteic killer systems (Bravo *et al.*, 1987; Tsuchimoto *et al.*, 1988). Under normal growth conditions the *parD* operon is inactive (Bravo *et al.*, 1987). As with the *parDE* locus of RK2, it only functions when plasmid copy numbers drop to very low levels (Ruiz-Echevarría *et al.*, 1995b; Sobecky *et al.*, 1996). Unlike the other proteic killer systems, the *parD/pem* locus increases stability by retarding growth, not by killing cells. This approach is at least in part responsible for the very modest stabilisation of 5 to 10-fold exerted by *parD* on a R1 replicon (Jensen *et al.*, 1995). The loci encode the PemI/Kis (9.3 kDa) antidote and the PemK/Kid (11.9 kDa) poison proteins which are transcribed from a single promoter upstream of the *pemI* gene (Tsuchimoto *et al.*, 1988). The promoter contains 2 pairs of inverted repeats to which a complex of PemI and PemK binds co-operatively to repress transcription of the *pem* operon (Tsuchimoto and Ohtsubo, 1993). It is thought that two units of PemI bind to 2 units of PemK through their C-terminal ends. This inactivates the poison (Ruiz-Echevarría *et al.*, 1995a). While PemI which binds the operator sites through its N-terminal end, is capable of auto-repression, this only

achieves 30% - 40% of the effect exerted by the PemI:PemK complex (Ruiz-Echevarría *et al.*, 1991; Ruiz-Echevarría *et al.*, 1995a). This separation of the autoregulatory and inhibitory regions of the antidote molecule is also found in CcdA (Bernard and Couturier, 1991). The promoter produces two transcripts, one of 700 nucleotides which encodes PemK and PemI, and another, generated by a ρ-independent terminator, of 370 nucleotides in length coding only for PemK. This possibly ensures that there is an excess of antidote to poison (Bravo *et al.*, 1988; Ruiz-Echevarría *et al.*, 1991). In plasmid free daughter cells Lon protease rapidly degrades PemI and the active PemK then inactivates DnaB preventing correct primosome assembly and causing retarded growth (Ruiz-Echevarría *et al.*, 1995; Tsuchimoto *et al.*, 1992). The retarding of growth can be more severe in some strains, and in these cases this leads to killing of plasmid free cells (Tsuchimoto and Ohtsubo, 1989). The finding that the concerted action of PemI and PemK can stimulate ColE1 replication 280% clearly indicates that the mechanism of stabilisation used by the *pem* locus is perhaps not as simple as it appears (Ruiz-Echevarría *et al.*, 1995a).

1.3.2.4 *phd/doc* of P1

The *phd/doc* system of plasmid P1 resides on a 700 bp fragment located >20 kb from the origin of replication (Lehnherr *et al.*, 1993). The Doc protein (13.6 kDa) encodes the poison while the Phd protein (8.1 kDa) encodes the antidote which is produced in 14-fold excess compared to the poison. The excess of Phd over Doc is due partly to *doc* containing 16 rare codons while *phd* has only one. Also responsible is the lack of a recognisable ribosome binding site for *doc*, which lies down stream of *phd*, and overlaps the stop codon of *phd*. Downstream of these two genes lies a 66-codon ORF which is dispensable to the system It is, however, not clear what role this protein serves or if it forms part of the *phd/doc* operon. While the mechanism of *phd/doc* is thought to function along similar lines to other proteic killer systems, it uses ClpXP protease to selectively degrade Phd thus ensuring Doc activation in plasmid free cells (Lehnherr and Yarmolinsky, 1995). The target of Doc is at present unknown. However, it is unlikely to be the GyrA subunit of DNA gyrase since no SOS induction by Doc has been observed and cell filamentation is not a characteristic symptom of P1 plasmid loss (Lehnherr *et al.*, 1993)

1.3.2.5 Characteristics common to protein-based poison - antidote killer systems

Unlike most gene families which share sequence homology, the poisons and antidotes of proteic killer system have poor sequence homology (Ruiz-Echevarría *et al.*, 1991b). Homology searches using the *pem* system have detected chromosomal homologs but deletion of these has been shown to have no effect on cell growth or viability (Masuda *et al.*, 1993; Masuda and Ohtsubo, 1994). This raises the question of whether these chromosomal homologs act in some subtle way to ensure chromosomal segregation or are they relics of plasmid integration, serving no purpose to the chromosome. Only a very thorough investigation of these chromosomal genes will help answer this question. Since sequence homology cannot be used as the basis for grouping proteic killer systems, we must look at the mechanisms of these systems to find unifying features. The systems all contain two small essential genes with the gene encoding the antidote lying upstream of the poison gene. There is, however, a recent exception to the antidote before toxin layout; the *higB* poison gene of plasmid Rts1 lies upstream of the *higA* gene which encodes the antidote (Tian *et al.*, 1996). The production of excess antidote and the more rapid degradation of the antidote than of the poison are common to all proteic killer systems in which these aspects have been studied. The mechanism whereby different amounts of poison and antidote are produced differs in the various systems with codon bias, ribosome binding site variation and overlapping reading frames contributing to the differences. Also notable is the fact that all systems are autoregulated either by the antidote, as happens for the *parDE* system of RK2 (Roberts *et al.*, 1993), or by a complex consisting of inactive poison and antidote (Ruiz-Echevarría *et al.*, 1995a; Tam and Kline,1989b). The proteic killer systems appear to have originated from a common ancestor and although the poison and antidote have co-evolved to effect a number of different targets, the mechanism of regulation and gene layout have been conserved.

1.3.2.6 Restriction and modification enzymes as stability determinants

The proteic killer systems discussed to this point all use direct poison:antidote interactions to ensure plasmid containing cells are not killed by the poison. Restriction and modification cassettes have been shown to confer stability on unstable plasmids (Kulakauskas *et al.*, 1995; Naito *et al.*, 1995). The poison (the restriction enzyme) and the antidote (the modification enzyme) do not directly interact, but exert opposing effects on the target of the poison (DNA).

Figure 1.8 shows how, in plasmid containing cells, the modification enzyme modifies the target DNA thus preventing restriction by its partner enzyme. Should the plasmid be displaced or lost, the modification enzyme which is produced in smaller amounts than the restriction enzyme decays rapidly and leaves the cell DNA susceptable to the more stable restriction enzyme. This explanation carries the implication that only newly synthesised DNA will be restricted, existing DNA which is methylated will not loose its methylation and will remain intact. It would thus seem that replication is the trigger for cell death. However, this remains to be determined.

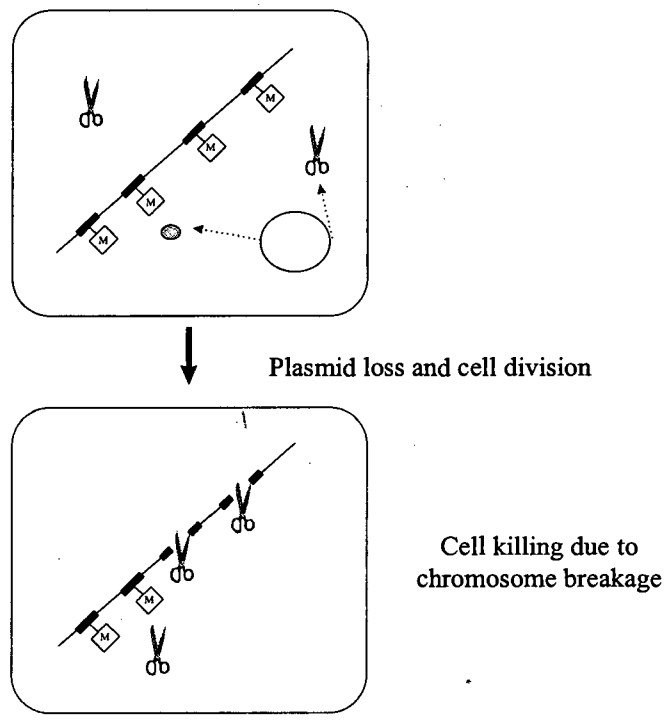


Figure 1.8 Post-segregational killing by restriction enzymes. The restriction enzymes are indicated as ✂ and the less stable modification enzyme as ●. The plasmid containing the restriction and modification genes is indicated as a circle and the modified sites on the genome are indicated as squares containing the letter "M". (adapted from Kobayshi *et al.*, 1996)

1.3.2.7 Bacteriocins in plasmid stability

Colicins and their corresponding immunity proteins are theoretically capable of stabilising plasmids in a manner similar to that of the proteic killer systems. The colicin genes and antibiotic production pathways often occur clustered with their cognate resistance genes and thus possess the mechanics with which to provide plasmid stability in addition to providing the

cells with a selective advantage over other cells in the population. The colicin and immunity proteins of classes E2 and E3 have been shown to directly interact to prevent the toxicity caused by the colicin (Jakes, 1982). Unlike other killer systems, colicin genes and antibiotic production pathways act not only within a single cell but also on a population, as their products are released into the environment and are able to effect other cells in the population. To date the question of extra-cellular release of components of proteic killer systems has not been addressed. A secretory signal sequence is not required for release of colicins as a lysis protein fulfils this role (Jakes and Model, 1979; Luria and Suit, 1982). No protein capable of fulfilling the role of a lysis protein has been found in proteic killer systems, and it thus seems unlikely that the components of proteic killer systems are released from the host cells.

1.4 *kil/kor* of RK2

The *kil/kor* system of RK2 if a highly complex set of genes that has been implicated in various plasmid encoded functions such as conjugation, tellurite resistance and plasmid stability (Thomson *et al.*, 1993; Grewal, 1990; Motallebi-Veshareh *et al.*, 1990). The *kil* loci confer a phenotype of host lethality which can only be suppressed when the relevant *kor* loci are supplied in *trans*. The *kil* loci, *kilA*, *kilB*, *kilC* and *kilE* can only be cloned in the presence of the appropriate *kor* genes (kil-override) which negatively regulate their transcription (Figurski *et al.*, 1982; Kornacki *et al.*, 1993). The central control operon consists of the *korA, incC, korB, korF* and *korG* genes while the *korE* gene is located separately within the *kilE* locus (Pansegrau *et al.*, 1994). The genetic organisation of the IncPα (RK2) genome is shown in Figure 1.9. The *kilA* locus encodes the proteins, KlaA (28.4 kDa), KlaB (42.2 kDa) and KlaC (28.8 kDa) and is repressed by KorA, KorB, and KorE (Figurski *et al.*, 1982; Goncharoff *et al.*, 1991; Walter *et al.*, 1991; Young *et al.*, 1985). Hydrophobicity analysis indicates that KlaA and KlaB are likely to be soluble cytoplasmic proteins while KlaC has four potential membrane-spanning domains and is probably located in the cell membrane (Goncharoff *et al.*, 1991). Each of these genes is capable of expressing a host lethal phenotype while all three genes were required to express tellurite resistance or inhibition of conjugal transfer of IncW plasmids (Goncharoff *et al.*, 1991). Tellurite resistance and inhibition of transfer occur even in the presence of KorA and KorB which are known to repress transcription from the *kilA* promoter (Goncharoff *et al.*, 1991). This activity from a supposedly repressed operon may either be due to read-through

from the upstream *kilE* locus or only partial inhibition of *kilA* due to the lack of KorE which has also been shown to inhibit *kilA* (Goncarhoff *et al.*, 1991; Pansegrau *et al.*, 1994).

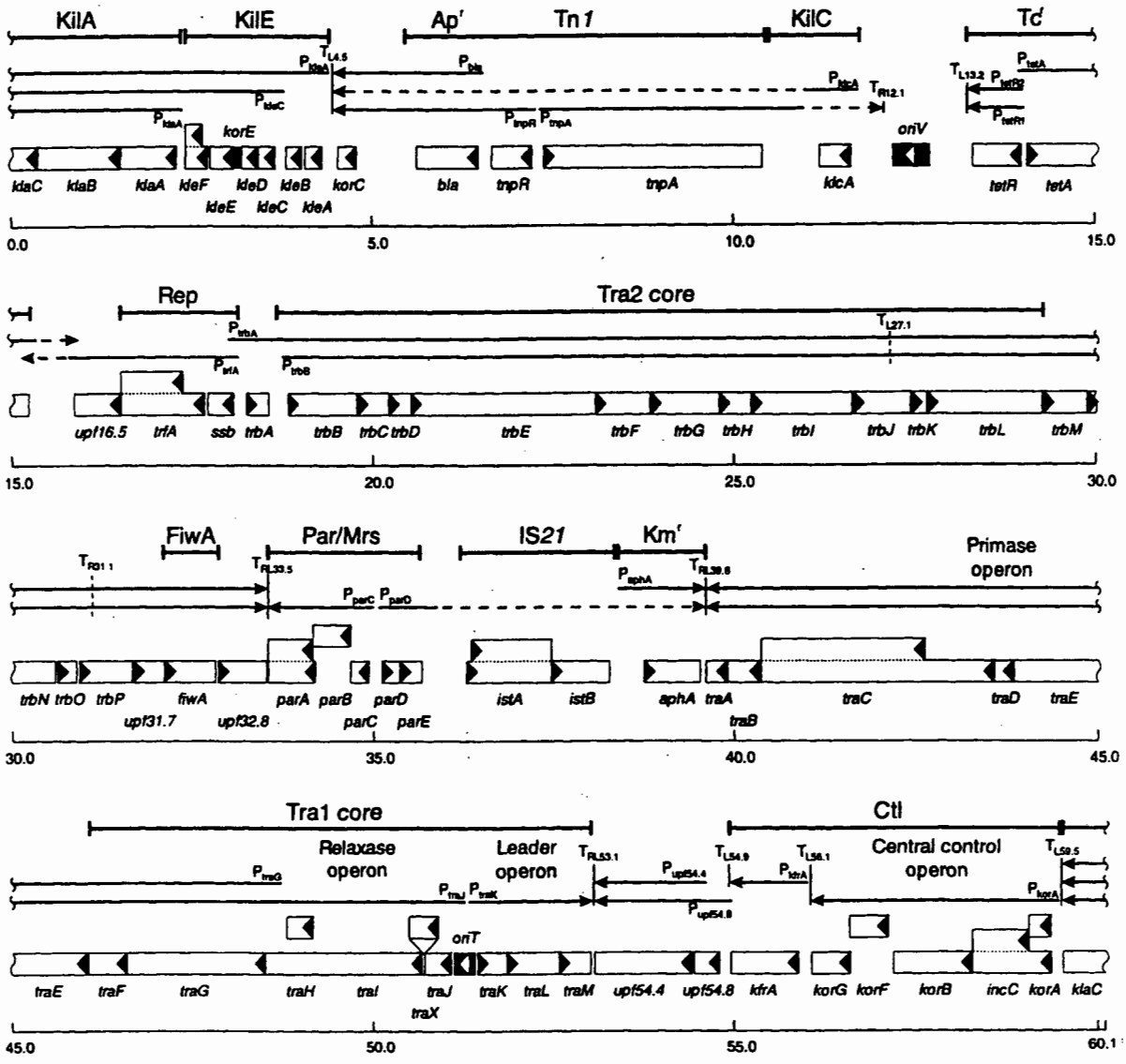


Figure 1.9 Genetic organisation of the IncPα genome. Distinct regions on the IncPα map are marked by bold lines. The extension of transcripts is indicated by light arrows. Where transcription end points are not defined, lines are broken. Boxes with triangles mark the position of genes. Filled triangles indicate the direction of translation. Black boxes represent intergenic regions of defined function, i.e. origins of vegetative and transfer DNA replication. Open triangles within black boxes indicate the direction of vegetative and transfer replication respectively. (from Pansegrau *et al.*, 1994)

Unregulated expression of *klaA* has been shown to inhibit cell growth, alter the outer cell membrane and inhibit cell division, further reinforcing the suggestion that it is membrane

associated (Saltman *et al.*, 1991). The *kilB* locus forms part of the Tra2 core region which is involved in conjugal transfer and surface exclusion. The *klbB* (*trbC*) gene is responsible for the host lethal phenotype (Thomson *et al.*, 1993). The *kilB* locus which is negatively regulated by *korB* is the only *kil* locus which has a single *kor* regulator (Figurski *et al.*, 1982). Over-expression of the *trbB* gene causes reduced expression from *trfA* and *kilB* operon fusions, but Thomson *et al.* (1993) have shown that this is insufficient for suppression of host lethality (Jagura-Burdzy *et al.*, 1992; Thomson *et al.*, 1993). The KlbB protein (15.0 kDa) is predicted to be a membrane protein. The KlbA protein (35.0 kDa) has an ATP binding domain, and KlbC (12.1 kDa) has a leucine zipper motif common to eukaryotic regulatory proteins (Thomson *et al.*, 1993). The *kilC* locus which is regulated by the *korA* and *korC* genes encodes the *klcA* gene which has 56% similarity to the ArdB antirestriction protein of the IncN plasmid pKM101 (Belogurov *et al.*, 1993; Pansegrau *et al.*, 1994). This possible antirestriction role differs from that of the *hok/sok* system in that *klcA* is some 20 kb from the leading region of conjugal transfer and thus is not ideally situated to ensure plasmid establishment during conjugation. The *kilE* locus which requires KorA or KorC in *trans* for inhibition of the host lethal phenotype consists of 2 operons both of which are preceded by sequences resembling strong σ^{70} promoters (Kornacki *et al.*, 1993). The *kleA* operon encodes the KleA (8.7 kDa) protein and the KleB (7.6 kDa) protein (Kornacki *et al.*, 1993). The *kleC* operon encodes the KleC (9.2 kDa), KleD (8.0 kDa), KleE (12.2 kDa) and KleF (11.3 kDa) proteins (Kornacki *et al.*, 1993). The promoters of both operons are similar and contain two palindromes A and C. KorA binds to A and KorC to C, in a synergistic manner (Kornacki *et al.*, 1993). The *kilE* locus is required for plasmid stabilisation in *Pseudomonas aeruginosa* but is dispensable in *E. coli* (Wilson *et al.*, 1996).The central control operon which is autorepressed by *korA* and *korB* together with the *krfA* gene have been shown to stabilise low copy number plasmids. Inactivation of the *incC* region removes this stabilising effect (Motallebi-Veshareh *et al.*, 1990). The *kfrA* gene encodes a protein of 308 amino acids which auto-represses its own production by binding through its N-terminal domain to an operator which overlaps the promoter (Jagura-Burdzy and Thomas, 1992; Thomas *et al.*, 1990). Repression is also exerted by *korA*, *korB*, *korFI* (*korF*) and *korFII* (*korG*) (Jagura-Burdzy *et al.*, 1991). The C-teminal end of the KfrA protein forms a rod-like domain which could allow attachment to cellular structures while the N-terminal end interacts with the *kfrA* operator (Jagura-Burdzy and Thomas, 1992). The *incC* polypeptide shows strong similarity to SopA of F and ParA of P1 while the KorB polypeptide shows strong similarity to SopB and ParB (Motallebi-Veshareh *et al.*, 1990). These similarities and the similar layout of

the central control operon to the Sop system of F and Par of P1 suggest that this region has a similar role in partition. Both Sop and Par have *cis* acting regions downstream of their operons. It thus interesting to note that all the gene products of the central control operon with the exception of IncC bind to the operator of *kfrA* which lies in an analogous position to *sopC* or *parS*. The inter-linking of replication control, conjugation and stability shown by RK2 is unique. Although components of the *kil/kor* system show similarity to other partition systems and the central control operon is capable of conferring stability, very little is know about how exactly *kil/kor* stabilises plasmids.

1.5 *par* of pSC101

Although extensively studied, the *par* locus of pSC101 is poorly understood and the extent to which it affects replication rather than acting as a stability determinant remains to be decided. The *par* locus covers approximately 370 bp and contains two lateral direct repeats (a and a′ in Figure 1.10) as well as central inverted repeat (b in Figure 1.10) which has the potential to form a hairpin loop with either of the direct repeats (Meacock and Cohen, 1980; Tucker *et al.*, 1984). Miller *et al.*, 1983 and Tucker *et al.*, 1984 claim that the region contains no ORF, yet Linder *et al.*, 1983 claims the region contains two ORFs of 49 codons each. These discrepancies need to be addressed by analysis of the transcription-translation products, if any, from the par locus. The *par* locus is capable of efficient stabilisation of both pSC101 and the unrelated replicons p15A and R1 (Meacock and Cohen, 1980; Nordström *et al.*, 1981). Deletion of either a or a′ results in a stable plasmid which has reduced ability to compete with the wild type pSC101 and is denoted by the Cmp$^-$ phenotype (Tucker *et al.*, 1984). The deletion of any two of the repeats or all three repeats results in a loss of stability and is denoted Par$^-$ (Tucker *et al.*, 1984). These results suggest that formation of a stem-loop by the a or a′ and b may be essential for stabilisation and that differences in stability are due to differences in the ability of plasmids to form these stem loops. When all three repeats are removed the plasmid is actively destabilised, as plasmid loss is greater than would be assumed for a randomly segregated plasmid with no stabilisation system. This phenotype is denoted Super-par$^-$ (Tucker *et al.*, 1984). The Super-par$^-$ plasmid's high rate of loss appears due to the entire pool of plasmids being segregated as a single unit (Tucker *et al.*, 1984). As is the case with some of the active partition systems, the *par* locus has been shown to bind to the cell outer membrane fraction (Gustafsson *et al.*, 1983). DNA gyrase has been shown to bind-specifically to the *par* locus. Its function at this site is not

known and plasmids deleted for the entire *par* locus show neither altered copy number nor altered resolution of multimers (Wahle and Kornberg, 1988).

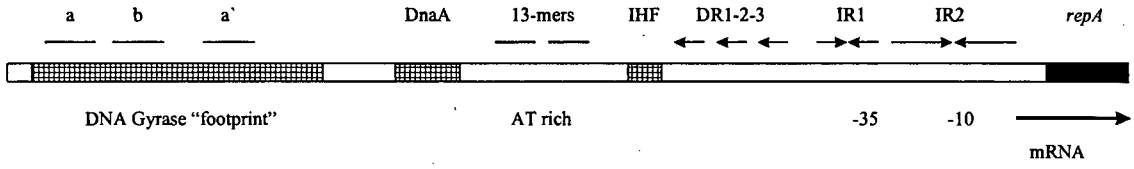


Figure 1.10 pSC101 origin of replication and *par* locus. IHF, DnaA and gyrase binding sites are indicated by hatched boxes. The partially repeated sequences of the *par* locus are indicated by the letters a, b, a′ and horizontal lines. The AT rich repeats of the origin of replication are indicated by horizontal lines and the label '13-mers', while direct repeats and inverted repeats are indicated by the prefix DR and IR respectively. (adapted from Conley and Cohen, 1995a.)

Replacement of the *par* locus by a gyrase binding site of comparable strength failed to stabilise plasmid inheritance, indicating that gyrase binding may be incidental to stability (Yang and Ames, 1988; Biek and Cohen, 1992). The *par* locus could, however, be partially replaced by a strong promoter which causes increased negative supercoiling (Beaucage *et al.*, 1991). A localised increase in negative supercoiling at the origin of replication has been shown to stabilise plasmids containing deletions of one or more of the *par* repeats (Miller *et al.*, 1990; Conley and Cohen, 1995a). The removal of the *par* locus to a site further from the origin of replication causes impaired stability and reduced supercoiling at the origin. This implies that the localised supercoiling at the origin of replication is at least partially responsible for stability (Conley and Cohen, 1995a). Manen *et al.*, (1990) found that the lambda O gene is capable of substituting for the *par* locus, this gene has strong homology to the *par* locus and may perform the same function as *par* in lambda. Unlike other researchers Manen and co-workers detected a drop in copy number of Par$^-$ plasmids. They also found that the O gene was capable of restoring the stability of deletion derivatives of the *par* locus but not of complete restoration of the copy number. This discrepancy is hard to explain but may be due to Manen and co-workers using plasmids containing the replication origins of both pSC101 and pBR322. The use of a PolA$^-$ strain in some of their experiments and the assumption that in such a strain the pBR322 origin exerts no effect may not be valid as ColE1 can affect replication of a plasmid containing

two origins of replication in PolA$^-$ strains (personal observations). The presence of the *par* locus has been shown to increase protein binding at the origin of replication which lies more that 200 bp from the *par* locus (Ingmer and Cohen, 1993). This increased binding may be due to the increased negative supercoiling caused by the *par* locus however how this relates to increased stability, especially in heterologous replicons is unclear. In contrast mutations in the replication protein RepA of pSC101 have been found to cause increased stability of Par$^-$ pSC101 (Conley and Cohen, 1995b). This increased stability may be due to enhanced formation of a RepA-Ori-DnaA complex which in addition to causing increased replication also has some stabilising effects (Biek and Cohen, 1989; Beaucage *et al.*, 1991; Miller and Cohen, 1993).

The *par* locus of pSC101 appears to ensure stability in at least two ways. In a heterologous plasmid it acts by altering supercoiling and this, via an unknown mechanism, ensures stable inheritance. In pSC101 in addition to increasing the supercoiling or perhaps because of the increase, the *par* locus ensures efficient formation of the replication initiation complex which appears to increase stability in a manner that is disproportionate to the increase in copy number. Strains containing a mutation in *topA* have been shown to increase the stability of unstable Par$^-$ P1, Par$^-$ F plasmids and Par$^-$ pSC101 (Austin and Eichorn, 1992; Miller *et al.*, 1990). It has been proposed that an increase in superhelical density which can be achieved either in a *topA* mutant or by adding the pSC101 *par* site allows plasmids to dissociate from each other and segregate randomly which they are not able to do normally (Austin and Eichorn, 1992). The inability of an equally strong gyrase binding site to substitute for the *par* locus indicates that factors other than gyrase binding are necessary to ensure a sufficient change in topology for increased plasmid stabilisation.

The stability mechanism of pSC101 is only slightly better understood now than it was 17 years ago when it was first detected. The interplay between gyrase and topoisomerase which acts to regulate supercoiling appears to be central to plasmid stability. How exactly increased supercoiling causes increases plasmid stability remains to be elucidated. The metabolic and replication cost of increased supercoiling to the plasmid also requires investigation.

1.6 Combined systems

While even a single stabilisation system can dramatically increase the stability of a plasmid there is a tendency for plasmids to acquire multiple systems. This ensures that should a system malfunction or be lost due to a mutation, there is always another to replace it. Should a plasmid not be segregated by one system, another system will take care of its segregation. The F plasmid has the *sopABC* system, the *ccd* system, the *flm* system, *srnB* locus as well as the *resD* encoded resolvase.

Partition systems of F

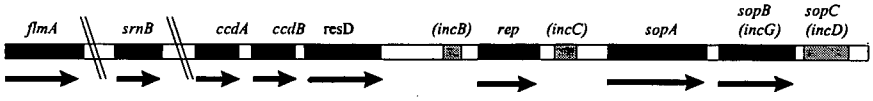


Figure 1.11 Plasmid stability systems found on the F plasmid. Adapted from Mori *et al.*, 1986.

RK2 encodes a *parCBA* system, the *parDE* system and the ParA resolvase. The prophage P1 encodes not only the *parABS* locus, the *phd/doc* system and *lox/cre* resolvase system but also the *Eco*P1 restriction and modification system (Bickle and Kruger, 1993). The types of partition system that a plasmid possess are usually different, in the case of the F plasmid however the *flm* and *srnB* loci can probably viewed as functionally overlapping. By possessing different systems, a plasmid ensures 'fail safe' stability. The cumulative effect of several stabilisation systems has been clearly illustrated on plasmids containing both the *hok/sok* and the *parDE* genes. The combined systems decrease the probability of plasmid loss by 10^{-6} compared to a system containing only *parDE* and by 10^{-3} compared to a system containing only *hok/sok* system (Pecota *et al.*,1997). Now that many of the questions relating to how stability loci function have been determined, the question of when they function should be asked. Do the systems function all the time or do certain systems only function under certain conditions such as low copy number or high copy number? The expression of a multimer resolution system would be wasted at low copy numbers when multimers are less likely to form, while post-segregational killing is energetically wasteful in high copy number plasmids where the possibility of a plasmid free segregant is low. Sobecky *et al.* (1996) have started to

answer some of these questions by showing that while the *parCBA* system was sufficient for stabilisation under normal conditions, the *parDE* system came into play only when the copy number dropped, at which time efficient partitioning is critical. Study of the roles of plasmid conjugation genes on stability of promiscuous plasmid populations, and hemi-methylation on stability of individual plasmids may yet provide more examples of plasmid stabilisation mechanisms.

1.7 Aims of this project

The initial aim of this project was determine the roles of ORF 3, 4 and 5 in the control of pTF-FC2 replication. The IncQ plasmid RSF1010 has a similar layout of replication genes to pTF-FC2 and the E and F genes exist in a position similar to ORFs 3, 4 and 5. The E and F genes have a similar size range to ORFs 3, 4 and 5, this leads to speculation that they might have a similar function. Since the F gene has been implicated in regulation of a promoter which transcribes the *repA* and *repC* genes of RSF1010, it was thought that ORFs 3, 4 and 5 may also regulate essential replication proteins. After many frustrating experiments designed to investigate plasmid regulation by ORFs 3, 4 and 5, the problem was resolved when it was discovered that spontaneous mutants in which ORFs 3, 4 and 5 had been deleted yet were able to replicate at unchanged copy numbers. This indicated that these genes were not involved in plasmid replication.

The results that were obtained suggested that ORFs 3, 4 and 5 were a proteic plasmid stabilisation system even though the products of the three genes did not show any sequence homology to known partition systems of that type (chapter 2). The project's aim was thus shifted to determining the mechanism of plasmid stabilisation and identifying the role of each of the ORFs (chapter 3). The considerable differences in stability conferred by the pTF-FC2 genes in different hosts also prompted investigations into the factors which determined host range-specificity (chapter 4).

Chapter 2

pTF-FC2 encodes a proteic plasmid stabilisation system

2.1 Introduction

Thiobacillus ferrooxidans strain FC1 is a member of a consortium of bacteria that formed the inoculum of a biooxidation plant used for the pretreatment of gold-bearing arsenopyrite ores at the Fairview mine, South Africa (Rawlings and Silver, 1995). Plasmid pTF-FC2 (Figure 2.1) was one of three plasmids resident in strain FC1 (Rawlings *et al.*, 1983). The nucleotide sequence of pTF-FC2 has been determined (Genbank accession numbers M64981 and M35249). Three regions of the plasmid have been identified, a replicon, a mobilisation region and a transposon-like region.

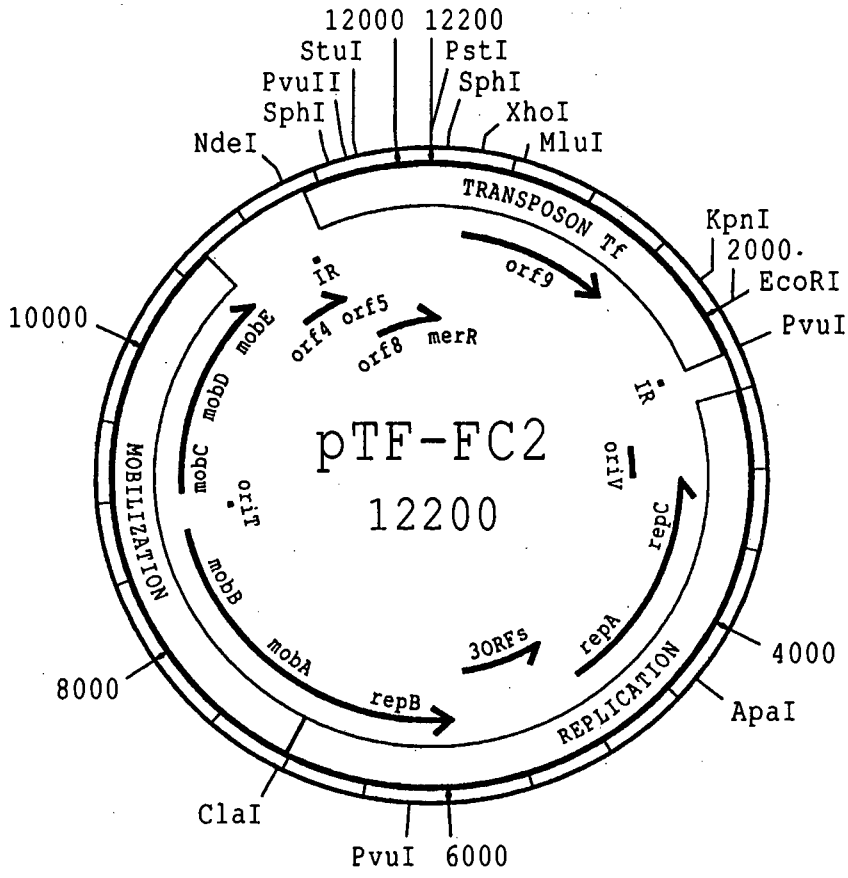


Figure 2.1 pTF-FC2 showing the replication, mobilisation and transposon regions. The numbers indicate distance from the unique *Pst*I site.

The pTF-FC2 replicon clearly resembles that of the IncQ plasmids RSF1010, R300B and R1162 (Dorrington and Rawlings, 1990). The RepB primase shows 23% amino acid identity to the repB of RSF1010. The RepA helicase and RepC origin binding proteins genes show 47% and 60% amino acid identity to their RSF1010 counterparts respectively (Figure 2.2). Like the IncQ plasmids, pTF-FC2 has a broad host range and has been found to replicate in a large number of Gram-negative bacteria (Rawlings and Kusano, 1994). Nevertheless, plasmid pTF-FC2 is compatible with the IncQ plasmids in *E. coli* and the IncQ plasmid R300B was unable to complement *repA, repB* or *repC* mutants of pTF-FC2. The mobilization region most closely resembles the TraI region of the IncPα plasmids RP4 or RK2 and the IncPβ plasmid R751 (Pansegrau *et al.*, 1994). Plasmid pTF-FC2 is very efficiently mobilised by either type of IncP plasmid (Rohrer and Rawlings, 1992). A 3.5 kb transposon, Tn*5467*, which has 38 bp terminal repeat sequences which are identical to those of Tn*21* is present on pTF-FC2 (Clennel *et al.*, 1995). Tn*5467* is inactive but is capable of transposition and resolution if the Tn*21* transposase and resolvase genes are provided in *trans*. The transposon encodes a functional glutaredoxin-like protein and has two ORFs which resemble those of multidrug resistance efflux proteins and transcriptional regulators (Clennel *et al.*, 1995).

Three small ORFs are located between the *repB* and *repA* genes of pTF-FC2 (Figure 2.2). The small size and lack of clear homology to any known genes suggested that these ORFs served either a regulatory or some other non-enzymatic role. Furthermore, the two genes that are located in the equivalent position in the case of the IncQ plasmid RSF1010, encode two small proteins (E and F) which have been implicated in the regulation of plasmid replication (Maeser *et al.*, 1990). This chapter is a report of the experiments which led to the conclusion that the three small ORFs are not involved in the control of plasmid replication. Instead they comprise a proteic poison antidote stability mechanism.

2.2 Materials and methods

2.2.1 General methods

Unless otherwise stated, nucleic acid extractions, digestions, ligations and other manipulations were performed according to Sambrook *et al.* (1989). All strains and vectors used in this work

are listed in appendix 1. All constructs produced or used are listed in appendix 2. Southern hybridisation was performed using Hybond-N+ membrane (Amersham), Boehringer-Mannheim Dig DNA labeling kit (cat # 1175033) and the Boehringer-Mannheim Dig nucleic acid detection kit (prod # 1175041) according to the manufacturer's instructions. Northern hybridisation was performed according to the method of Church and Gilbert (1984) using GIBCO BRL molecular weight markers (Cat. # 5620SA). α-P^{32} was supplied by Amersham (UK). Site directed mutagenesis was performed using the pMa and pMc plasmid pair according to the method of Stanssens *et al.* (1989). Oligonucleotide primers used to perform mutagenesis, amplify DNA and sequence DNA are listed in appendix 3. All constructs made from PCR (Polymerase Chain Reaction) amplified DNA were sequenced to confirm the fidelity of amplification and cloning. DNA sequencing was performed using a ALF Express automated sequencer for 12 hours at 60 W and 25 mA at 55°C. Templates were prepared using the Pharmacia Autoread Kit (prod # 2729002) according to the manufacturers instructions.

2.2.2 Media and growth conditions

All *E. coli* strains were maintained at 37°C on LA (Luria Bertani Agar) or grown in LB (Luria Bertani Broth) with the appropriate antibiotic selection unless otherwise stated. Normal working concentrations were 100 µg/ml ampicillin, 50 µg/ml kanamycin, 30 µg/ml chloramphenicol, 20 µg/ml tetracycline and 25 µg/ml streptomycin. *Pseudomonas putida* was grown at 30°C on LNG (LB or LA supplemented with 4% glucose and 4% KNO_3) plates or broth and selected for with 300 µg/ml chloramphenicol and 250 µg/ml kanamycin.

2.2.3 Isolation of total DNA

Cells were grown overnight in 30 ml LB containing the appropriate antibiotics and harvested by centrifugation at 5000 rpm for 10 minutes. The pellet was resuspended in 10 ml solution I (25 % sucrose, 10 mM Tris HCl pH 8.0, 10 mM EDTA) and lysozyme was added to a final concentration of 1 mg/ml. After incubation at 37°C for 15 minutes the sample was cooled on ice and 5 ml of 0.25M EDTA was added. After a further 5 minutes on ice, 10 ml of solution II (2% SDS, 10 mM Tris HCl pH 8.0, 10 mM EDTA) was added and the sample was incubated at room temperature for 30 minutes. RNAse was added to a final concentration of 50 µg/ml and

the sample was incubated at 50°C for a further 30 minutes after which proteinase K was added to a final concentration of 50 µg/ml and the sample incubated for an hour at 50°C. The sample was then repeatedly extracted with phenol, and precipitated with 1/10 volume 3 M sodium acetate and 2 volumes 100% ethanol before being re-suspended in TE buffer.

2.2.4 Determination of plasmid copy number

Total DNA was isolated from cells carrying the relevant plasmids by the method listed in section 2.2.3. A range of amounts of total DNA was dot-blotted onto Hybond N+ as was a range of quantities of plasmid DNA. The membrane was then probed with Dig-labelled plasmid DNA. The intensity of the hybridisation signals obtained from known quantities of plasmid DNA was compared to similar signals from total DNA. Since equivalent signals indicated equivalent quantities of plasmid DNA, it was possible to calculate the ratio of plasmid DNA to chromosomal DNA. Using the relationship below the copy number of the constructs was determined.

$$\frac{\text{plasmid size * copy number}}{\text{chromosome size}} = \frac{\text{amount of plasmid DNA with signal equivalent to total DNA}}{\text{amount of total DNA with signal equivalent to plasmid DNA}}$$

2.2.5 Tests for plasmid stability

Plasmid stability was determined by growing the plasmid containing cells in batch culture for 100 generations in Terrific Broth (24 g/l yeast extract; 12 g/l tryptone; 4 ml/l glycerol; 17 mM KH_2PO_4; 72 mM K_2HPO_4) without antibiotic selection. Samples taken at 20 generation intervals were, in the case of pOU82-based constructs, plated to LA containing X-Gal (40 µg/ml) and the ratio of blue plasmid containing cells to total cells was determined. Cells containing other constructs were plated first to LA plates and, following overnight growth, 100 colonies were toothpicked onto LA containing the appropriate antibiotic and the percentage survival recorded. The loss frequency (LF) was calculated according to the method of Gerdes *et al.* (1985).

$$LF = 1 - (V_1/V_2)^{[1/(g_1-g_2)]}$$

Where V_1 and V_2 are the frequency of plasmid-containing cells after g_1 and g_2 generations respectively.

2.3 Results

2.3.1 Mutagenesis of ORFs 3, 4 and 5 of pTF-FC2

In order to determine their function, site directed mutagene2sis was used to introduce a frame shift mutation into each of ORFs 3, 4 and 5.

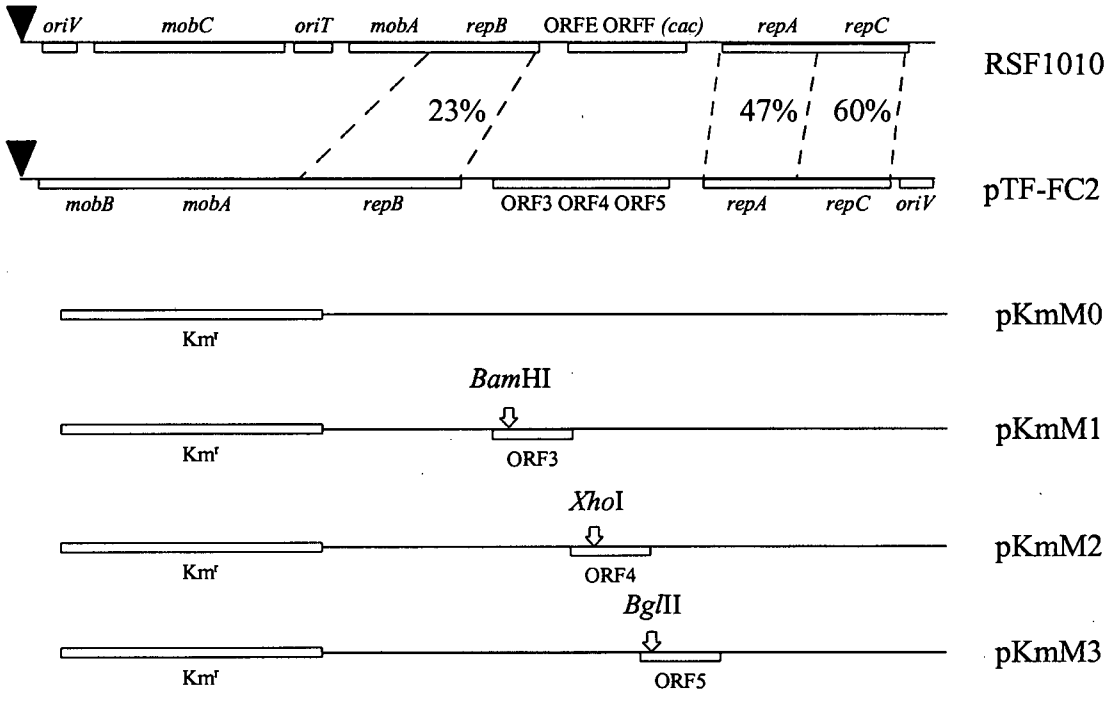


Figure 2.2 A comparison of the IncQ plasmid RSF1010 and pTF-FC2 which has been simplified by excluding the transposon present in each plasmid with the position of each transposon indicated by an inverted triangle. Broken lines between plasmids indicate similar genes with the amino acid identity of the gene products shown as a percentage. The region of pTF-FC2 used to construct pKmM0 and the positions of the three frame-shift mutations introduced into ORFs 3, 4 and 5 are shown below the map of pTF-FC2.

The entire pTF-FC2 insert from pTF200 (appendix 2) was excised and a *Eco*RI - *Hinc*II fragment and was cloned into the *Eco*RI - *Sma*I sites of the mutagenesis plasmid pMa. The primers MP1, MP2 and MP3 were then used to introduce a *Bam*HI site into ORF 3, an *Xho*I site into ORF 4 and a *Bgl*II site into ORF 5 respectively. The introduction of these sites

resulted in a frame shift in each of these ORFs thus causing them to produce nonsense proteins. An unmutated replicon and each of the mutant replicons were excised from the pMa mutagenesis vector using *Eco*RI and *Sal*I and ligated to the kanamycin resistance marker from Tn5 which had been digested with *Eco*RI and *Sal*I. This produced the pTF-FC2 replicon-based plasmids, pKmM0 (unmutated), pKmM1 (ORF 3 mutant), pKmM2 (ORF 4 mutant) and pKmM3 (ORF 5 mutant) as indicated in Figure 2.2. The integrity of the mutants was confirmed by restriction mapping and sequencing. The effect of the mutations in ORFs 3, 4 and 5 on the pTF-FC2 replicon was tested by determining the plasmid copy number in *E. coli* JM105. Plasmids pKmM0, pKmM2 and pKmM3 were found to have a copy number of 10 to 14 per genome which is identical to that for pTF-FC2 (Dorrington and Rawlings, 1990). However, the copy number of pKmM1 was reduced to one under conditions which selected for the plasmid, the host bacterium also grew poorly. In order to test whether plasmid host-range had been affected, the unmutated replicon and mutants were transformed into the *E. coli polA* mutant GW125a and *P. putida.* Only plasmids pKmM0, pKmM2 and pKmM3 retained the ability to replicate in these hosts. The region containing ORF 3 was excised from pTV4293 (appendix 2) on a *Hin*dIII - *Sma*I fragment and was cloned into the *Hin*dIII - *Nru*I sites of the broad-host-range vector pSUP106 to create pORF 3. When this construct was present in the *E. coli polA* mutant or *P. putida* cells, the ability of coresident pKmM1 to replicate was restored, whereas coresident pSUP106 vector controls were unable to replicate. Initially it was assumed that the frame-shift mutation in ORF 3 had effected the replication control mechanism of the replicon resulting in a reduction in copy number and host-range. However, when pKmM1 was introduced by itself into either *P. putida* or the *E. coli polA* mutant, a small number of spontaneous mutants arose that were able to replicate in these hosts. Plasmids were extracted from thirty six of these mutants and analysed by restriction enzyme mapping. This revealed that two types of plasmid deletion, which allowed the mutants to regain the ability to replicate in *P. putida* or *E. coli polA* mutants, had occurred. A representative of each type of deletion mutant was selected and sequenced. One type of deletion mutant (pKmM1del1) had a 711 bp deletion bordered by a 9 bp repeated sequence (5'-CGCAGAAGA-3') and the deletion had removed all of the DNA encoding ORFs 3 and 4 as well as the first 134 bp of ORF 5 (Figure 2.3). The other type of mutant (pKmM1del2) had a 104 bp deletion and was bordered by a different 7 bp repeat sequence (5'-CAGGAGC-3'). This deletion had removed 44 bp from the 5' end of *rep*B and the region upstream of the start codon of ORF 3. Both the deletion mutants had copy numbers that were indistinguishable from the unmutated replicon.

CCGCCGATCTTGGCAAGTACCGGCAGCAGTGGGAGAAGCTGGAAGGGCGCGAGCCTGTACGACAGG
A A D L G K Y R Q Q W E K L E G R E P V R Q

↓Deletion 2 start ↓Deletion 1 start
AGGAGCAGGCAAAGGCGCAGAAGATCGAGCGCGACAACTCGCCGGGAATGAGTCTCTAGCGTTGCG
Q E Q A K A Q K I E R D N S P G M S L *

Deletion 2 end↓ *pasA* start ➢
TGGTGGTTGTGATATACTTGTATAGCGTTTTCAGAACAGGAGCCGAAACATGCTTGCAATCCGACT
M L A I R L

T insertion to give frame shift and *BamHI* site ⩛
GCCCGCCGAAGTGGAAACCCGCCTTGAAGCACTGGCGCAGGCCACAGGGCGGACCAAGACTTTCTA
P A E V E T R L E A L A Q A T G R T K T F Y

TGCCCGCGAAGCCATCCTTGAGCACTTGGATGACCTCGAAGATTTGTACCTTGCAGAGCAACGCCT
A R E A I L E H L D D L E D L Y L A E Q R L

pasB start ➢
GATCGACATTCGCGCAGGCAAAACCCAAACCGTGCCACTC**GAAG**AAGTGATGAAACGCTATGGCAT
I D I R A G K T Q T V P L E E V M K R Y G M
M A

⩛ G insertion to give frame shift and *XhoI* site
GGAAGGTTGAACTCGACCCAGCCGCCGAGCGCGAGCTAGGCAAGATCGACCAGCAGACCGCCCGCC
E G *
W K V E L D P A A E R E L G K I D Q Q T A R

GCATCCTCGCTTTTTTGCATGGCCGTGTCGCCCAGCTCGACGACCCGCGCAGCATTGGCGAAGCCC
R I L A F L H G R V A Q L D D P R S I G E A

TCAAAGGCTCCAAACTGGGAGCCTTCTGGAAATACCGCGTTGGGGATTGGCGAATCATCGCCAGCA
L K G S K L G A F W K Y R V G D W R I I A S

pasC start ➢
TCGAGGACGGTGCTTTGCGCATCCTCGTTATGCGCATCGGCAATCGTA**AGGAG**GTTTACCGCCAAT
I E D G A L R I L V M R I G N R K E V Y R Q
M

⩛ T insertion to give frame shift and *BglII* site
GATCGAATACAGCTACCAGATCGACCCGCGCCCCTCCGACCTTGGCGGCGGCTGGCGGTTGCGCCT
*
I E Y S Y Q I D P R P S D L G G G W R L R L

Deletion 1 end ↓
GTTGGAAAGCGGCGAGGAAGTCGGCGGCGGAGTGTTCCCGTTGTCCGAGTACGCCACAGCAGAGAA
L E S G E E V G G G V F P L S E Y A T A E N

CGCAGAAGAAGCGGCCACGTACGCCTATGAGGACGCCTTGGCCGAGGCTTCGGCGTGGCTGGCATC
A E E A A T Y A Y E D A L A E A S A W L A S

GAGGGGCGAAAATTGAGCGGCGCGGCAGGGGGATTGCGGCCCCGGCAGCGCCTAACCACAACTGTC
R G E N *

Figure 2.3 (previous page) Nucleotide sequence of the region of pTF-FC2 encoding ORFs 3, 4 and 5 as well as the immediate flanking regions. The genes for each ORF have been renamed *pasA, B* and *C*, respectively, and the amino acid sequence of each gene product is indicated below the nucleotide sequence. Sites of insertion of bases into each gene, T insertion into *pasA* (creating pKmM1), G insertion into *pasB* (creating pKmM2) and T insertion into *pasC* (creating pKmM3) are shown above the sequence. The end positions of the spontaneous deletions 1 and 2 are indicated by vertical arrows (↓), and the repeated sequences which flank each spontaneous deletion are highlighted. Potential ribosome binding sites for each gene are indicated in bold. The region which was amplified and fused in-frame to a *lacZ* reporter gene is underlined (section 3.3.4). The sequence above overlaps two sequences which have been deposited in the Genbank data base and extends from nt 1511 (M64981, Dorrington *et al.*, 1991) to nt 835 (M35249, Dorrington and Rawlings, 1990).

2.3.2 Effects of mutations in ORFs 3, 4 and 5 on plasmid stability

The observation that the copy number and host range of the deletion plasmids was the same as for the unmutated plasmid indicated that ORFs 3, 4 and 5 were not essential for plasmid replication nor were they involved in the regulation of pTF-FC2 replication. The stability of each of the mutants in *E. coli* JM105 was then tested and the results are shown in Figure 2.4. In the absence of selection, plasmids containing the unmutated replicon (pKmM0) and the mutation in ORF 5 (pKmM3) were relatively stable and maintained at rates that are higher than would be predicted by random segregation of plasmids with a copy number of 10 to 14. The ORF 4 frame-shift mutant (pKmM2) was less stable than the unmutated plasmid, whereas the ORF 3 frame-shift mutant (pKmM1) was highly unstable. Deletion mutants pKmM1del1 and pKmM1del2 were both unstable, with the plasmid with the more extensive deletion (pKmM1del1) being more stable than the plasmid with the smaller deletion. Thus the pTF-FC2 replicon was stable only when both ORFs 3 and 4 were intact and expressed.

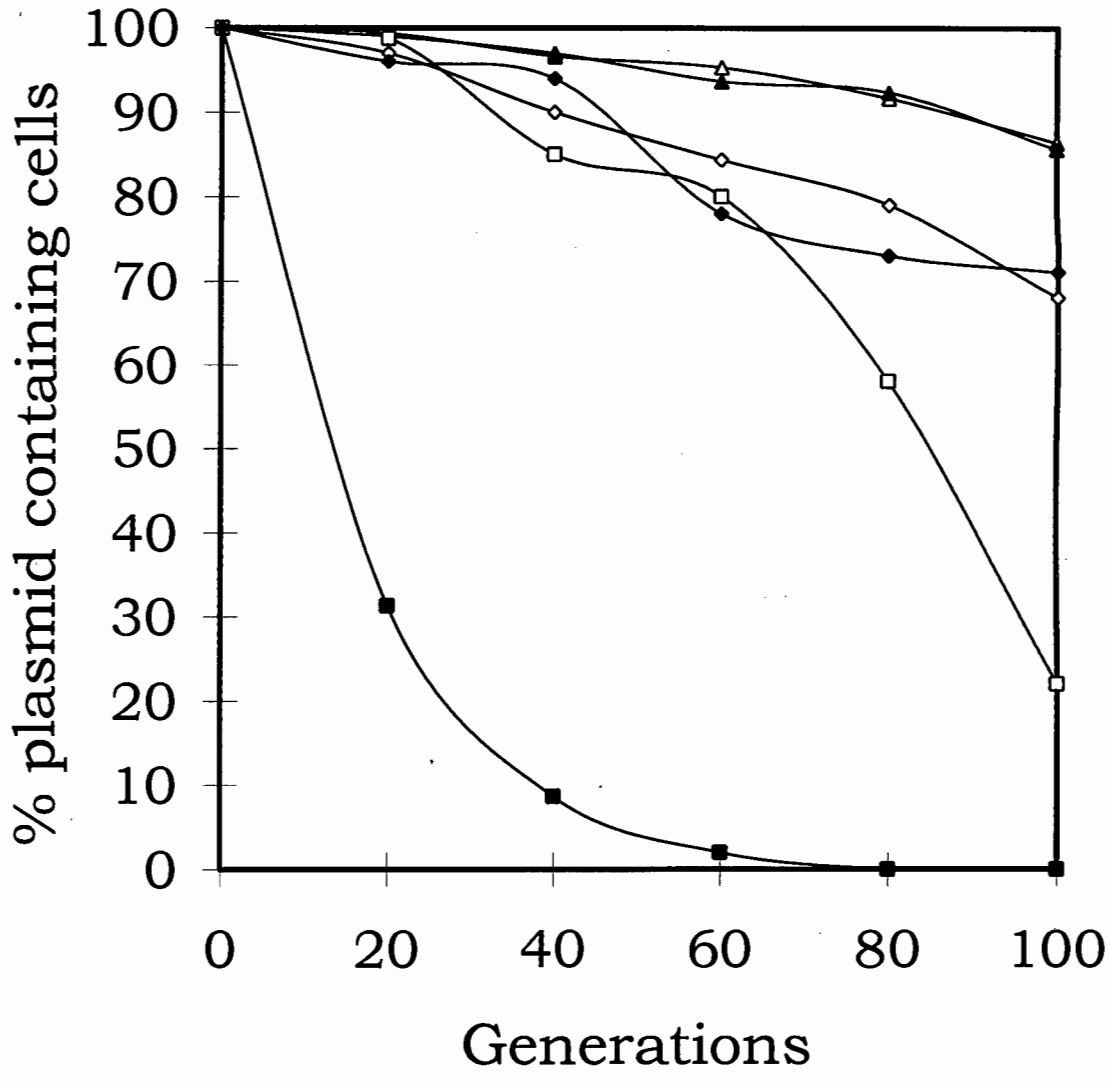


Figure 2.4 Stability of the unmutated pTF-FC2 replication region compared with the stability of plasmids with mutations in ORFs 3, 4, 5, and two spontaneous deletion mutants. Unmutated plasmid pKmM0, (Δ); ORF 3 mutant, (■); ORF 4 mutant, (◇); ORF 5 mutant, (▲); deletion mutant pKmM1del1, (◆) and pKmM1del2 (□). Each data point is the mean of three separate experiments.

2.3.3 Stabilisation of a heterologous replicon by the region encoding ORFs 3, 4 and 5

To confirm that the region encoding ORFs 3, 4 and 5 functions as a plasmid stability system, the ability of this region to stabilise a heterologous R1 plasmid replicon was investigated. Since the plasmid RSF1010 contains two small ORFs in a position analogous to ORFs 3, 4 and 5,

ORFs E and F of RSF1010 were also tested for the ability to confer stability on the heterologous replicon. The region of DNA from 174 bp upstream of the ORF 3 start codon to 23 bp downstream of the stop codon of ORF 5 was amplified from pTV101 (appendix 2) using primers PTF-ECO and PTF-BAM (appendix 3). The ORFs E and F of RSF1010 were amplified from pSUP106 (appendix 1) using primers RSF-ECO and RSF-BAM (appendix 3). These primers amplified a DNA fragment which extended from 96 bp upstream of the start codon of ORF E to 21 bp downstream of the stop codon of ORF F, this region includes the P_4 promoter known to initiate transcription of ORF E and F (Maeser *et al.*, 1990). In both cases the primers had extra bases added to produce an *Eco*RI site on one side of the amplified fragment and a *Bam*HI restriction site on the other. The amplification products were digested with *Eco*RI and *Bam*HI before being cloned into the *Eco*RI and *Bam*HI sites of the unstable R1 derivative, pOU82 (Jensen *et al.*, 1995) to create pOU-PTF (ORFs 3, 4 and 5) and pOU-RSF (ORFs E and F) (Figure 2.5). The stability of these constructs as well as that of pOU82 was tested in *E.coli* JM105 (Figure 2.6). After 100 generations without selection, the control plasmid pOU82 showed a loss frequency of approximately 0.02 as did pOU-RSF while pOU-PTF showed a 2.2-fold increase in stability with a loss frequency of approximately 0.009. This confirmed that the region containing ORFs 3, 4 and 5 was able to confer plasmid stability and the genes were named *pasA* (ORF 3), *pasB* (ORF 4) and *pasC* (ORF 5) (plasmid addiction system).

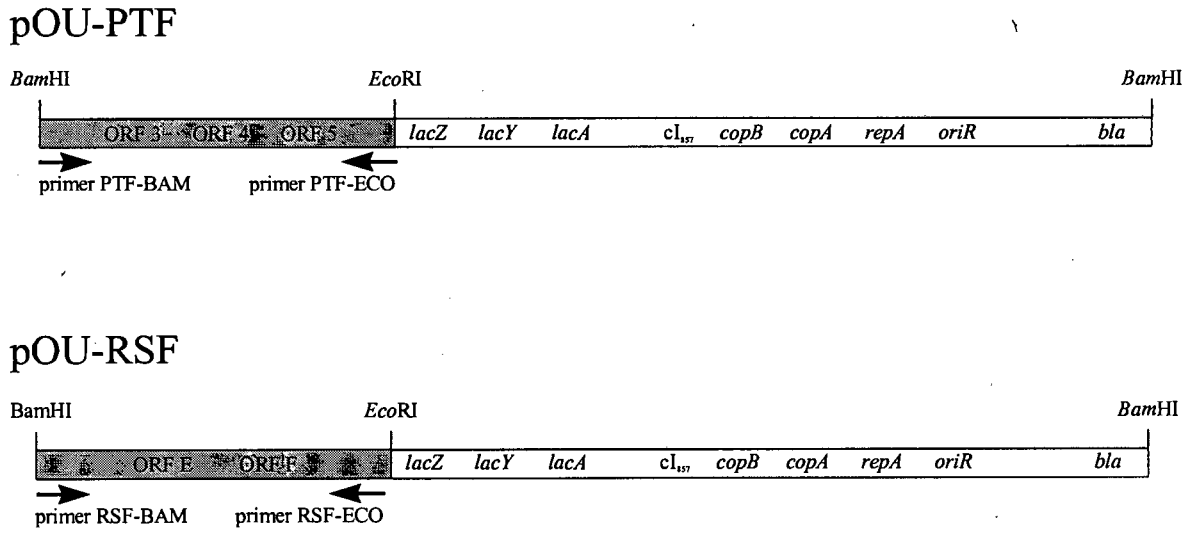


Figure 2.5 Schematic drawings of pOU-PTF and pOU-RSF. The vector pOU82 is indicated in white and the inserts are indicated in grey. The positions and names of primers used to amplify the inserts are shown below the drawing.

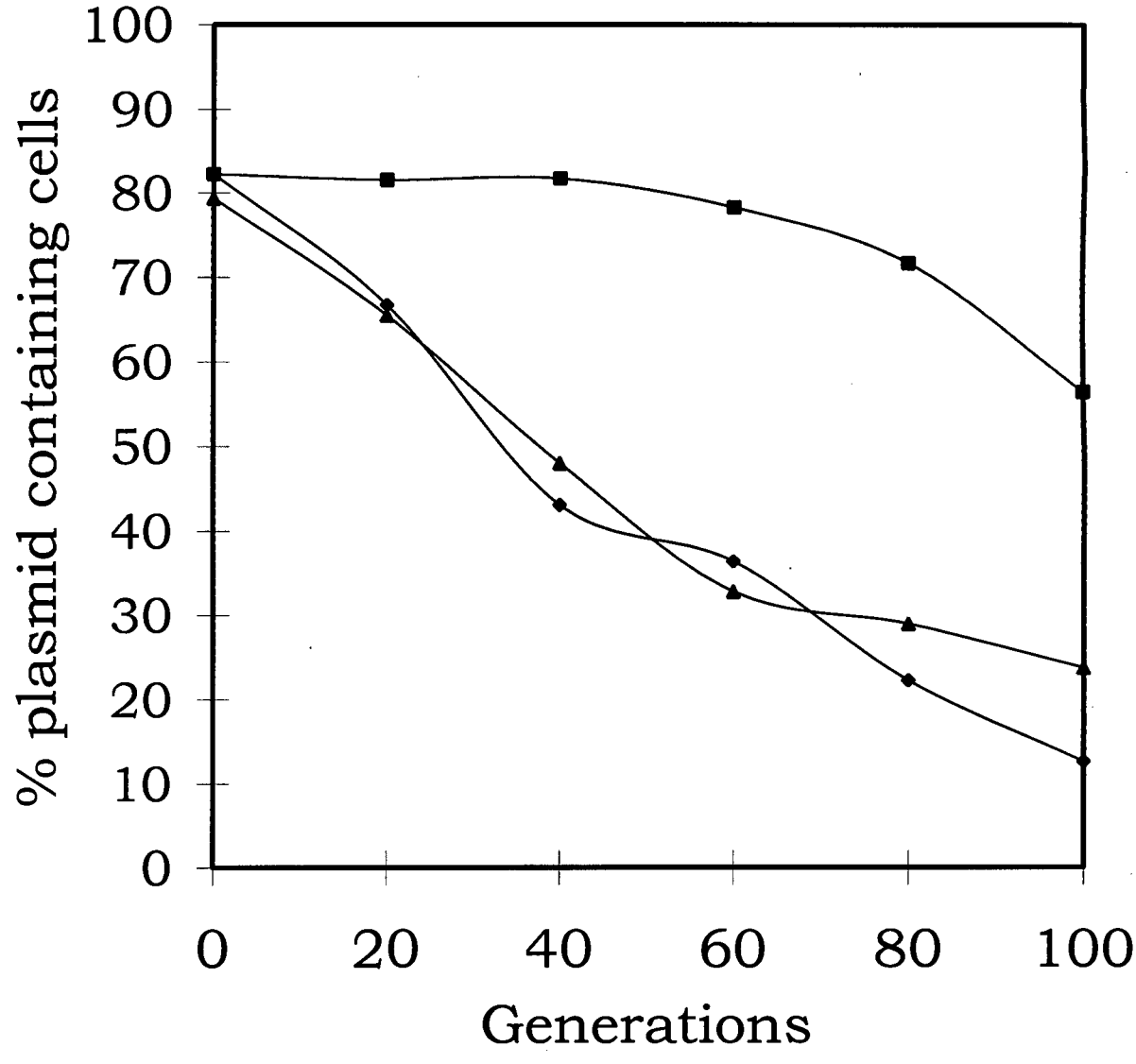


Figure 2.6 Comparison of the stability of pOU82 (◆), pOU-PTF (■) and pOU-RSF (▲). Each data point is the mean of three separate experiments.

2.3.4 Analysis of RNA transcripts produced by the minimal replicon of pTF-FC2

Because ORFs 3, 4 and 5 were situated between the *repB* and *repAC* genes, the transcripts produced by the minimal replicon of pTF-FC2 were analysed to determine if the genes responsible for plasmid stability were co-transcribed with any of the surrounding replication genes. RNA was extracted from *E. coli* JM109 cells, as well as JM109 cells containing pKmM0 and pKmM1del1. The RNA was electrophoresed in triplicate and blotted before being probed with one of three probes, *repB*, ORFs 3, 4 and 5 or *repAC* (Figure 2.8). The *repB* probe was made by excising the *repB* gene from pTV151 (appendix 2) using *Cla*I and *Bam*HI. The ORF 345 probe was made by excising these genes from pTac-pasABC (section 3.3.1) using *Eco*RI and *Hin*dIII. The *repAC* probe was made by excising these genes from pMRepM4 (appendix 2) using *Xba*I and *Eco*RI. All the excised genes were purified by electro-elution from agarose gels before being labeled.

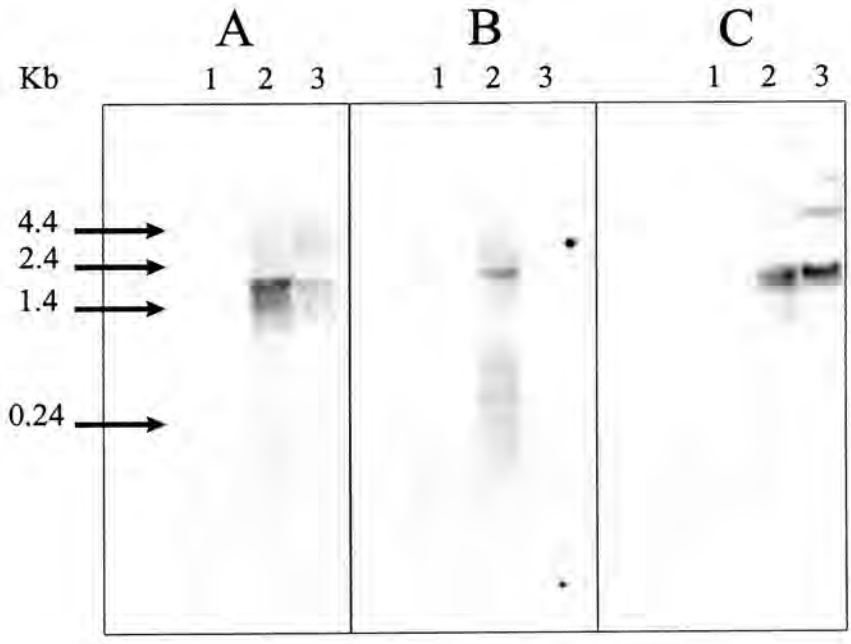


Figure 2.7 Northern blot of transcripts of the minimal replicon of pTF-FC2. The arrows on the left indicate the distances migrated by RNA molecular weight markers. In each panel, lane 1 contains RNA from *E. coli* JM109 (control), lane 2 RNA from *E. coli* JM109 containing pKmM0 and lane 3 contains RNA from *E. coli* JM109 containing pKmM1del1. Panel A was probed with the *repB* gene, panel B with ORFs 3, 4 and 5 and panel C with the *repAC* genes.

The *repB* and *repAC* genes hybridised to an approximately 2 kb transcript from both pKmM0 and pKmM1del1 (Figure. 2.7) The ORF 3, 4 and 5 probe hybridised to an approximately 2 kb transcript from pKmM0 only. No hybridisation was seen to pKmM1del1 by the ORFs 3, 4 and 5. None of the probes hybridised to the control JM109 RNA. Sequence data suggests that a transcript encoding *repB* and ORFs 3, 4 and 5 would be at least 1806 bases in size and that a transcript encoding *repAC* would be at least 1750 bases in size. The best fit for this data is shown in Figure 2.8.

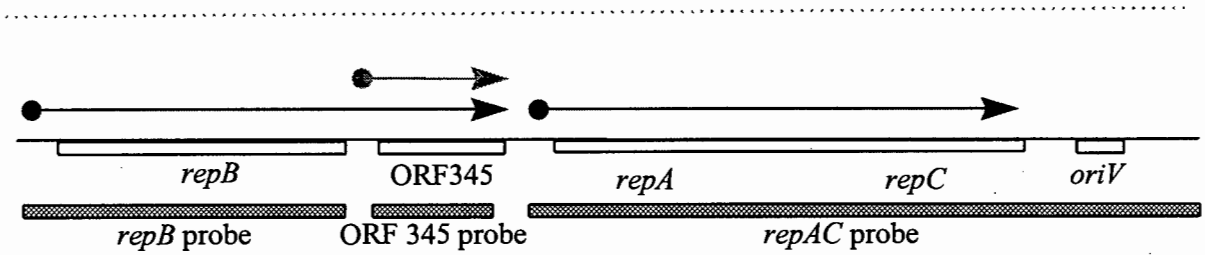


Figure 2.8 Position of probes and proposed transcripts of the minimal replicon of pTF-FC2. Regions used as probes are shown as hatched boxes under the genes they include, and transcripts are indicated as black lines above the genes with arrows indicating the direction of transcription. The grey solid line shows the position of a transcript which based on *lacZ* fusion experiments (chapter 3) is thought to occur but was not detected. Reasons for the inability to detect this transcript are discussed in section 2.4 The grey dotted line indicates a portion of a large transcript that traverses the entire minimal replicon of pTF-FC2.

The transcript sizes suggest that the minimal replicon produces two approximately 1.8 - 2.0 kb transcripts. One transcript appears to originate from a promoter upstream of *repB* and terminates between *pasC* and *repA*. The other transcript appears to originate from a promoter between *pasC* and *repA* and terminates downstream of *repC*. However, this does not explain why pKmM1del1 produced a 2kb transcript which hybridises to *repB*. The deletion of ORFs 3, 4 and 5 should reduce the size of the transcript by 711 bp. Since the 2 kb transcripts migrate at approximately the same speed as 23s rRNA, it is possible that the shortened transcript become entangled with the rRNA and was prevented from migrating to it's correct position. Faint bands of approximately 4.2 - 4.4 kb in the case of pKM1del1 and 4.8 - 5.0 kb in the case of pKmM0 may represent read-through transcripts originating within the vector kanamycin resistance gene which is upstream of *repB*. These larger transcripts would include the regions encoding *repB*,

ORFs 3, 4 and 5 and *repAC*. A transcript encoding these genes is predicted from sequence data to be greater than the sum of *repB*, ORFs 3, 4 and 5 and *repA* and *repC* (3623 bp), the actual size depending upon where in the kanamycin resistance gene the transcript starts and at which point beyond *repC* it terminates (grey dotted line Figure 2.8). This size difference between these transcripts is in approximate agreement with the 711 bp difference between pKmM0 and pKmM1del1. These transcripts are not observed when ORFs 3, 4 and 5 were used to probe pKmM1del1 confirming the absence of ORFs 3, 4 and 5 on the deletion construct.

2.4 Discussion

The discovery that ORFs 3, 4 and 5 function as a plasmid stabilisation system was unexpected. The sequence of ORFs 3, 4 and 5 was determined by Dorrington and Rawlings (1990), and these authors reported that they found no significant similarity to any DNA sequence published in the GenEmbl database. Due to the proximity of genes involved in plasmid replication and the similarity of size and location to the genes E and F of RSF1010 which have been implicated in replication control (Maeser *et al.*, 1990), initial investigations were centred around the possible involvement of ORFs 3, 4 and 5 in replication. The low copy number seen for pKmM1 under conditions which selected for the plasmid suggested that ORF 3 was involved in copy number control. Initially it was assumed that *polA* mutants and *P. putida* were inefficient hosts for pTF-FC2-based replicons and that the mutation in pKmM1 effected the copy number to such an extent that the plasmid was unable to replicate in these hosts. The discovery of deletion mutants pKmM1del1 and pKmM1del2 which have normal copy numbers indicted that ORF 3 was not necessary for the maintenance of normal plasmid copy number.

The effect of the mutant pKmM1del2, in which only 104 bp were deleted, on plasmid host range and copy number was similar to that of mutant pKmM1del1 which has both ORF 3 and ORF 4 completely deleted. Since the deletion of a small region upstream of the ORFs had the same effect as deletion of the ORFs, it is probable that the upstream region contains a promoter for the ORFs. The presence of a promoter within 91 bp upstream of the start codon of ORF 3 is confirmed in Chapter 3.

Since it appeared that the ORFs were not involved in the replication control of pTF-FC2, it was decided to test the stability of the mutants. Several of the mutations in the ORF 3, 4 and 5 gene

cluster caused a loss of plasmid stability without a change in copy number, suggesting that the ORFs may be involved in plasmid stabilisation. The results indicated that the product of ORF 5 either did not play a role in plasmid stability or that its role was subtle (confirmed in section 3.3.2). The involvement of the ORFs in plasmid stabilisation was confirmed by the ability of ORFs 3, 4 and 5 to stabilise a plasmid containing an heterologous replicon. The unstable R1 derivative pOU82 was used for this purpose. Despite similarities in size and position between the ORF 3, 4 and 5 gene cluster and the EF gene cluster of RSF1010, the EF gene cluster was unable to stabilise pOU82. The presence of the ORF 3, 4 and 5 gene cluster resulted in a modest but definite 5-fold stabilisation of pOU82 in *E. coli* JM109. The stability assays indicated that only when ORF 3 and ORF 4 were both intact and expressed, was the replicon stable. Inactivation of ORF 3 without concurrent inactivation of ORF 4 resulted in a highly destabilised plasmid.

A number of plasmid encoded toxin-antidote systems exist which counteract the loss of plasmids by selectively killing plasmid-free segregants (Jensen *et al.*, 1995). Phenotypically, this phenomenon leads to plasmid stability. The results presented in this chapter could be explained if the products of ORFs 3 and 4 were a poison-antidote plasmid addiction system (proteic killer system) with ORF 3 encoding the antidote and ORF 4 the poison. A mutation in ORF 3 (pKmM1) would result in rapid plasmid loss because production of the toxin in the absence of a functional antidote would cause death of plasmid containing cells. The inability of pKmM1 to replicate in *polA* mutants and *P. putida* is probably due to increased sensitivity of these strains to the toxin compared to *E. coli* JM109. The low copy number of pKmM1 may be due to the conflicting selection pressures experienced by cells containing pKmM1. The kanamycin in the growth medium selects for cells containing the plasmid while overproduction of the toxin selects against cells containing the plasmid. The conflicting selection pressures may favour a low copy number where the conflict is minimised. A copy number of one would allow the plasmid to withstand the effect of kanamycin in the medium while minimising the gene dosage of the toxin. In strains which are able to tolerate the toxin, a gene dosage of one would allow plasmid containing cells to survive. In strains which are more sensitive to the toxin, a gene dosage of one is possibly still lethal. The effects of the ORF 3, 4 and 5 gene cluster in different strains are examined in Chapter 4. Plasmids lacking the gene for the poison (pKmM2), the genes for the poison and antidote (pKmM1del1) or the promoter for the poison and antidote (pKmM1del2) would be lost at rates which are higher than those of a plasmid with

an intact stability mechanism. The increased instability of the smaller deletion may be due to readthrough from the *repB* gene. The deletion removes the 19 C-terminal amino acids from RepB however read-through to the new termination site causes an increase in size of approximately 6.5 kDa. The deletion is within the gene for the antidote, however, the toxin gene is still intact. Readthrough from the weak *repB* promoter would produce both inactive antidote and active toxin. Due to the weakness of the *repB* promoter (LacZ fusion results not shown), overexpression of the toxin would not be as lethal as when driven by the *pas* promoter.

Amino acid sequences between different poison-antidote systems are generally poorly conserved. However, limited sequence homology has been detected between the Ccd antitoxin of plasmid F and the ParD antitoxin of plasmid R1 (Ruiz-Echevarria *et al.*, 1991). Alignment of the predicted amino acid sequences of the three Pas proteins with other plasmid proteic systems was carried out and no clear homology was detected with either the PasB toxin or the PasC protein. However, a weak but clear similarity (31% amino acid sequence identity) was detected between the PasA antitoxin of pTF-FC2 and the ParD antitoxin of the RK2/RP4 plasmids (Figure. 2.9).

```
PasA      M..LAIRLPAEVETRLEALAQATGRTKTFYAREAILEHLD..DLEDL
              : :          : :::   :.:   ::    :: :   :  :
ParD      MSRLTIDMTDQQHQSLKALAALQGKTIKQYA....LERLFPGDA.DA

PasA      ........YLAEQRLIDIRAGKTQTVPLEEVMKRY..GMEG     74
                       .:. :  :::  :  . :.      :
ParD      DQAWQELKTMLGNRINDGLAGKVSTKSVGEILDEELSGDRA     83
```

Figure 2.9 Alignment of the amino acid sequences of ParD of RK2 and PasA of pTF-FC2. Identical amino acids and conservative amino acid changes are indicated between the alignments by double and single dots, respectively.

All proteic plasmid stabilisation systems investigated to date consist of a promoter immediately upstream of the antidote and toxin genes. A promoter upstream of ORF 3 would be expected to produce a transcript of either 750 bp (ORFs 345) or approximately 2550 bp (ORFs 345, *repA* and *repC*). No such transcript was detected. Other proteic killer systems have been shown to be

self regulating (Tam and Kline, 1989a; Roberts *et al.*, 1993; Tsuchimoto *et al.*, 1988) and it is probable that this is also the case for ORFs 3, 4 and 5. The inability to detect an ORF 3, 4 and 5 transcript may be because the operon is repressed for most of the cell cycle. ORFs 3, 4 and 5 appear to be transcribed together with *repB*. It is possible that ORFs 3, 4 and 5 are transcribed constitutively at a low level from a promoter upstream of *repB,* which would ensure production of sufficient repressor to keep the promoter upstream of ORFs 3, 4 and 5 repressed. Only when the copy number falls to very low levels would the promoter upstream of ORF 3 become active. The promoter upstream of *repB* is not essential for plasmid stability since an amplified fragment containing only ORF 3, 4 and 5 was able to stabilise pOU82. The promoter may serve only to fine tune the stability system. The transcripts produced by other proteic killer systems are insufficiently studied and it is not possible to tell if this arrangement of overlapping transcripts is common. The E and F genes of RSF1010 are present on 2 transcripts (Maeser *et al.*, 1990). One originates upstream of the *repB* gene and terminates downstream of the *repC* gene. The other originates upstream of the E gene and terminates downstream of the *repC* gene. Given the similarities between RSF1010 and pTF-FC2 replication regions, it was expected that they might function similarly. This however, was not the case as shown by the inability of E and F to stabilise pOU82. Since the deletion of the promoter upstream of ORF 3 does not alter the copy number of pTF-FC2 it is probable that unlike the P_4 promoter of RSF1010 which transcribes *repA* and *repC*, the ORF 3, 4 and 5 promoter is responsible for the transcription of ORFs 3, 4 and 5 only. No role for the F gene product other than to autoregulate P_4 has, as yet, been reported.

In order to indicate that the ORFs 3, 4 and 5 encode a plasmid addiction system, these genes were renamed *pasA, pasB* and *pasC.*

Chapter 3

Regulation and control of the *pas* system

3.1 Introduction

Proteic plasmid stabilisation systems share a common gene layout consisting of a promoter followed by the gene for the antidote and finally the gene for the toxin. The systems are, however, regulated in a number of different ways. The *ccd* system of the F plasmid requires the CcdA:CcdB complex in order to autoregulate the *ccd* operon (Salmon *et al.*, 1994). In contrast, only the ParD protein is required to autoregulate the *parDE* operon of RK2 (Roberts *et al.*, 1993). The more rapid decay rate of the antidote as opposed to the toxin has been found to be due to the Lon protease in the *ccd* system, and the ClpX protease in the *phd/doc* system (van Melderen *et al.*, 1994; Lehnherr and Yarmolinsky, 1995). Most systems appear to be bacteriocidal. However, the *parD* system of R1 appears to rely on inhibiting, rather than killing, plasmid free cells (Jensen *et al.*, 1995). Because of these and other differences, the various systems cannot be assumed to be regulated in identical ways.

The *pas* system is unique in being composed of three, rather than two, genes. Since other proteic systems are able to function adequately with only an antidote and a toxin, it is probable that *pasC* serves an ancillary role. The stability of pKmM3 has shown that *pasC* may be inactivated without affecting stability of the parent replicon in *E. coli* JM105, supporting the premise that *pasC* codes for an ancillary protein. The role of the *pasC* gene, the regulatory mechanisms of the *pas* system, and how the various proteins interact to stabilise plasmids is explored in this chapter.

3.2 Materials and methods

3.2.1 General methods

DNA amplification by the polymerase chain reaction was performed using DynazymeII (Finnzymes Oy, catalogue # F500L) in a JDI 2500 thermal cycler. The following cycle was

used: (1) initial denaturation at 95°C for 2 min, (2) annealing at 52°C for 30 sec, (3) elongation at 72°C for 1 min, (4) denaturation at 95°C for 30 sec. The annealing, elongation and denaturation stages (steps 2 to 4), were repeated thirty times. All constructs made from PCR amplified DNA were sequenced as described in Chapter 2 to confirm the fidelity of amplification and cloning. Fusions to the *tac* promoter were sequenced using a primer (TACPRI) (appendix 3) complementary to the *tac* promoter. Fusions of pTF-FC2 promoters to *lacZ* were sequenced using a primer (LACZPRI) (appendix 3) complementary to the *lacZ* gene. Translation products were detected using a DNA-directed transcription-translation kit (Promega, catalogue number L1020) and the products analysed by electrophoresis using a 15 % polyacrylamide gel (Maeser *et al.*, 1990). β-galactosidase assays were performed according to the method of Miller and Stadtman (1972).

3.2.2 Growth curves

Constructs containing the *tac* promoter were transformed into *E. coli* JM105 and grown overnight at 37°C in LB containing ampicillin (100 μg/ml) before being diluted 1:100 in LB containing ampicillin. After two hours growth, IPTG was added to a final concentration of 0.5 mM in order to induce the *tac* promoter. The OD_{600} was determined at hourly intervals for a total of seven hours.

3.2.3 Protease sensitivity assay

Cells containing both a pOU82-based construct and pKG339 were grown overnight at 35°C in LB containing 30 μg/ml ampicillin and 20 μg/ml tetracycline. The cultures were then diluted 1:100 in LB containing only tetracycline and after 2 hours growth IPTG was added to a final concentration of 2mM. Samples were taken at hourly intervals and the absorbance was measured at 600 nanometers. In tetracycline resistant strains the chloramphenicol resistant construct pKGCm was used instead of pKG339. This vector was selected by using 30 μg/ml chloramphenicol.

3.3 Results

3.3.1 Polypeptides from pas A, B and C and the effect of over-expression on growth

The *pas A, B* and *C* genes were independently amplified from pTV101 (appendix 2) using primers which added *Eco*RI and *Hin*dIII restriction sites to either end of each gene. Primers 3F and 3R were used to amplify *pasA*, primers 4F and 4R were used to amplify *pasB* and primers 5F and 5R were used to amplify *pasC*. Combinations of *pasA* and *B*, *pasB* and *C* and *pasA, B* and *C* were also amplified using different pairs of the same primers (Figure 3.1).

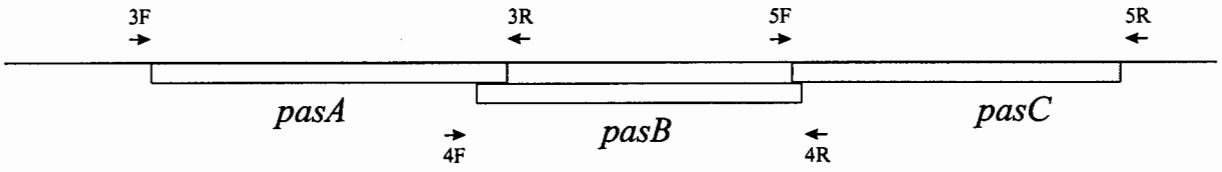


Figure 3.1 Position of primers used to amplify *pasA, B and C*. Genes are shown as white boxes, and primers as arrows. All primers ending in the letter F had *Eco*RI sites incorporated, and all primers ending in the letter R had *Hin*dIII sites incorporated.

The fragments were cloned into the *Eco*RI and *Hin*dIII sites of pKK223-3 to create transcriptional fusions pTac-pasA, pTac-pasB, pTac-pasC, pTac-pasAB, pTac-pasBC and pTac-pasABC. The polypeptides produced from these constructs using an *E. coli* derived transcription-translation system are shown in Figure 3.2. Polypeptides of about 8 kDa and 11 kDa were produced from the pTac-pasA and pTac-pasC constructs respectively (Figure 3.2 Panel A, lanes 2 and 4), but no protein band of the size predicted for PasB was observed in Panel A - lane 3 despite *pasB* being cloned downstream of the *tac* promoter. The toxin protein of poison-antidote systems is synthesised in small amounts relative to the antidote (Roberts and Helinski, 1992; Tsuchimoto *et al.*, 1992; Lehnherr *et al.*, 1993), presumably because the protein is poorly translated. The start of the *pasB* gene overlaps the end of the *pasA* gene and the ribosome binding site of *pasB* is less homologous to the consensus site and further from the AUG start than is usual for highly expressed genes. When combinations of PasA, B and C were

expressed from the vector *tac* promoter, a protein band of approximately 10.5 kDa was observed in addition to the 8 kDa band of PasA in the pTac-pasAB construct (Figure 3.2 Panel B, lane 1). The 10.5 kDa band was also visible in the expression products of pTac-pasABC (Figure 3.2 Panel B, lane 3) but not in those of pTac-pasBC (Figure 3.2 Panel B, lane 2) where the shadow-band is similar to that produced from pasC (Figure 3.2 Panel A, lane 4).

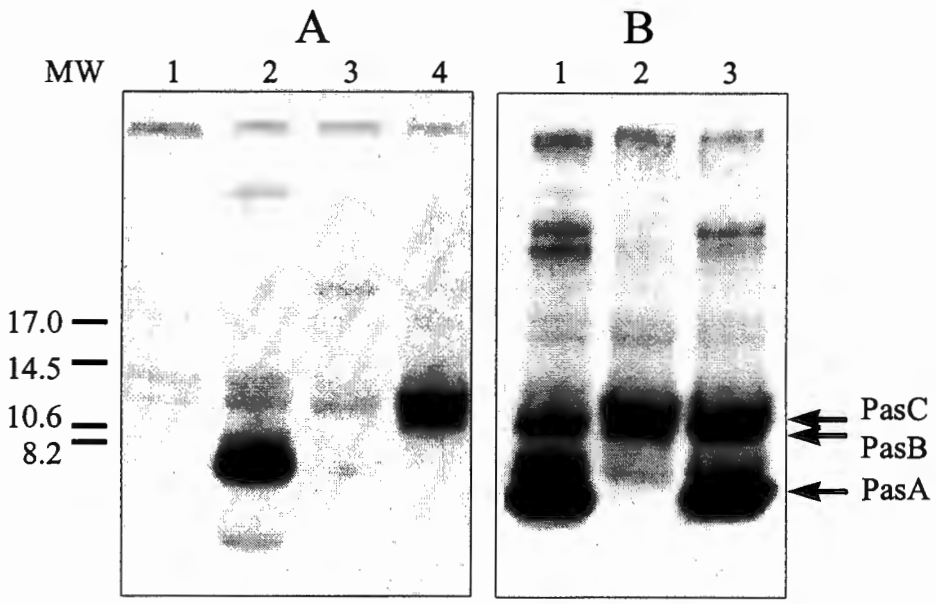


Figure 3.2 Autoradiograph of a SDS-PAGE gel analysis of polypeptides produced from *pasABC* genes under control of a *tac* promoter in an *E. coli*-derived *in vitro* transcription-translation system. Panel A: lane 1, pKK223-3; lane 2, pTac-pasA; lane 3, pTac-pasB; lane 4, pTac-pasC. Panel B: lane 1, pTac-pasAB; lane 2, pTac-pasBC; lane 3, pTac-pasABC.

It appears that translation of *pasA* is required for efficient production of PasB. Polypeptides of higher molecular weight (25-27 kDa) were produced from the pTac-pasAB (Figure 3.2 Panel B, lane 1) and pTac-pasABC (Figure 3.2 Panel B, lane 3) constructs. These proteins could possibly represent PasA-PasB fusions due to frame-shift translational read through or covalently linked PasA:PasB complexes since they were only detected in constructs in which PasA and PasB were both present (and PasB was not detectable when expressed alone).

Growth curves were conducted on *E. coli* JM105 cells containing the *tac* gene fusions so as to confirm which gene was the toxin and which the antidote. On induction of the *tac* promoter

with IPTG, growth was inhibited in *E. coli* cells containing pTac-pasB (Figure 3.3 B) but not in cells containing pTac-pasA (Figure 3.3 A) or pTac-pasC (Figure 3.3 C). This indicated that in

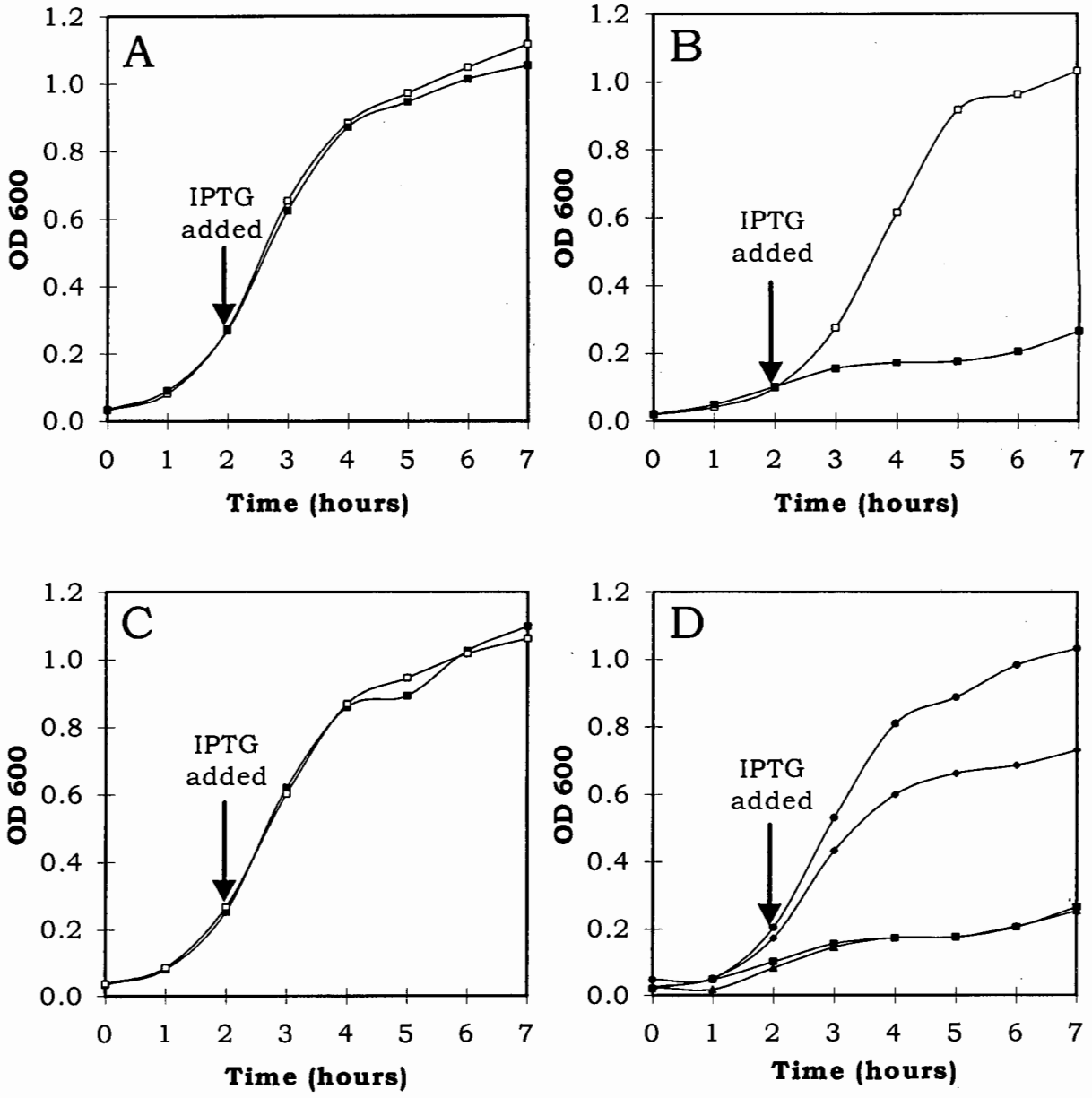


Figure 3.3 Growth curves of *E. coli* cells containing one or more of the *pas* genes under the control of the *tac* promoter. Un-induced cells are indicated by curves marked (□) and induced cells by curves marked (■). **A**, *E. coli* (pTac-pasA); **B**, *E. coli* (pTac-pasB); **C**, *E. coli* (pTac-pasC); and **D**, (▲) induced *E. coli* (pTac-pasB), (◆)induced *E. coli* (pTac-pasAB), (■) induced *E. coli* (pTac-pasBC) and (●) induced *E. coli* (pTac-pasABC) cells. Each data point is the mean of three separate experiments.

spite of the inability to detect PasB expression from pTac-pasB *in vitro,* on IPTG induction sufficient PasB was produced *in vivo* to inhibit growth of pTac-pasB containing *E. coli* cells. When pTac-pasAB was present, the toxic effect of the *pasB* gene product was moderated but not completely removed by the presence of *pasA*, while the presence of *pasC* (pTac-*pasBC*) had no affect on reducing toxicity (Figure 3.3 D). When all three genes (pTac-*pasABC*) were present the toxic effect of *pasB* was almost fully relieved to the same level as uninduced cells (Figure 3.3 D). From these results it was inferred that PasA was an antidote, PasB a toxin and that PasC enhanced the ability of PasA to neutralise the toxic effect of PasB.

3.3.2 The role of PasC in moderating the effect of PasB

In the previous section it became apparent that cells containing pTac-pasAB did not grow as efficiently as pTac-pasABC. In order to confirm that this enhanced growth rate is due to presence of PasC, two further experiments were performed. Growth curves were carried out using *E. coli* JM105 cells containing the *pasAB* genes under control of the *tac* promoter and the *pasC* gene *in trans* also under control of the tac promoter. Due to the possibility of incompatability between replicons and the need for different selectable markers, the genes supplied *in trans* were provided on a pACYC184-based construct. The construction of these constructs is described in section 3.3.4. As can be seen from Figure 3.4, when pTac-pasC-pACYC was placed *in trans* to pTac-pasAB, the growth rate was significantly increased over that observed when only pACYC was present *in trans*. The copy number of pACYC-based replicons is far lower than that of ColE1-based replicons, thus the gene dosage *pasC* would be far lower than that of *pasAB*. To ensure that the different replicons and the different gene dosages do not cause some artefactual increase in growth rate, the experiment was repeated after swapping the replicons. When cells containing pTac-pasC were grown with pTac-pasAB-pACYC *in trans*, a higher growth rate was seen than when cells contained pKK223-3 with pTac-pasAB-pACYC *in trans*. This confirmed that PasC was able to increase the growth of PasA and PasB containing cells and that the higher the *pasAB* gene dosage, the lower the growth rate of the host cells.

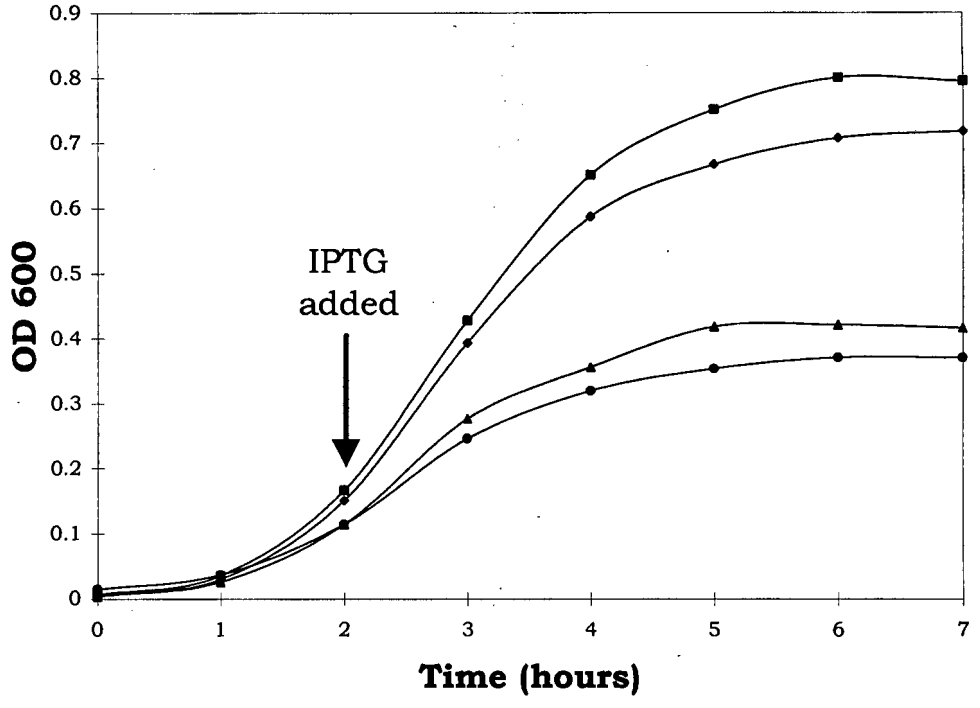


Figure 3.4 Effect of *pasC* on the growth rate of cells containing plasmids expressing *pasAB*. *E. coli* [pTac-pasAB, pACYC] cells (control) are marked (●); *E. coli* [pTac-pasAB, pTac-pasC-pACYC] (◆); *E. coli* [pKK223-3, pTac-pasAB-pACYC] (control)(▲); *E. coli* [pTac-pasC, pTac-pasAB-pACYC] (■) Each data point is the mean of three separate experiments.

To further confirm these findings, the *pas* genes were amplified from pKmM3 using the primers 3F and 5R. The amplification product was digested with *Eco*RI and *Hin*dIII and cloned into pKK223-3, which had been digested with *Eco*RI and *Hin*dIII, to create pTac-pasABC*. This construct is identical to pTac-pasABC except that it contains a T inserted in *pasC* which causes a frame shift and inactivates the gene. The cells containing the construct pTac-pasABC show greatly increased growth, compared to cells containing pTac-pasABC*, confirming the role of *pasC* in increasing the growth rate of *pasAB* containing cells (Figure 3.5). The increased growth rate and cell density at stationary phase observed for *E. coli* [pTac-pasAB-pACYC, pKK223-3] cells compared to *E. coli* [pTac-pasAB, pACYC] cells suggests that the higher levels of PasA and PasB caused by higher gene dosages in the ColE1-based vector may be deleterious to the cells.

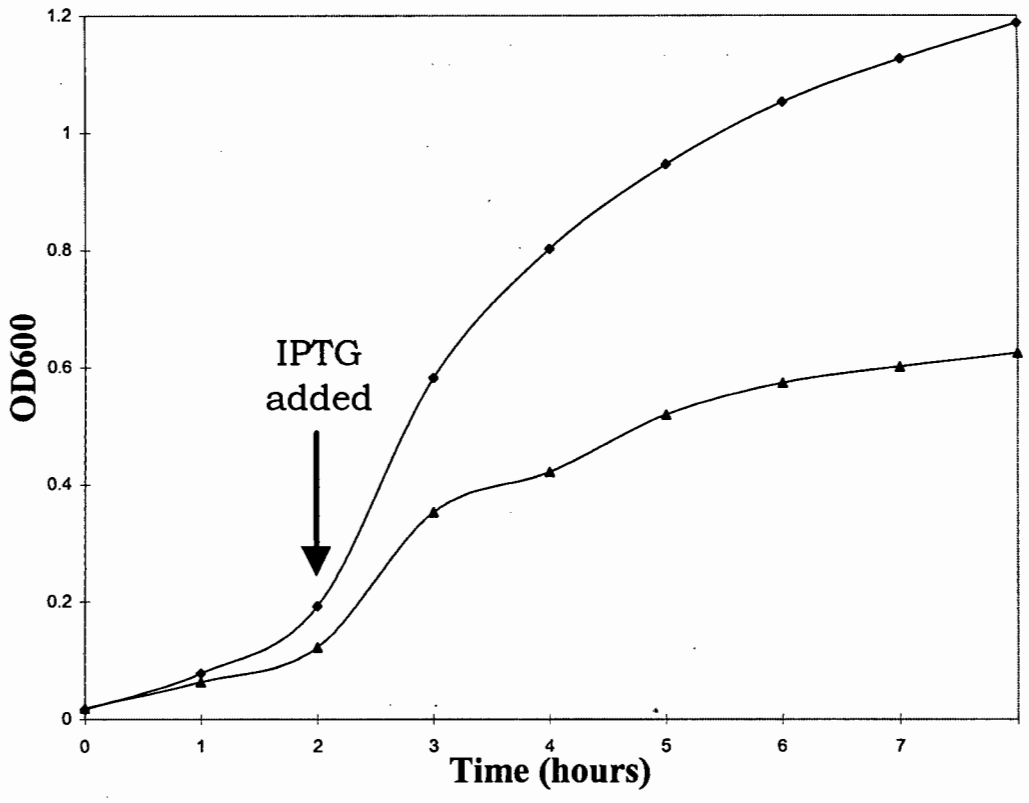


Figure 3.5 Comparison of the growth rate of *E. coli* JM105 cells containing pTac-pasABC and pTac-pasABC*. Data points for cells containing pTac-pasABC are shown as (◆) and pTac-pasABC* as (▲). Each data point is the mean of three separate experiments.

3.3.3 PasB - bacteriostatic or bacteriocidal ?

To determine whether PasB was bacteriocidal or bacteriostatic the construct pTac-pasB was grown in the presence of selection, either with or without IPTG induction, and the effect of *pasB* induction on the number of plasmid-containing cells was observed. If PasB was bacteriostatic, on induction of *pasB,* the number of plasmid-containing cells would be expected to remain constant over time as cells would be prevented from dividing but would not be killed. Only when the antibiotic became depleted would 'break out' of plasmid-free cells occur. If PasB was bacteriocidal, there would be a sharp decrease of plasmid-containing cells on *pasB* induction.

In induced cultures, the total number of plasmid-containing cells fell rapidly from 10^8 to 10^3 within an hour due to cell death caused by PasB thus confirming the bacteriocidal nature of

PasB, (Figure 3.6). No noticeable toxic effect was observed in cells in which *pasB* was not induced and the number of plasmid-containing cells increased from 10^8 to 10^{10} over 5 hours.

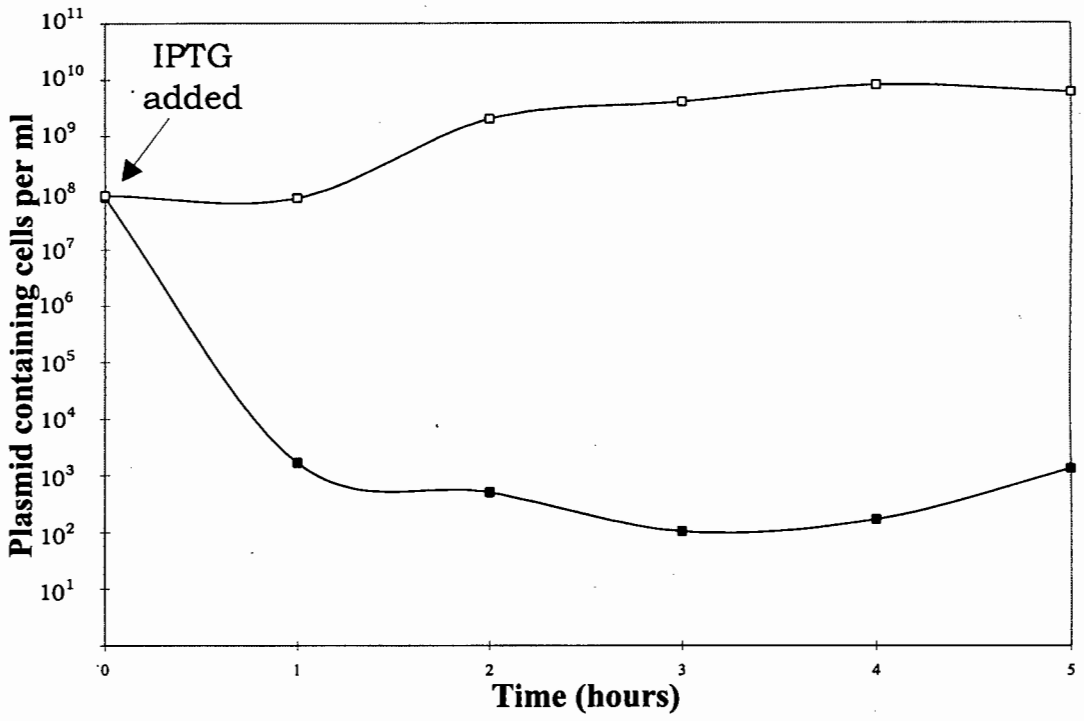


Figure 3.6 The effect of PasB on cell viability. Plasmid retention in uninduced cells is indicated by (□) and in cells in which *parB* was induced by (■) symbols. Each data point is the mean of three separate experiments.

3.3.4 The effect of PasA, B, and C on expression of promoters 1, 2 and 3

To investigate any possible involvement of *pasA*, *B* and *C* in the regulation of pTF-FC2 replication as well as stability, the regions upstream of *repB*, *pasA* and *repAC* were translationally fused to a reporter β-galactosidase gene. These experiments were performed before the discovery of mutants which had *pasA, B* and *C* deleted, the results presented here confirm that *pasA, B* and *C* are not involved in regulating replication. The region upstream of *repB* (nts 1 to 257) was amplified from pTV101 using the primers -1212F and FP1. Primers -1212F and FP2 were used to amplify nucleotides 1240 to 1362 upstream of *pasA* from pTV400 (appendix 2), and primers -1212F and FP3 amplified nucleotides 1854 to 2162 upstream of *repA* from pTV4101 (appendix 2). In each case the amplification products were digested with *Eco*RI which cut within the amplified vector MCS and *Bam*HI which cut at a site introduced by

the primers FP1-3. The digested products were cloned into the *Eco*RI and *Bam*HI sites of pMC1403 so as to create the three in-frame translational fusion reporter plasmids. The *lacZ* fusions are pP1H (nts 1 - 257, the region upstream of *repB*), pP2H (nts 1240 - 1362, the region upstream of *pasA*) and pP3H (nts 1854 - 2162, the region upstream of *repA*)(see Figure 3.7). The constructs produced blue colonies when transformed into *E. coli* JM109 and plated on LA containing ampicillin and X-Gal indicating that the amplified fragments contained promoter elements. The three site-directed mutant plasmids and the intact replicon were introduced *in trans* to the promoter constructs and the effect of each of the mutations on β-galactosidase expression was determined.

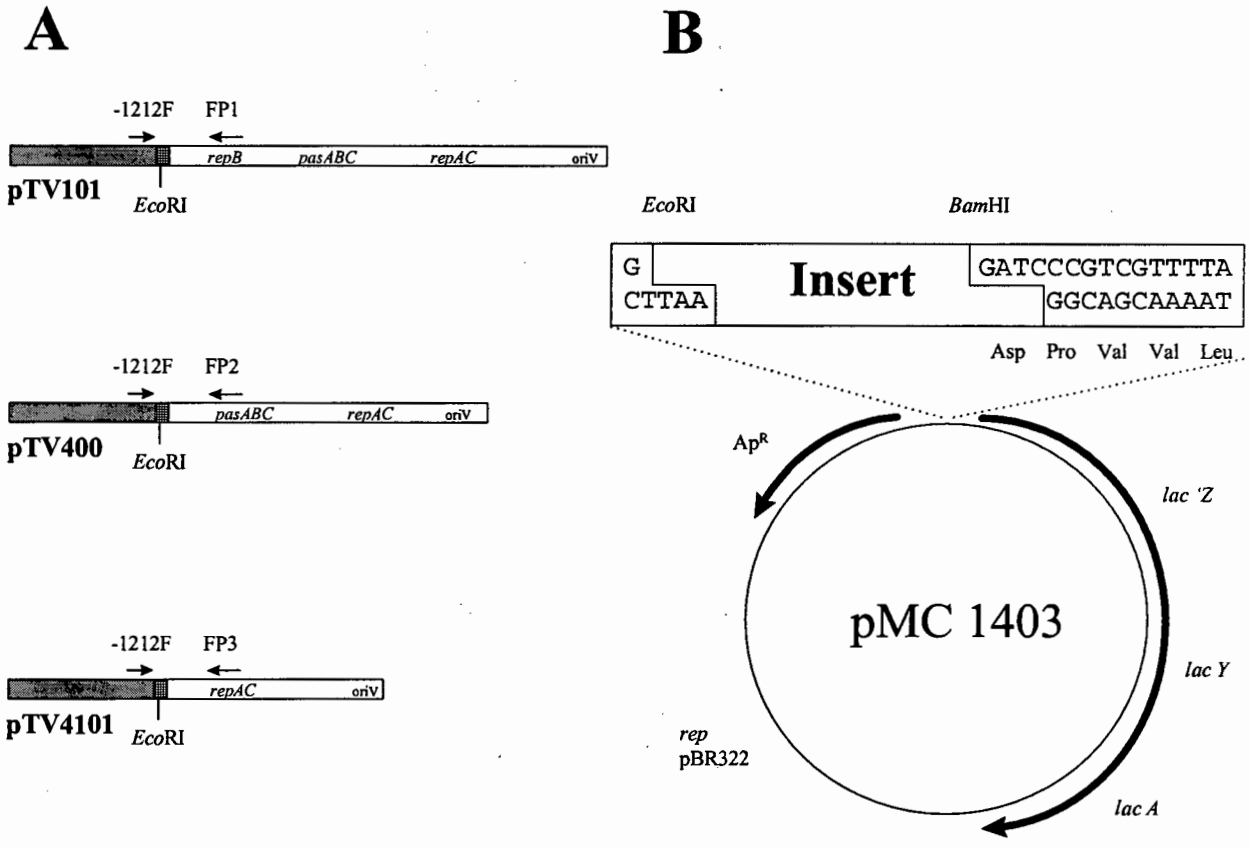


Figure 3.7 Cloning of pTF-FC2 promoters into pMC1403. Panel A shows pTV101, pTV400 and pTV4101 and the orientation of the primers used to amplify the promoter regions. The primers are shown as arrows, the pUC19 vector as a grey box, the multiple cloning site of pUC19 as a hatched box and the pTF-FC2 replicon as a white box. Panel B shows pMC1403 and the orientation of the insert.

Table 3.1 Effects of *pasABC* site-directed mutant plasmids on the promoters of the minimal replicon of pTF-FC2.

lacZ plasmid	Coresident plasmid	Inactive gene	Average β-galactosidase expression (Miller units)
pP1H	(none)		57 ± 4
pP1H	pKmM0	none	61 ± 4
pP1H	pKmM1	*pasA*	71 ± 3
pP1H	pKmM2	*pasB*	62 ± 8
pP1H	pKmM3	*pasC*	46 ± 3
pP2H	(none)		1555 ± 575
pP2H	pKmM0	none	69 ± 23
pP2H	pKmM1	*pasA*	755 ± 515
pP2H	pKmM2	*pasB*	158 ± 59
pP2H	pKmM3	*pasC*	86 ± 27
pP3H	(none)		10 ± 4
pP3H	pKmM0	none	18 ± 9
pP3H	pKmM1	*pasA*	6 ± 4
pP3H	pKmM2	*pasB*	11 ± 3
pP3H	pKmM3	*pasC*	14 ± 5

The results of various combinations of mutant plasmids and promoters are shown in Table 3.1. The promoter upstream of *repB* exhibited modest strength (57 Miller units) and was largely unaffected by inactivation of *pasA, B* or *C*. The promoter upstream of *repA* was very weak (10 Miller units) and also did not respond to the inactivation of any of the *pas* genes. Despite showing considerable scatter, the results for the *pas* promoter upstream of the *pas* genes showed that not only was the promoter very strong (1555 Miller units) but it also appeared to be negatively regulated by *pasA*. Inactivation of the *pasA* gene caused a 10-fold increase in expression levels. The inactivation of the *pasB* gene caused only a two-fold increase in expression from the *pas* promoter. While inactivation of the *pasC* gene had no significant effect on the level of β-galactosidase expression.

Each of the mutant plasmids supplied *in trans* also contained the *pas* promoter which in turn must also be regulated by the *pas* gene products. This complex state of affairs makes it difficult to accurately determine the effects of each of the *pas* genes. By supplying the *pas* genes under control of the *tac* promoter, any negative feedback is bypassed. However, the pP2H and the pTac-pas constructs both use ColEI origins of replication and are both ampicillin resistant. The pTac-pas fusions were excised using *Pvu*I and *Bam*HI, the *Pvu*I site was then blunted and the fragment subcloned into pACYC184 which had been cut with *Cla*I (blunted) and *Bam*HI. This avoided plasmid incompatability when the pTac-pas constructs were placed *in trans* to pP2H and also provided a different selectable marker. These constructs, pTac-pasA-pACYC, pTac-pasB-pACYC, pTac-pasC-pACYC, pTac-pasAB-pACYC, pTac-pasABC-pACYC were transformed into *E. coli* JM105 cells containing pP2H. Although the levels of β-galactosidase expression were again found to have considerable scatter, the results were clear (Table 3.2).

Table 3.2 Auto-regulation of the *pasABC* genes.

lacZ plasmid	Coresident plasmid	Average β-galactosidase expression (Miller units)	Percentage activity
pP2H	pACYC184	348 ± 205	100%
pP2H	pTac-pasA-pACYC	146 ± 142	42%
pP2H	pTac-pasB-pACYC	*	
pP2H	pTac-pasAB-pACYC	34 ± 9	10%
pP2H	pTac-pasABC-pACYC	95 ± 11	27%

* lethal construct could not be tested.

The level of β-galactosidase activity decreased from 348 to 146 units when the *pasA* gene was provided *in trans* on its own. This repression was enhanced (34 units) when *pasAB* or when *pasABC* (95 units) was present. The *pasABC* promoter, therefore, appears to be autorepressed (2.5-fold) by PasA and this repression increased when both PasA and PasB were present (10-fold). Regulation by *pasB* or *pasBC* could not be tested due to the lethal effects of PasB in the absence of PasA. It is possible that moderation of the repression shown by PasA and PasB combined when PasC was also present may be caused by PasC binding to a PasA:PasB complex. This may make the complex a less effective repressor of the *pasABC* operon but may have the advantage that the PasA:PasB complex is more stable in the presence of PasC and

therefore is less toxic to the cells. This is consistent with the effect of PasC shown earlier (Figures 3.3 and 3.4)

3.3.5 The role of the Lon protease in plasmid stabilisation

The pOU82/pKG339 conditional replication system (Jensen *et al.*, 1995) was used to determine which *E. coli* host protease is involved in the selective degradation of the PasA antidote. When pKG339 is provided *in trans* to pOU82, replication of pOU82 can be halted by the addition of IPTG.

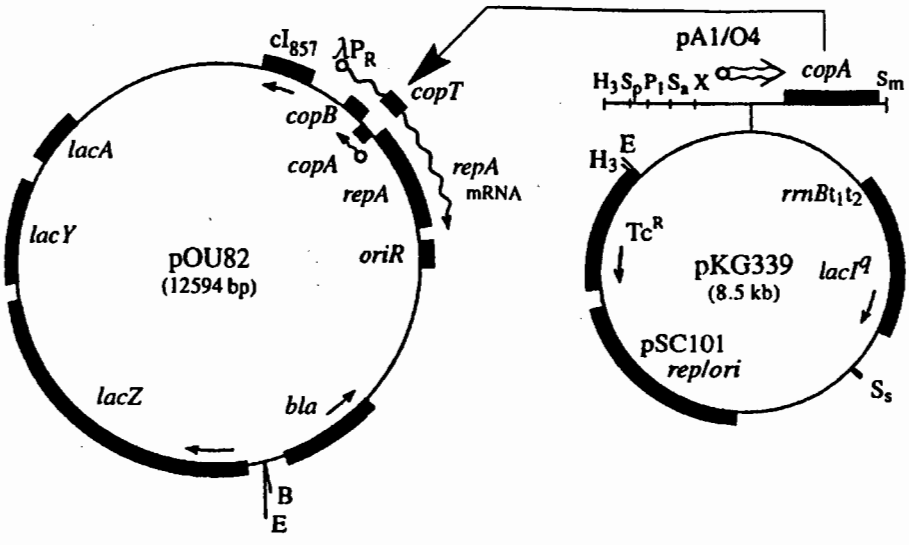


Figure 3.8 The pOU82/pKG339 conditional replication system from Jensen *et al.* (1995). Plasmid pKG339 is a pSC101 derivative that contains the *copA* gene downstream of the strong LacI-regulated pA1/04 promoter. Thus, pKG339-containing cells produce large amounts of CopA RNA upon addition of IPTG to the growth medium. Tc^R denotes the tetracycline-resistance gene and *rrnBt1t2* is the tandem pair of translational terminators from the *E. coli* *rrnB* operon. Plasmid pOU82 is a mini-R1 derivative containing the λPr promoter in front of the replication-control region. The temperature-sensitive lambda repressor encoded by the λ*c*I 857 allele represses the λPr promoter at 30°C. At 42°C the λPr promoter is derepressed, resulting in uncontrolled run-away replication. At 30°C, the copy number of pOU82 is the same as that of wild-type mini-R1 replicons (Larsen *et al.*, 1984). The replication-control elements of R1 are shown. *copT* denotes the target of the CopA antisense RNA in the *repA* mRNA (shown as a wavy arrow). Arrows inside the circles denote directions of transcription. *bla* denotes the ampicillin-resistance gene.

Figure 3.8 continued. Transcription of the *lacZYA* operon of pOU82 is driven by a truncated derivative of the *deoP2* promoter (Gerdes *et al.*, 1985). Restriction sites are shown, as follows: E, *Eco*RI; B, *Bam*HI, H3, *Hin*dIII; Sp, *Sph*I; P1, *Pst*I; Sa, *Sal*I; Sm, *Sma*I; and Ss, *Ssp*I.

When pOU82 contains a proteic plasmid stabilisation system, halting replication will result in plasmid free cells containing the toxin:antidote complex. The protease which is responsible for the selective degradation of the antidote will then digest the antidote and release the toxin resulting in cell death and a reduction in cell growth rate. In mutants deficient in the specific protease, the antidote will not preferentially decay and no cell death will occur. The Lon mutant strain SG22095 contains a tetracycline resistance marker and thus it was not possible to select for pKG339 in this strain as pKG339 requires tetracycline for selection (Figure 3.8). The replication control cassette of pKG339 was excised using *Xho*I and *Hin*dIII and ligated into pACYC184 digested with *Sal*I and *Hin*dIII to create pKGCm (The *Xho*I site lies adjacent to the *Ssp*I site shown in figure 3.8) . The ability of this construct to halt replication of pOU82 was tested by transforming this construct and pOU82 into *E. coli* SG22025, the parent strain of both the Lon and ClpX mutants. After overnight growth at 35°C with selection for both pOU82 and pKGCm, the transformants were then plated to LA containing either chloramphenicol or chloramphenicol and IPTG (2mM). Cells plated to IPTG-containing plates produced white colonies confirming the absence of pOU82, while cells plated to LA lacking IPTG gave blue colonies indicating the presence of pOU82.

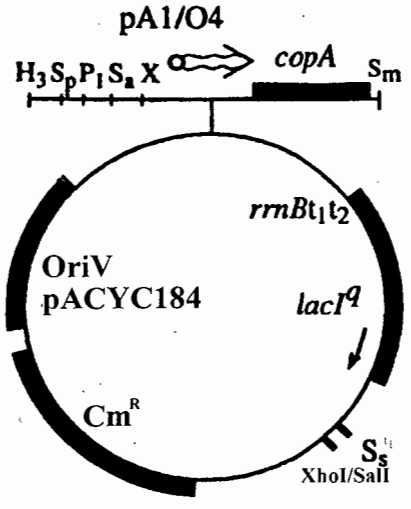


Figure 3.9 Plasmid map of pKGCm. Adapted from Gerdes *et al.*, 1985.

Panel B of Figure 3.10 shows how the growth of the ClpX mutant strain SG22093 was retarded when the replication of pOU-PTF was halted. This growth retardation was also observed in the parent strain SG22025 when pOU-PTF replication was halted (data not shown). Panel A shows that when replication of pOU-PTF was halted in the Lon mutant, no growth retardation was observed when compared to pOU82. This data suggests that the Lon protease is responsible for the selective degradation of the *pasA* antidote.

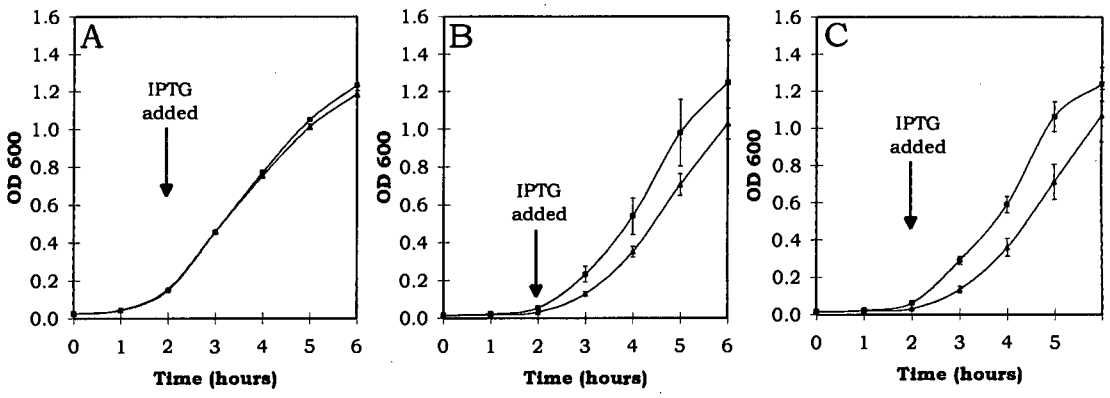


Figure 3.10 Growth curves of protease mutants. Panel A shows the growth curve of Lon mutant SG22095 containing either pOU82 and pKGCm (control)(■) or pOU-PTF and pKGCm (▲). Panel B shows a growth curve of the ClpX mutant SG22093 containing either pOU82 and pKGCm (control)(■) or pOU-PTF and pKGCm (▲). Panel C shows a growth curve of the parent strain SG22025 containing either pOU82 and pKGCm (■) or pOU-PTF and pKGCm (▲). Each data point is the mean of three separate experiments.

To confirm the involvement of Lon, the stability of pOU-PTF was tested in both the ClpX and Lon mutant strains as well as the *E. coli* SG22025 parent strain. In the parent strain, pOU-PTF showed a loss frequency of 1.4×10^{-2}, in the ClpX mutant it showed a loss frequency of 1.5×10^{-2} while in the Lon mutant the loss frequency was 3.5×10^{-2}. This comparatively small loss of stability seen in the Lon mutant strain supports the notion that Lon is at least in part responsible for the degradation of PasA. The level of stabilisation shown by pOU-PTF in these strains is negligible compared with the results obtained in *E. coli* JM105. It should be noted that the level of stabilisation was highly dependent on the strain used. What is important is relative difference in stabilising ability of pOU-PTF between these strains which are isogenic except

for the mutation in their proteases. The effect of strain differences on the effectiveness of the *pasABC* system is explored in depth in Chapter 4.

3.4 Discussion

In this chapter supporting evidence that *pasA* encodes the antidote and *pasB* encodes the toxin of a proteic plasmid stabilisation system has been presented. The role of *pasC* in moderating the toxic effects of PasA:PasB complexes has also been shown.

The use of *tac* fusions to each of the *pas* genes allowed the effect of each gene product to be studied in isolation. Only PasB effected cell growth when expressed by itself. PasB appeared to act by killing cells, although its target is unknown. Over-expression of PasB did not cause induction of the SOS response and thus does not affect GyrA as is the case for CcdB of the *ccd* system (data not shown). When combinations of the *pas* genes were expressed it became evident that while *pasA* was able to partially neutralise *pasB*, *pasC* was unable to exert any effect on *pasB* by itself. Full neutralisation of PasB required not only PasA but also PasC. The involvement of a third gene in a proteic stabilisation system is unique. Since *pasC* by itself or in combination with *pasB* did not relieve inhibition of growth by PasB, it is probable that PasC acts by stabilising the interaction of PasA and PasB rather than by modifying the target of PasB or enhancing growth. Plasmid stability data of pKmM3 indicated that although PasC may moderate the toxic effects of PasA:PasB complexes, its effect on plasmid stability is negligible or can only be seen over growth periods of greater than 100 generations. PasC may assist PasA in preventing cell death or growth retardation in plasmid containing cells. Since the presence of PasC reduced PasA:PasB toxicity, *pasA, pasB* and *pasC* containing cells would be expected to outgrow cells which contained *pasA* and *pasB* alone. As shown in Figure 3.5 this does occur when the *pas* genes are expressed from the *tac* promoter. The growth rate of cells containing the unmutated pTF-FC2 replicon, pKmM0 also show faster growth rates than cells containing the *pasC* frame shift mutant pKmM3 (data not shown). Since PasA and PasB are probably not co-valently linked, there must be a dissociation equilibrium during which PasB will exist in an active form. It is speculated that PasC reduces the dissociation of the PasA:PasB complex and that this in turn reduces the toxicity of PasB and prevents inappropriate killing. Further studies should focus on cross-linking PasA and PasB in order to prove that they do bind to each other and to PasC. If possible the toxicity of cross-linked PasA:PasB and PasA:PasB:PasC should be

compared to that of PasB by itself, however, in order to quantify toxicity a better understanding of the target of PasB will be required.

Proteic killer systems are known to be autoregulated at the level of transcription. However, since it was not possible to isolate sufficient mRNA to determine exact starting point of the *pas* transcript, a translational fusion was constructed to the region upstream of the *pas* genes. This approach also ensured that should any activator sequences or protein binding sites be present in the region of the promoter, they would not be disrupted. These reporter gene fusions would indicate the cumulative effect of both transcriptional and translational regulation and thus provide a better measure of how the levels of proteins produced varied in accordance with the presence or absence of regulators than transcriptional fusions alone. While PasA partially repressed the *pas* operon, PasB was required for full repression. The apparent derepression caused by the addition of PasC suggests that a complex of PasA:PasB:PasC does not bind to the site of regulation as well as a PasA:PasB complex. Alternatively PasA and PasB may bind independently to the regulatory site. By stabilising PasA:PasB complexes, PasC would reduce the amount of free PasA and PasB available for binding to the regulatory site. This would de-repress the operon giving rise to the slightly increased expression seen when PasC was supplied in addition to PasA and PasB. The *pas* genes appear not to be involved in regulating any of the replication genes. The scatter seen when high levels of β-galactosidase are expressed was possibly due to plasmid instability caused by the toxicity of β-galactosidase at high concentrations. While using a low copy number promoter probe vector may reduce this toxicity, other disadvantages were found to occur. Low copy number pACYC-based derivatives of pMC1403 were constructed initially in this study. However, the levels of β-galactosidase were shown to vary considerably when ColEI-based constructs was introduced *in trans*. This variation was due to the reduction in copy number of the pACYC-based promoter probe vector (data not shown). It was unclear whether this reduction in plasmid copy number and hence in β-galactosidase level was due to the type of replicon present *in trans* or due to the size of the cloned insert. It was thus decided to use the widely used and documented pMC1403 vector despite the associated scatter in β-galactosidase activities obtained.

The cornerstone of proteic killer systems is that the toxin and the antidote should have different decay rates. Previous studies into the proteases responsible for antidote degradation in other systems have used temperature dependent replicons. A shift to a non-permissive temperature

has been used to cure cells of the plasmid containing the killing system. This, however, also causes changes in cell growth rate. The pOU82/pKG339 system has previously been used to cure cultures of plasmid and thus mimic the effects of natural plasmid loss. This system does not rely on a temperature shift with the associated change in growth rate. Furthermore it allows for a more rapid curing and does not require any special strains. Only in the Lon mutant strain was growth retardation of cells which were cured of pOU-PTF, not apparent. This indicated, that Lon may be required to degrade PasA for post-segregational killing to occur. The induction of *phd*-mediated killing in plasmid P1 caused a complete halt in increase in culture optical density (Lehnherr and Yarmolinsky, 1995). The *pas* system merely retarded growth in the strains in which it was used. This difference is due either to the lack of sensitivity of the strains used to the PasB toxin or the reduced availability of the active toxin in these strains. This is supported by the inherently low level of stability seen for pOU-PTF in both the parent and ClpX mutant strains. The approximately two-fold reduction in stability of pOU-PTF seen in the Lon mutant corroborated the growth inhibition experiments and supports the conclusion that Lon is required for *pas-mediated* stability. A low level of PasB toxicity in these strains probably accounts for the inability of the *pas* genes to stabilise pOU82 more than two-fold. The difference in strain sensitivity to the PasB toxin is explored in chapter 4.

Chapter 4

The effects of strain variation on the *pas* system

4.1 Introduction

An often forgotten component of proteic plasmid stabilisation systems is the host cell. The systems rely on host factors such as polymerases for transcription, ribosomes for translation and proteases for the degradation of both toxin and antidote. The ability of the polymerases of different hosts to recognise the promoter sequence of *pas* and hence express the *pas* proteins is expected to vary as is the association constant of the internal protease (e.g. Lon) for the antidote. The growth medium used has also been shown to effect the degree of stabilisation conferred by proteic stabilisation systems (Roberts *et al.*, 1994; Jensen *et al.*, 1995). These and other as yet undetected factors all influence how well the plasmid is stabilised in a particular host. Roberts and Helinski (1992) have shown that while the *parD* locus of RK2 is able to stabilise plasmids in a wide range of organisms, the degree of stabilisation varies considerably between the different hosts. *E. coli* JM107 containing a plasmid with the *parDE* genes showed a loss rate of 0.30 % per generation. *E. coli* JM109 with the same plasmid showed a zero percent loss rate while *E. coli* GM1859 showed a 1.26% loss rate. Strain differences in the stabilising effect of the *pem* system have been observed by Tsuchimoto and Ohtsubo (1989).

There has however been no systematic attempt other than the work of Roberts and Helinski (1992) to study which differences between strains are responsible for the different rates of stabilisation. This chapter attempts to elucidate which host dependent functions most influence the stabilising ability of the *pas* system in strains of *E. coli*. The results presented here apply only to the *pas* system, other proteic systems are likely to interact differently with their hosts and thus are likely to be affected by different host elements.

4.2 Results

4.2.1 Stability in different hosts

Table 4.1 shows loss frequencies for pOU-PTF in a number of *E. coli* strains. While the *pas* genes could stabilise pOU82 approximately 100-fold in *E. coli* CSH50, it was capable of only approximately 5-fold stabilisation in *E. coli* JM105. The *pas* genes could not stabilise the plasmid in either *E. coli* JM107 or JM109. The control plasmid pOU82 had a similar stability in *E. coli* CSH50, JM105 and JM107, however, in JM109 it was ten-fold more stable. Of the strains tested *E. coli* JM107 is isogenic to JM109 except that JM107 is *recA*$^+$ and JM109 is *recA*$^-$. The *recA* system may therefore have an effect on the inherent stability of pOU82. This is supported by the work of Roberts and Helinski (1992) who found mini-RK2 plasmids to be three-fold more stable in *E. coli* JM109 than in *E. coli* JM107.

Table 4.1 The effect of the *pas* locus in different hosts.

E. coli strain	Plasmid	Loss Frequency (LF)
CSH50 ara, Δ(lacpro), strA, thi	pOU82	2×10^{-2}
	pOU-PTF	3×10^{-4}
JM105 endA1, thi, rpsL, sbc15, hsdR4, Δ(lac-proAB),[F`, traD36, proAB, lacI^{q}ZΔM15]	pOU82	2×10^{-2}
	pOU-PTF	9×10^{-3}
JM107 endA1, thi, gyrA96, hsdR17, (r_k^-, m_k^-), relA1, supE44, Δ(lac-proAB), [F`, traD36, proAB, lacI^{q}ZΔM15]	pOU82	5×10^{-2}
	pOU-PTF	5×10^{-2}
JM109 endA1, recA1, thi, gyrA96, hsdR17, (r_k^-, m_k^-), relA1, supE44, Δ(lac-proAB), [F`, traD36, proAB, lacI^{q}ZΔM15]	pOU82	5×10^{-3}
	pOU-PTF	4×10^{-3}

4.2.2 Effect of host on PasA, PasB and PasC interaction

The effect of the expression each of the *pas* genes on the growth of *E. coli* JM105 after induction with IPTG was also examined in *E. coli* CSH50, JM107 and JM109. While the JM strains all contain the lacI^q gene, CSH50 does not. The F′ episome was therefore transferred from *E. coli* JM109 to CSH50 (Pro$^-$) by conjugation to create the Pro$^+$ lacI^q auxotroph CSH50-

I^q which was used in these experiments. As can be seen from Figure 4.1 the growth rates of bacteria was highest in cells containing the control plasmid pKK223-3 irrespective of the strain used. The PasB toxin was able to retard growth in all strains tested but this retardation was most severe in *E. coli* CSH50-I^q. In all strains tested PasB was able to retard growth, with the effect being most noticeable in *E. coli* CSH50-I^q and least noticeable in *E. coli* JM109 with *E. coli* JM107 and JM105 showing an intermediate level of retardation. When PasA and PasB were both present, PasA significantly counteracts the growth retardation exerted by PasB in all strains except *E. coli* CSH50-I^q which was still severely retarded. When all three of the *pas* genes were present the toxic effects of PasB were relieved further in all strains except *E. coli* JM107 which showed a similar growth rate to that obtained when only PasA and PasB were present. Despite the moderating effect of PasC, the growth rate and stationary phase cell density of *E. coli* CSH50-I^q cell containing all the *pas* genes was far lower that that obtained when the control pKK223-3 was used.

4.3 Discussion

Initial experiments to test the stabilising influence of the *pas* genes on pOU82, when conducted in *E. coli* JM105 revealed an apparently ineffective system providing only 5-fold stabilisation. In order to allow these results to be directly comparable with those obtained by Jensen and co-workers (1995), the experiments were repeated using *E. coli* CSH50. The 100-fold stabilisation conferred by the *pas* genes in CSH50 suggested that strain differences play a very important role in determining how effective the *pas* system is at stabilising a plasmid. The inability of *pas* to function in *E. coli* JM107 and JM109 as well as the very poor level of stabilisation seen in the Lon protease mutant and its parent strain (section 3.3.5) further confirmed the importance of strain differences.

It is possible that the strength of the *pas* promoter could vary in the different strains. However, different levels of transcription from the *pas* promoter would not alter the normal concentration of toxin and antidote within the cell on which stability depends. A reduced rate of transcription would increase the time required to achieve the normal cellular concentration of toxin and antidote. Only when the rate of transcription was so slow that the normal concentration of toxin and antidote was not reached before cell division occurred, would there be any potential effect on stability. Even if toxin and antidote concentrations were sub-optimal, the effect on stability

may be very small indeed. A more rapid rate of transcription would merely speed up the time taken to achieve self-regulating levels of antidote and toxin.

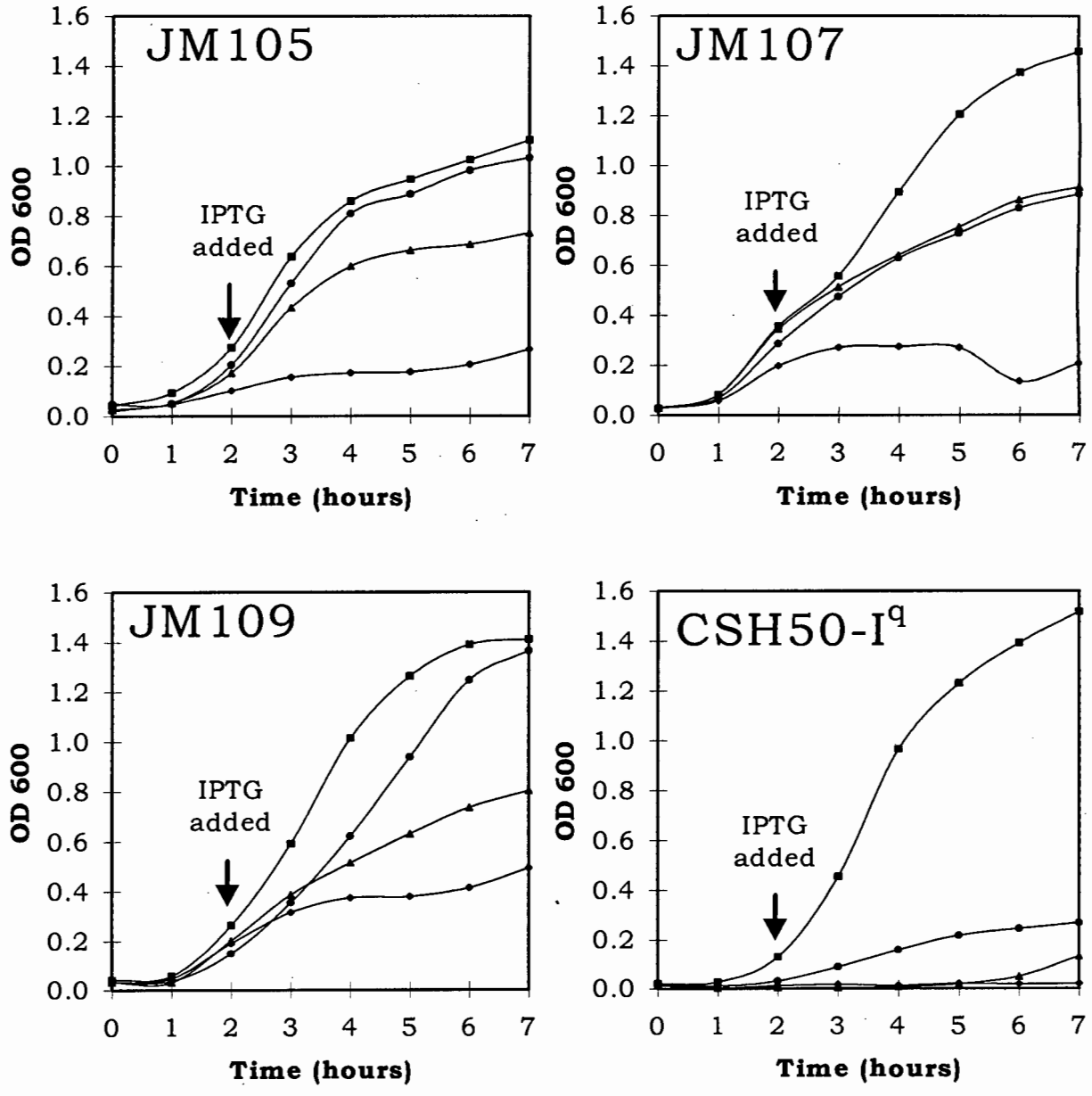


Figure 4.1 Growth curves of various *E. coli* strains over-expressing the *pas* genes. Each graph shows pKK223-3 (control)(■); pTac-pasB (◆); pTac-pasAB (▲); and pTac-pasABC (●). Each data point is the mean of three separate experiments.

The effect of various combinations of the *pas* genes on the growth of the host strains provided more insight into why *pas-mediated* plasmid stability varied in different strains. While the overall growth rate and stationary phase cell density of CSH50-I^q containing different

combinations of the *pas* genes was far lower that that of the other strains tested, it appeared that PasA did not effectively neutralise PasB in the absence of PasC and that neutralisation of the toxin was poor even in the presence of PasC. PasB has been shown to be highly toxic even in strains in which the *pas* locus was unable to stabilise pOU82. It may be that the increased stability of pOU-PTF in CSH50 was due to greater toxicity of PasB in this strain. An alternative explanation for the increased stability may be that in *E. coli* CSH50, the interaction between PasA and PasB is relatively weak even when supported by PasC. This would ensure that in plasmid free cells PasA would be able to be degraded as it was less tightly bound to PasB and therefore more available for preferential degradation by the Lon protease. A facet of the interaction of toxin and antidote in proteic systems not extensively studied is whether the protease is able to degrade antidote when bound to the toxin. It is possible that the host protease requires unbound or loosely bound antidote as a substrate and the stronger the toxin antidote interaction the less the probability that the antidote will be preferentially degraded. Recent work by van Melderen and co-workers (1996) has shown that the Lon protease is more effective against free CcdA than bound CcdA. In strains such as JM109 a tightly bound antidote may be degraded by non-specific proteases at the same rate as the toxin while in strains such as CSH50 the antidote - toxin binding equilibrium might be shifted such that the antidote is unbound for sufficient time for it to be selectively degraded. If the antidote is transiently unbound, the toxin will also be transiently unbound. When the *pas* system is present this will cause cellular toxicity which will result in the slower growth rates and the lower stationary phase cell densities seen for CSH50-I^q. When expressed from its native promoter, the *pas* locus will produce lower concentrations of Pas proteins since the *pas* promoter is autoregulated, this will ensure that inappropriate toxicity by PasB will be minimised. The inability of PasC to moderate the effects of PasA and PasB in JM107 is confusing. The initial experiment was carried out in triplicate and then repeated. The results were consistent indicating that this is a real effect. However, since the *pas* system was unable to stabilise pOU-PTF in both this strain and the closely related strain JM109, it is unlikely that the effect of PasC determines the efficiency of stabilisation in these strains. Variation in plasmid stability is probably due to the relative toxicity of the poison in different *E. coli* host strains. The more toxic the poison is to the host strain, the greater the ability of the *pas* system to stabilise the plasmid. Stability due to the *pas* system appears to rely on a number of factors including the strength of the interaction between toxin and antidote. Very strong binding of the antidote to toxin will prevent the toxin functioning at the appropriate time, however, if binding is too

weak, the toxin will kill plasmid containing cells and contribute little to plasmid stability (see Figure 4.2).

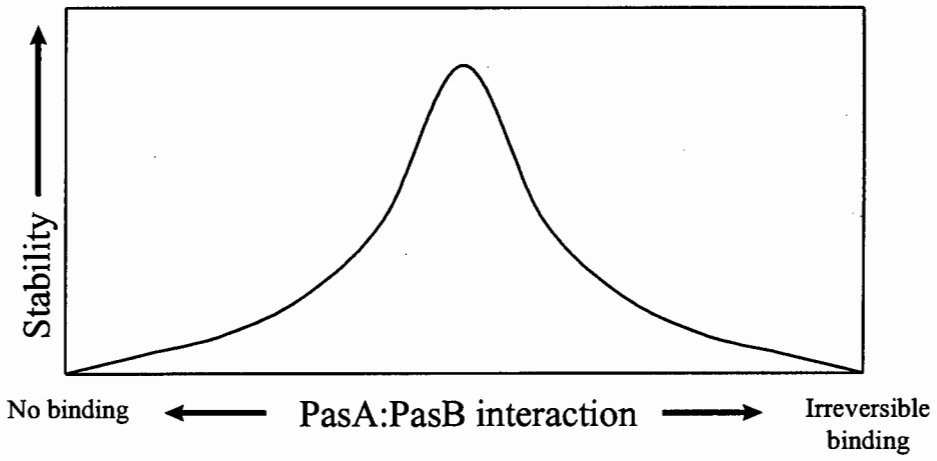


Figure 4.2 The effect of the strength of PasA:PasB interaction on stability.

The model put forward here relies on the antidote bound to the toxin being less or not susceptible to degradation by the Lon protease than free antidote. In order to prove this assumption it would be necessary to construct an *in vitro* assay that was able to monitor the degradation of bound PasA compared to unbound PasA. If purified PasA and PasB were mixed and separated on a non-denaturing PAGE gel, a band corresponding to PasA:PasB complexes should be visible in addition to bands for each of the unbound proteins. If purified Lon protease was added to the mixture of PasA and PasB and samples taken over a range of time intervals the intensity of the band corresponding to unbound PasA should decrease faster than the band corresponding to the PasA:PasB complexes. If, however, there is no difference in the rates of Lon-mediated degradation of bound and unbound PasA, the intensity of the bands for PasA and PasA:PasB should decrease at similar rates.

Chapter 5

General Conclusions

Plasmid pTF-FC2 has been reported to be remarkably stable within *Pseudomonas flourescens* in the absence of selection (Herrera *et al.*, 1992) in spite of the lack of any obvious stability mechanism. The stability mechanism that this study has uncovered is both simple and effective in certain *E. coli* strains. Figure 5.1 below outlines a model of how the *pas* system functions.

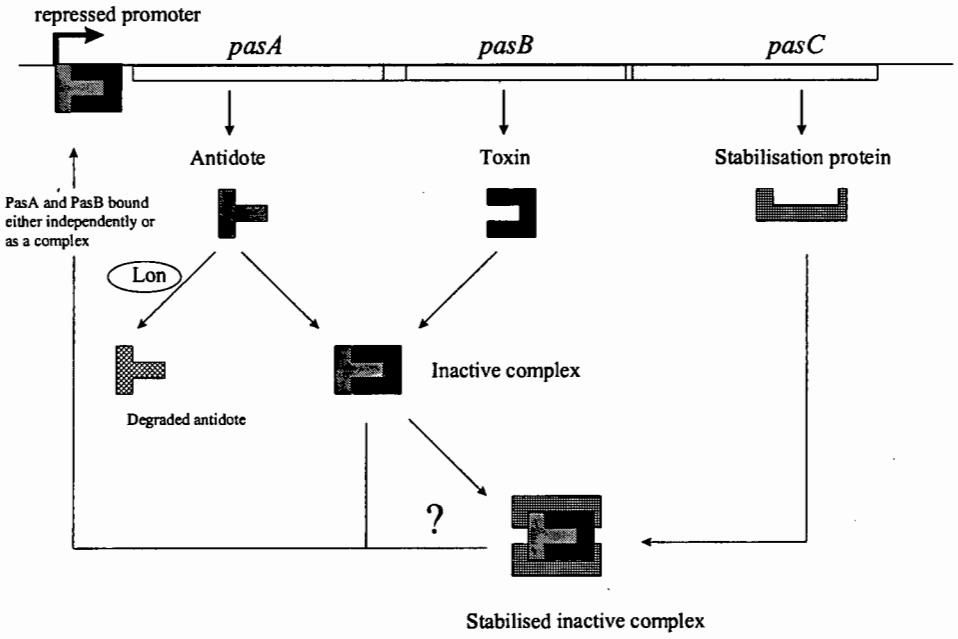


Figure 5.1 Model of the mechanism of *pas*-mediated plasmid stabilisation.

Once introduced into a cell the *pas* operon is strongly expressed from a promoter which must be located within a region of approximately 90 bp upstream of the *pasA* gene. Large amounts of antidote (PasA) and stabilising protein (PasC) are produced, while only a small amount of toxin (PasB) is produced. The proportionally smaller amount of toxin is due to both a poor ribosome binding site and a translational overlap between *pasA* and *pasB*. This excess of antidote ensures that there is no free toxin or that when the toxin becomes dissociated from the antidote, it is rapidly neutralised. The excess of stabilising protein ensures that the toxin

remains bound to the antidote. Once sufficient proteins have been produced the operon will become repressed either by toxin:antibody complexes, toxin:antidote:stabilising protein complexes or by independent binding of toxin and antidote. Future experiments consisting of band shift assays performed with purified Pas proteins should assist in determining the number and combination of proteins which regulate the *pas* promoter. As the cell volume grows, the concentration of toxin and antidote will drop, this in turn will de-repress the *pas* promoter and the concentration of toxin and antidote will be restored to a repressing level. In plasmid-free cells, the antidote will be degraded by the Lon protease and the more stable unbound toxin will be free to cause cell death. Differences in the stability conferred by the *pas* locus in different strains appears at least in the case of CSH50 and JM105 to be due to differences in the association of toxin and antidote. A balance between bound and free toxin appears required for stability to be achieved.

The *pas* system has several of the features commonly found in proteic plasmid stabilisation systems. The antidote gene is located at the start of the operon and is followed by the toxin gene. Both the toxin and antidote genes encode small proteins and the operon is negatively autoregulated. In the *pas* systems, as in all other proteic stabilisation systems, factors which impede the translation of the gene for the toxin are present. These include a poor ribosome binding site and overlap with the terminus of the gene for the antidote, which ensures that the toxin is expressed at far lower levels than the antidote. The poor sequence homology shown by the *pas* system to other proteic systems is almost a unifying feature of proteic stabilisation systems described to date and should not be regarded as a indication that *pas* might not be a proteic system. The presence of a third gene, *pasC*, although unique in proteic stabilisation systems, does not fundamentally alter how the system functions. The presence of a protein which enhances the neutralising effect of the antidote can be regarded as an adaptation to the weak neutralising effect exerted by the pas antidote. Because the extent of the neutralisation of the PasB toxin by the PasA antidote is not completely effective, cells which contain the *pasAB* system grow more slowly than plasmid free cells. Under such conditions there would be a strong selection pressure for plasmids in which either: (a) PasA and PasB binding was more effective or (b) for a protein which would assist in PasA and PasB binding. The *pas* stabilisation system of pTF-FC2 appears to have used the second approach in that PasC enhances the ability of the PasA antidote to neutralise the PasB toxin. Cells containing plasmids with the complete PasA:PasB:PasC complex would out-compete cells containing only

the PasA:PasB complex and this selection pressure may have helped in the evolution of the three component system.

It has been suggested by some researchers (Proceedings of Plasmid Biology 96, Graz, Austria) that plasmids may be considered to be modular in nature. This means that many plasmids have discrete modules which control replication, conjugation or stability. The *pas* operon, although unassociated with replication of pTF-FC2 replicon, exists amongst the replication genes. This layout suggests that the *pas* operon is an evolutionary module which has been inserted into the replicon in such a way that plasmid replication was not noticeably affected. The stabilising ability of the module would ensure that the module is strongly selected for and is perpetuated in the population. The E and F genes of the closely related plasmid RSF1010 do not appear to constitute a similar module since the F protein is responsible for regulating the P_4 promoter which controls of replication proteins RepA and RepC. The E and F proteins can both be produced in large amounts with no ill effect on the cells in which they are produced, thus it is unlikely that either is a toxin of a proteic stabilisation system. The related IncQ-like plasmid pIE1107 has only one open reading frame at the position analogous to that of the E and F genes (pers comm. Erhard Tietze) suggesting that regulation of replication may be different for different IncQ-like plasmids. pTF-FC2 does not carry any readily identifiable antibiotic resistance determinants or other obvious beneficial elements and thus would be expected to be easily lost from a population. It is possible that the IncQ ancestor plasmid did not contain a proteic stabilisation system and that some decedent plasmids have acquired this system. Plasmids such as RSF1010 and pEI1107 which are found in host species which may benefit from antibiotic resistance could have acquired resistance genes to ensure their selection. *T. ferrooxidans* lives in an environment where antibiotics are chemically unstable, thus resistance to antibiotics would not be beneficial. Due to a lack of appropriate plasmid borne selective markers, PTF-FC2 may have acquired the proteic stabilisation system despite the toxic load it causes. It equally probable that the IncQ ancestor plasmid possessed a proteic stabilisation system which has, where possible, been substituted with less toxic selective traits such as antibiotic markers. It must be noted that the *pas* system was studied in *E. coli* and that how it performs in *T. ferrooxidans* is unknown. Unfortunately being a highly acidophilic, obligately chemolithotrophic bacterium, a workable genetic system for *T. ferrooxidans* is not available. As a result the performance of *pas* mutants in the host from which the plasmid was isolated cannot be studied as yet.

Certain questions still remain to be answered about the *pas* system. The target of the poison is as yet unknown as are the exact ratios in which poison, antidote and stabilising protein combine to repress the *pas* promoter or neutralise the poison. Does PasA contain an autoregulatory domain and a separate PasB neutralising domain? It is also not known how the lack of PasC will effect the stability of the system in different *E. coli* and other hosts or over periods of greater than the 100 generations examined in this study. Of great interest is the question of whether autoregulation is required for plasmid stability, can a promoter insensitive to the Pas proteins substitute for the *pas* promoter? Data presented here suggests that the Lon protease may require the antidote to be unbound for it to be degraded and that a fine balance in the strength of toxin:antidote interaction may be required for optimum plasmid stabilisation. These and other questions raised elsewhere in this thesis should form the basis for further studies of this fascinating system.

Literature Cited

Abeles, A.L., Friedman, S.A., and Austin, S.J. (1985) Partition of unit-copy mini-plasmids to daughter cells. III. The DNA sequence and functional organisation of the P1 partition region. *J Mol Biol* **185:** 261-272.

Abeles, A., Brendler, T. and Austin, S. (1993) Evidence of two levels of control of P1 oriR and host oriC replication origins by DNA adenine methylation. *J Bacteriol* **175:** 7801-7807.

Austin, S., Ziese, M., and Sternberg, N. (1981) A novel role for site-specific recombination in the maintenance of bacterial replicons. *Cell* **25:** 729-736.

Austin, S., and Wierzbecki, A. (1983) Two mini-F-encoded proteins are essential for equipartitioning. *Plasmid* **10:** 73-81.

Austin, S., and Abeles, A. (1985) The partition functions of P1, P7, and F mini-plasmids. *Basic Life Science* **30:** 215-226.

Austin, S.J., and Eichorn, B.G. (1992) Random diffusion can account for *tpA*-dependent supression of partition defects in low-copy-number plasmids. *J Bacteriol* **174:** 5190-5195.

Bailone, A., Brandenburger, A., Levine, A., Pierre, M., Dutreix, M., and Devoret, R. (1984) Indirect SOS induction is promoted by ultraviolet light-damaged mini-F and requires the mini-F *lynA* locus. *J Mol Biol* **179:** 367-390.

Beaucage, S.L., Miller, C.A., and Cohen, S.N. (1991) Gyrase-dependent stabilisation of pSC101 plasmid inheritance by transcriptionally active promoters. *EMBO J* **10:** 2583-2588.

Belogurov, A.A., Delver, E.P., Rodezevich, O.V. (1993) Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences. *J Bacteriol* **175:** 4843-4850.

Bernard, P., and Couturier, M. (1991) The 41 C-terminal residues of the mini-F plasmid CcdA protein are sufficient to antagonise the killer activity of the CcdB protein. *Mol Gen Genet* **226:** 297-304.

Bernard, P., and Couturier, M. (1992) Cell killing by the F plasmid CcdB protein involves poisoning of the DNA-topoisomerase II complexes. *J Mol Biol* **226:** 735-745.

Bex, F., Karoui, H., Rokeach, L., Dreze, P., Garcia, L., and Couturier, M. (1983) Mini-F encoded proteins: Identification of a new 10.5 kiloDalton species. *EMBO J* **2:** 1853-1861.

Bickle, T.A., and Kruger, D.H. (1993) Biology of DNA restriction. *Microbiol Rev* **57:** 434-450.

Biek, D.P., and Cohen, S.N. (1989) Involvement of integration host factor (IHF) in maintenance of plasmid pSC101 in *Escherichia coli*: Characterization of pSC101 mutants that replicate in the absence of IHF. *J Bacteriol* **171:** 2056-2065.

Biek, D.P., and Cohen, S.N. (1992) Propagation of pSC101 plasmids defective in binding of integration host factor. *J Bacteriol* **174:** 785-792.

Bork, P., Sander, C., and Valencia, A. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci USA* **89:** 7290-7294.

Bravo, A., de Torrontegui, G., and Díaz, R. (1987) Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. *Mol Gen Genet* **210:** 101-110.

Bravo, A., Ortega, S., de Torrontegui, G., and Díaz, R. (1988) Killing of *Escherichia coli* cells modulated by components of the stability system ParD of plasmid R1. *Mol Gen Genet* **215:** 146-151.

Brosius J., and Holy A. (1984) Regulation of ribosomal RNA promoters with a synthetic *lac* operator. *Proc. Natl. Acad. Sci. USA* **81**: 6929-6933.

Burkardt, H., Reiss, G., and Phüler, A. (1979) Relationships of group P1 plasmids revealed by heteroduplex experiments: RP1, RP4, R68 and RK2 are identical. *J Gen Micro* **114:** 341-348.

Campbell, J.L. and Kleckner, N. (1990) *E. coli* oriC and *dnaA* gene promoter are sequestered from dam methyltransferase follwing the passage of the chromosomal replication fork. *Cell* **62:** 967-979.

Casadaban M.J., Martinez-Arias A., Shapira S.K., and Chou J. (1983) β-galactosidase gene fusions for analysing gene expression in *Escherichia coli* and Yeast. *Methods in Enzymology*. **100**: 293-308.

Church, G., and Gilbert, W. (1984) Church Blots. *Proc Natl Acad Sci* **81:** 1991-995

Clennel, A-M., Johnston, B., and Rawlings, D.E. (1995) Structure and function of Tn5467, a Tn21-like transposon located on the *Thiobacillus ferrooxidans* broad-host-range plasmid pTF-FC2. *Appl Environ Microbiol* **61:** 4223-4229.

Colloms, S.D., Sykora, P., Szatmari, G., and Sherratt, D.J. (1990) Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. *J Bacteriol* **172:** 6973-6980.

Conley, D.L., and Cohen, S.N. (1995a) Effects of the pSC101 partition (*par*) locus on *in vivo* DNA supercoiling near the plasmid replication origin. *Nucleic Acids Res* **23:** 701-707.

Conley, D.L., and Cohen, S.N. (1995b) Isolation and characterisation of plasmid mutations that enable partitioning of pSC101 replicons lacking the partition (*par*) locus. *J Bacteriol* **177:** 1086-1089.

Dam, M., and Gerdes, K. (1994) Partitioning of plasmid R1: Ten direct repeats flanking the *parA* promoter constitute a centromere-like partition site *parC*, that expresses incompatibility. *J Mol Biol* **236:** 1289-1298.

Davis, T.L., Helinski, D.R., and Roberts, R.C. (1992a) Transcription and autoregulation of the stabilising functions of the broad-host-range plasmid RK2 in *Escherichia coli, Agrobacterium tumefaciens and Pseudomonas aeruginosa*. *Mol Micro* **6:** 1981-1994.

Davis, M.A., Martin, K.A., and Austin, S.J. (1992b) Biochemical activities of the ParA partition protein of the P1 plasmid. *Mol Micro* **6:** 1141-1147.

de Boer, P.A.J., Crossley, R.E., and Rothfield, L.I. (1989) A division inhibitor and a topological specificity factor encoded by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* **56:** 641-649.

Dorrington, R.A., and Rawlings, D.E. (1990) Characterization of the minimum replicon of the broad-host-range plasmid pTF-FC2 and similarity between pTF-FC2 and the IncQ plasmids. *J Bacteriol* **172:** 5697-5705

Dorrington, R.A., Bardien, S., and Rawlings, D.E. (1991) The broad-host-range plasmid pTF-FC2 requires a primase-like protein for autonomous replication in *Escherichia coli*. *Gene* **108:** 7-14.

Eberl, L., Givskov, M., and Schwab, H. (1992) The divergent promoters mediating transcription of the *par* locus of the plasmid RP4 are subject to autoregulation. *Mol Micro* **6:** 1969-1979.

Eberl, L., Kristensen, C.S., Givskov, M., Grohmann, E., Gerlitz, M., and Schwab, H. (1994) Analysis of the multimer resolution system encoded by the *parCBA* operon of the broad-host-range plasmid RP4. *Mol Micro* **12:** 131-141.

Ezaki, B., Ogura, T., Niki, H., and Hiraga, S. (1991) Partitioning of a mini-F plasmid into anucleate cells of the *mukB* null mutant. *J Bacteriol* **173:** 6643-6646.

Figurski, D.H., Pohlman, R.F., Bechhofer, D.H., Prince, A.S., and Kelton, C.A., (1982) Broad-host-range plasmid RK2 encodes multiple *kil* genes potentially lethal to *Escherichia coli* host cells. *Proc Nat Acad Sci USA* **79:** 1935-1939.

Franklin, F.C.H., Bagdasarian, M., Bagdasarian, M.M., and Timmis, K.N. (1981) Molecular and functional analysis of the TOL plasmid pWW0 from *Pseudomonas putida* and cloning of the genes for the entire regulated aromatic ring *meta*-cleavage pathway. *Proc Natl Acad Sci USA* **78:** 7458-7462.

Funnell, B.E. (1988a) Participation of *Escherichia coli* integration host factor in the P1 plasmid partition system. *Proc Natl Acad Sci USA* **85:** 6657-6661.

Funnell, B.E. (1988b) Mini-P1 plasmid partotioning: Excess ParB protein destabilises plasmids containing the centromere *parS*. *J Bacteriol* **170:** 954-958.

Funnell, B.E. (1991) The P1 partition complex at *parS*: The influence of *Escherichia coli* integration host factor and substrate topology. *J Biol Chem* **266:** 14328-14337.

Funnell, B.E., and Gagnier, L. (1993) The P1 plasmid partition complex at parS. *J Biol Chem* **268:** 3616-3624.

Funnell, B.E., and Gagnier, L. (1994) P1 plasmid partition: Binding of P1 ParB protein and *Escherichia coli* integration host factor to altered *parS* sites. *Biochimie* **76:** 924-932.

Gerdes, K., Larsen, J.E.L., and Molin, S. (1985) Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol* **161:** 292-298.

Gerdes, K., Rasmussen, P.B., and Molin, S. (1986a) Unique type of plasmid maintenance: post-segregational killing of plasmid free cells. *Proc Natl Acad Sci USA* **83:** 3116-3120.

Gerdes, K., Bech, F.W., Jørgensen, S.T., Lobner-Olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlström, O., Molin, S., and von Meyenburg, K. (1986b) Mechanism of post-segregational killing by the *hok* gene product of the *parB* system of plasmid R1 and its homology with the *relF* gene product of the *E. coli relB* operon. *EMBO J* **5:** 2023-2029.

Gerdes, K., and Molin, S. (1986c) Partitioning of plasmid R1: Structural and functional analysis of the *parA* locus. *J Mol Biol* **190:** 269-279.

Gerdes, K., Poulsen, L.K., Thisted, T., Nielsen, A.K., Martinussen, J., and Andreasen, P.H. (1990a) The *hok* killer gene family in gram-negative bacteria. *New Biologist* **2:** 946-956.

Gerdes, K., Thisted, T., and Martinussen, J. (1990b) Mechanism of post-segregational killing by the *hok/sok* system of the plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Mol Micro* **4:** 1807-1818.

Gerdes, K., Nielsen, A., Thorsted, P., and Wagner, E.G.H. (1992) Mechanism of killer gene activation. Antisense RNA-dependent RNAse III cleavage ensures rapid turn-over of the stable Hok, SrnB and PndA effector messenger RNAs. *J Mol Biol* **226:** 637-649.

Gerlitz, M., Hrabak, O., and Schwab, H. (1990) Partitioning of broad-host-range plasmid RP4 is based on a complex function including a site-specific recombination system. *J Bacteriol* **172:** 6194-6203

Goncharoff, P., Saadi, S., Chang, C-H., Saltman, L.H., and Figurski, D.H. (1991) Structural, molecular, and genetic analysis of the *kilA* operon of broad-host-range plasmid RK2. *J Bacteriol* **173:** 3463-3477.

Grewal, K.K. (1990) Plasmid RP4 host-lethal function *kilA* and its repressor *korA* sequences are part of the cryptic tellurite resistance transposon Tn521. *FEMS Micro Lett* **69:** 97-100.

Gustafsson, P., Eolf-Watz, H., Lind, L., Johansson, K-E., and Nordström, K. (1983) Binding between the *par* region of plasmids R1 and pSC101 and the outer membrane fraction of the host bacteria. *EMBO J* **2:** 27-32.

Hanai, R., Lui, R., Benedetti, P., Caron, P.R., Lynch, A.S., and Wang, J.C. (1996) Molecular dissection of a protein SopB essential for *Escherichia coli* F plasmid partition. *J Biol Chem* **271:** 17469-17475.

Helsberg, M., and Eichenlaub, R. (1986) Twelve 43-base-pair repeats map in a *cis*-acting region essential for partition of plasmid mini-F. *J Bacteriol* **165:** 1043-1045.

Herrera, G., Sntman, S.J., and Thomson, J.A. (1992) Construction of a bioinsecticidal strain of *Pseudomonas flourescens* active against the sugar can borer, *Eldana saccharina. Appl Environ Microbiol* **60:** 682-690.

Hiraga, S. (1992) Chromosome and plasmid partition in *Escherichia coli. Annu Rev Biochem* **61:** 283-306.

Hiraga, S., Ichinose, C., Kido, M., Hirano, M., and Yamazoe, M. (1996) Abstracts of Plasmid Biology 96: International symposium on plasmid molecular biology. p67.

Ingmer, H. and Cohen, S.N. (1993) The pSC101 *par* locus alters protein-DNA interactions *in vivo* at the plasmid replication origin. *J Bacteriol* **175:** 6046-6048.

Jaffé, A., Ogura, T., and Hiraga, S. (1985) Effects of the *ccd* function of the F plasmid on bacterial growth. *J Bacteriol* **163:** 841-849.

Jagura-Burdzy, G., Ibbotson, J.P., Thomas, C.M. (1991) The *korF* region of broad-host-range plasmid RK2 encodes two polypeptide products with transcriptional repressor activity. *J Bacteriol* **173:** 826-833.

Jagura-Burdzy, G., and Thomas, C.M. (1992) *krfA* gene of broad-host-range plasmid RK2 encodes a novel DNA-binding protein. *J Mol Biol* **225:** 651-660.

Jagura-Burdzy, G., Ibbotson, J.P., and Thomas, C.M. (1992) Crosstalk between plasmid vegetative replication and conjugative transfer: repression of the *trfA* operon by *trbA* of broad-host-range plasmid RK2. *Nucleic Acids Res* **20:** 3939-3944.

Jakes, K.S. (1982) The mechanism of action of colicin E2, colicin E3 and cloacin DF13. In *Molecular action of toxins and viruses*. Cohen, P., and van Heyningen, S. (eds). Elsevier Biomedical Press, Amsterdam. 131-167.

Jakes, K.S., and Model, P. (1979) Mechanism of export of colicin E1 and colicin E3. *J Bacteriol* **138:** 770-778.

Jensen, R.B., and Gerdes, K. (1995) Programmed cell death in bacteria: proteic plasmid stabilisation systems. *Mol Micro* **17:** 205-210.

Jensen, R.B., Grohmann, E., Schwab, H., Díaz-Orejas, R., and Gerdes, K. (1995) Comparison of *ccd* of F, *parDE* of RP4, and *parD* of R1 using a novel conditional replication control system of plasmid R1. *Mol Microbiol* **17:** 211-220. passim

Johnson, E.P., Strom, A.R., and Helinski, D.R. (1996) Plasmid RK2 toxin protein ParE: Purification and interaction with the ParD antitoxin protein. *J Bacteriol* **178:** 1420-1429.

Kobayshi, I., Nakayama, Y., Naito, T., Handa, N., and Kusano, K. (1996) Abstracts of Plasmid Biology 96: International symposium on plasmid molecular biology.

Kornacki, J.A., Chang, C-H., and Figurski, D.H. (1993) *kil-kor* regulon of promiscuous plasmid RK2: Structure, products, and regulation of two operons that constitute the *kilE* locus. *J Bacteriol* **175:** 5078-5090.

Kulakauskas, S., Lubys, A., and Ehrlich, S.D. (1995) DNA restriction-modification systems mediate plasmid maintenance. *J Bacteriol* **177:** 3451-3454.

Kusukawa, N., Mori, H., Kondo, A., and Hiraga, S. (1987) Partitioning of the F plasmid: overproduction of an essential protein for plasmid partition inhibits plasmid maintenance. *Mol Gen Genet* **208:** 365-372.

Lane, D., de Feyter, R., Kennedy, M., Phua, S-H., Simon, D. (1986) D protein of mini-F plasmid acts as a repressor of transcription and as a site-specific resolvase. *Nucleic Acids Res* **14:** 9713-9728.

Lane, D., Rothenbuehler, R., Merrillat, A-M., and Aiken, C. (1987) Analysis of the F plasmid centromere. *Mol Gen Genet* **207:** 406-412.

Larsen, J.E.L., Gerdes, K., Light, J., and Molin, S. (1984) Low-copy-number plasmid-cloning vectors amplifiable by depression of an inserted foreign promoter. *Gene* **28:** 45-54.

Lehnherr, H., and Yarmolinsky, M.B. (1995) Addiction protein Phd of plasmid prophage P1 is a substrate of the ClpXP serine protease of *Escherichia coli. Proc Natl Acad Sci USA* **92:** 3274-3277.

Lehnherr, H., Maguin, E., Jafri, S., and Yarmolinsky, M.B. (1993) Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of the prophage, and *phd*, which prevents host death when prophage is retained. *J Mol Biol* **233:** 414-428.

Lin, Z. and Mallavia, L.P. (1994) Identification of a partition region carried by the plasmid QpH1 of *Coxiella burnetii. Mol Micro* **13:** 513-523.

Loh, S.M., Ray, A., Cram, D.S., O'Gorman, L.E., and Skurray, R.A. (1986) Location of a second partition region (ParL) on the F plasmid. *FEMS Micro Lett* **37:** 179-182. 29

Loh, S.M., Cram, D.S., and Skurray, R.A. (1988) Nucleotide sequence and transcriptional analysis of a third function (*Flm*) involved in F-plasmid maintenance. *Gene* **66:** 259-268.

Luria, S.E., and Suit, J.L. (1982) Transmembrane channels produced by colicin molecules. In *Membranes and transport.* Martonosi, A. (ed). Plenum Publishing, New York. **2:** 279-284.

Lynch, S.A., and Wang J.C. (1995) SopB protein-mediated silencing of genes linked to the *sopC* locus of *Escherichia coli* F plasmid. *proc Natl Acad Sci USA* **93:** 1896-1900.

Maeser, S., Scholtz, P., Otto, S., and Scherzinger, E. (1990) Gene F of the plasmid RSF1010 codes for a low-molecular-weight repressor protein that autoregulates expression of the *repAC* operon. *Nucleic Acids Res.* **18:** 6215-6222.

Maki, S., Takiguchi, S., Miki, T., and Horiuchi, T. (1992) Modulation of DNA supercoiling activity of *Escherichia coli* DNA gyrase by F plasmid proteins. *J Biol Chem* **267:** 12244-12251.

Maki, S., Takiguchi, S., Horiuchi, T., Sekimizu, K., and Miki, T. (1996) Partner switching mechanisms in the inactivation and rejuvenation of *Escherichia coli* DNA gyrase by F plasmid proteins LetD (CcdB) and LetA (CcdA). *J Mol Biol* **256:** 473-482.

Malki, A., Kern, R., Kohiyama, M., and Hughes, P. (1992) Inhibition of DNA synthesis at the hemimethylated pBR322 origin of replication by a cell membrane fraction. *Nucleic Acids Res* **20:** 105-109.

Manen, D., Goebel, T., and Caro, L. (1990) The *par* region of pSC101 affects plasmid copy number as well as stability. *Mol Microbiol* **4:** 1839-1846.

Martin, K.A., Friedman, S.A., and Austin, S.J. (1987) Partition site of the P1 plasmid. *Proc Natl Acad Sci USA* **84:** 8544-8547.

Masuda, Y., Miyakawa, K., Nishimura, Y., and Ohtsubo, E. (1993) *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. *J Bacteriol* **175:** 6850-6856.

Masuda, Y., and Ohtsubo, E. (1994) Mapping and disruption of the *chpB* locus in *Escherichia coli*. *J Bacteriol* **176:** 5861-5863.

Meacock, P.A., and Cohen, S.N. (1980) Partitioning of bacterial plasmids during cell division. A *cis*-acting locus that accomplishes stable plasmid inheritance in cell populations. *Cell* **20:** 529-542.

Miki, T., Yoshioka, K., and Horiuchi, T. (1984a) Control of cell division by sex factor F in *Escherichia coli*: I. The 42.84-43.6 F segment couples cell division of the host bacteria with replication of plasmid DNA. *J Mol Biol.* **174:** 605-625.

Miki, T., Chang, Z-T., Horiuchi, T. (1984b) Control of cell division by sex factor F in *Escherichia coli*: II. Identification of the genes for inhibition protein and trigger protein on the 42.84-43.6 F segment. *J Mol Biol* **174:** 627-646.

Miki, T., Orita, T., Furuno, M., and Horiuchi, T. (1988) Control of cell division by sex factor F in *Escherichia coli*. III. Participation of the *groES* (*mopB*) gene of the host bacteria. *J Mol Biol* **201:** 327-338.

Miki, T., Park, J.A., Nagao, K., Murayama, N., and Horiuchi, T. (1992) Control of segregation of chromosomal DNA by sex factor F in *Escherichia coli*: Mutants of DNA gyrase A suppress *letD* (*ccdB*) product growth inhibition. *J Mol Biol* **225:** 39-52.

Miller, C.A., Tucker, W.T., Meacock, P.A., Gustafsson, P., and Cohen, S.N. (1983) Nucleotide sequence of the partition locus of *Escherichia coli* plasmid pSC101. *Gene* **24:** 309-315.

Miller, C.A., Beaucage, S.L., and Cohen, S.N. (1990) Role of DNA superhelicity in partioning of the pSC101 plasmid. *Cell* **62:** 127-133.

Miller, C.A., and Cohen, S.N. (1993) The partition (*par*) locus of pSC101 is an enhancer of plasmid incompatability. *Mol Microbiol* **9:** 695-702.

Mori, H., Kondo, A., Ohshima, A., Ogura, T., and Hiraga, S. (1986) Structure and function of the F plasmid genes essential for partitioning. *J Mol Biol* **192:** 1-15.

Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A., Hiraga, S. (1989) Purification and characterisation of SopA and SopB proteins essential for F plasmid partitioning. *J Biol Chem* **264:** 15535-15541.

Motallebi-Veshareh, M., Rouch, D.A., and Thomas, C.M. (1990) A family of ATPases involved in active partitioning of diverse bacterial plasmids. *Mol Micro* **4:** 1455-1463.

Murayama, N., Shimizu, H., Takiguchi, S., Yoshifumi, B., Amino, H., Horiuchi, T., Sekimizu, K., and Miki, T. (1996) Evidence for involvement of *Escherichia coli* genes *pmbA*, *csrA* and a previously unrecognised gene *tldD*, in the control of DNA gyrase by *letD* (*ccdB*) of sex factor F. *J Mol Biol* **256:** 483-502.

Naito, T., Kusano, K., and Kabayashi, I. (1995) Selfish behavior of restriction-modification systems. *Science* **267:** 897-899.

Niki, H., Jaffe, A., Imamura, R., Ogura, T., and Hiraga, S. (1991) The new gene *mukB* codes for a 177 kDa protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *EMBO J* **10:** 183-193.

Nordström, K., and Austin, S.J. (1989) Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet* **23:** 37-69.

Nordström, K., and Austin, S. (1990) Partition-mediated incompatibility of bacterial plasmids. *Cell* **60:** 351-354.

Nordström, K., Molin, S., and Aagard-Hansen, H. (1981) Plasmid R1 incompatability. Contribution from the *cop/rep* and from the *par* systems. In *Molecular Biology, Pathogenicity, and Ecology of Bacterial Plasmids*. Clowes, R.C., Levy, S.B., and Koenig, E.L., eds. (New York: Plenum Press), 291-301.

Ogura, T., and Hiraga, S. (1983a) Gene-encoded products and a cis-acting region are involved in partition. *Cell* **32:** 351-360.

Ogura, T., and Hiraga, S. (1983b) Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc Natl Acad Sci USA* **80:** 4784-4788.

Ohnishi, Y., and Schlessinger, D. (1972) Total break-down of ribosomal and transfer RNA in a mutant of *Escherichia coli*. *Nature New Biol* **238:** 228-231.

Ohnishi, Y., and Akimoto, S. (1980) I-like R-plasmids promote degradation of stable ribonucleic acid in *Escherichia coli*. *J Bacteriol* **144:** 833-835.

Ono, K., Akimoto, S., and Ohnishi, Y. (1987) Nucleotide sequence of the *pnd* gene product in plasmolysis. *Microbiol Immunol* **31:** 1071-1083.

Pansegrau, W., Lanka, E., Barth, P.T., Figurski, D.H., Guiney, D.G., Haas, D., Helinski, D.R., Schwab, H., Stanisich, V.A., and Thomas, C.M. (1994) Complete nucleotide sequence of the Birmingham IncPα plasmids: Compilation and comparative analysis. *J Mol Biol* **239:** 623-663.

Patient, M.E., and Summers, D.K. (1993) ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. *Mol Microbiol* **9:** 1089-1095

Pecota, D.C., and Wood, T.K. (1996) Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *J Bact* **178:** 2044-2050.

Pecota, D.C., Kim, C.S., Wu, K., Gerdes, K., and Wood, T.K. (1997) Combining the *hok/sok*, *par*DE and *pnd* postsegregational killer loci to enhance plasmid stability. *Appl. and Env. Micro.* **63:** 1917-1924.

Poulsen, L.K., Larsen, N.W., Molin, S., and Andersson, P. (1989) A family of genes encoding a cell killing function may be conserved in all gram negative bacteria. *Mol Micro* **3:** 1463-1472.

Poulsen, L.K., Refn, A., Molin, S., and Andersson, P. (1991) The *gef* genes from *Escherichia coli* is regulated at the level of translation. *Mol Micro* **5:** 1639-1648.

Prentki, P., Chandler, M., and Caro, L. (1977) Replication of the prophage P1 during the cell-cycle of *Escherichia coli*. *Mol Gen Genet* **152:** 71-76.

Rabussay, D. (1983) Regulation of gene expression transcription: phage evoked changes in RNA polymerase. In Mathews, C., Kutter, E.M., Mosig, G., and Berget, P.B. (ed.), *Bacteriophage T4*. American Soceity for Microbiology, Washington, D.C. 167-173.

Rawlings, D.E., Gawith, C., Petersen, A., and Woods, D.R. (1983) Characterisation of plasmids and potential genetic markers in *Thiobacillus ferrooxidans*. In Rossi, G.,

and Torma, A.E. (eds) *Progress in Biohydrometallurgy,* Associazione Mineraria Sarda, Cagliari. 555-570.

Rawlings D.E., Pretorius I.-M., and Woods D.R. (1986a) Expression of *Thiobacillus ferrooxidans* plasmid functions and the development of genetic systems for Thiobacilli. In Biotechnology and Bioengineering Symp. No. 29. John Wiley & Sons Inc. pp. 291-297.

Rawlings D.E., Sewcheran R., and Woods D.E. (1986b) Characterization of a broad-host-range mobilizable *Thiobacillus ferrooxidans* plasmid and the construction of *Thiobacillus* cloning vectors. In *Proc. of the International Symp. on Biohydrometallurgy.* Elsevier. pp. 419-427

Rawlings, D.E., and Silver, S. (1995) Mining with microbes. *Bio/Technology* **13:** 773-778.

Read, T.D., Thomas, A.T., and Wilkins, B.M. (1992) Evasion of type I and type II DNA restriction systems by IncI1 plasmid ColIb-P9 during transfer by bacterial conjugation. *Mol Micro* **6:** 1933-1941.

Roberts, R.C., Burioni R., and Helinski, D.R. (1990) Genetic characterisation of the stabilising functions of a region of the broad-host-range plasmid RK2. *J Bacteriol* **172:** 6204-6216.

Roberts, R.C., and Helinski, D.R. (1992) Definition of a minimal plasmid stabilisation system from the broad-host-range plasmid RK2. *J Bacteriol* **174:** 8119-8132.

Roberts, R.C., Spangler, C., and Helinski, D.R. (1993) Characteristics and significance of the DNA binding activity of plasmid stabilisation protein ParD from the broad-host-range plasmid RK2. *J Biol Chem* **268:** 27109-27117.

Roberts, R.C., Ström, A.R., and Helinski, D.R. (1994) The *parDE* operon of the broad-host-range plasmid RK2 specifies growth inhibition associated with plasmid loss. *J Mol Biol* **237:** 35-51.

Rohrer J., and Rawlings D.E. (1994) Regulation of the mobilization of the broad-host-range plasmid pTFF-FC2. Mol. Microbiol. **9:** 1051-1059.

Rosenfeld, R., and Grover, N.B. (1993) Partition of plasmid R1: A computer simulation. *Plasmid* **29:** 117-124.

Ruiz-Echevarría, M.J., Berzal-Herranz, A., Gerdes, K., and Díaz-Orejas, R. (1991a) The *kis* and *kid* genes of the *parD* maintenance system of plasmid R1 form an operon that is autoregulated at the level of transcription by the co-ordinated action of the Kis and Kid proteins. *Mol Micro* **5:** 2685-2693.

Ruiz-Echevarría, M.J., de Torrontegui, G., Giménez-Gallego, G., and Díaz-Orejas, R. (1991b) Structural and functional comparison between the stability systems ParD of plasmid R1 and Ccd of plasmid F. *Mol Gen Genet* **225:** 355-362.

Ruiz-Echevarría, M.J., Giménez-Gallego, G., Sabariegos-Jareño, R., and Díaz-Orejas, R. (1995a) Kid, a small protein of the *parD* stability system of plasmid R1, is an inhibitor of DNA replication acting at the initiation of DNA synthesis. *J Mol Biol* **247:** 568-577.

Ruiz-Echevarría, M.J., de la Torre, M.A., and Díaz-Orejas, R. (1995b) A mutation that decreases the efficiency of plasmid R1 replication leads to the activation of *parD*, a killer stability system of the plasmid. *FEMS Micro Lett* **130:** 129-136.

Russell, D.W. and Zinder, N.D. (1987) Hemimethylation prevents DNA replication in *E. coli*. *Cell* **50:** 1071-1079.

Salmon, M.A., van Melderen, L., Bernard, P., Couturier, M. (1994) The antidote and autoregulatory functions of the F plasmid CcdA protein: a genetic and biochemical survey. *Mol Gen Genet* **244:** 530-538.

Saltman, L.H., Kim, K-S., and Figurski, D.H. (1991) The *kilA* operon of promiscuous plasmid RK2:the use of a transducing phage (λp*klaA*-1) to determine the effects of the lethal *klaA* gene on *Escherichia coli* cells. *Mol Micro* **11:** 2673-2683.

Sambrook, J., Maniatis, T., and Fritsch, F. (1989) Molecular cloning: a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M., Scherzinger, E., (1989) Complete nucleotide sequence and gene organisation of the broad-host-range plasmid RSF1010. *Gene* **75:** 271-288.

Sharp, P.A., Cohen, S.N., and Davidson, N. (1973) Electron microscope heteroduplex studies of sequence relationships among plasmids of *Escherichia coli*. *J Mol Biol* **75:** 235-255.

Shortle, D. (1983) A genetic system for analysis of staphylococcal nuclease. *Gene* **22:** 181-189.22

Sobecky, P.A., Easter, C.L., Bear, P.D., and Helinski D.R. (1996) Characterisation of the stable maintenance properties of the *par* region of broad-host-range plasmid RK2. *J Bacteriol* **178:** 2086-2093.

Stanssens, P., Opsomer, C., McKeown, Y.M., Kramer, W., Zabeau, M., and Fritz, H-J. (1989) Efficient oligonucleotide-directed construction of mutations in expression vectors

by the gapped duplex DNA method using alternating selectable markers. *Nucleic Acids Res* **17:** 4441-4453.

Sternberg, N., and Hamilton, P. (1981) Bacteriophage P1 site-specific recombination. I. Recombination between *lox*P sites. *J Mol Biol* **150:** 467-786.

Stirling, C.J., Szatmari, G., Stewart, G., Smith, M.C.M., and Sherratt, D.J. (1988) The arginine repressor is essential for plasmid-stabilising site-specific recombination at the ColE1 *cer* locus. *EMBO J* **7:** 4389-4395.

Stirling, C.J., Colloms, S.D., Collins, J.F., Szatmari, G., and Sherratt., D.J. (1989) *XerB*, an *Escherichia coli* gene required for plasmid ColE1 site-specific recombination, is identical to *pepA*, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase. *EMBO J* **8:** 1623-1627.

Stokes, H.W., Moore, J.J., and Krishnapillai, V. (1981) Complementation analysis in *Pseudomonas aeruginosa* of the transfer genes of the wide host range R plasmid R18. *Plasmid* **5:** 202-212.

Summers, D.K., and Sherratt, D.J. 1984 Multimerization of high copy number plasmids cause instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell* **36:** 1097-1103.

Summers, D.K. (1989) Derivatives of ColE1 *cer* show altered topological specificity in site-specific recombination. *EMBO J* **8:** 309-315.

Tam, J.E., and Kline, B.C. (1989a) Control of the *ccd* operon in plasmid F. *J Bacteriol* **171:** 2353-2360.

Tam, J.E., and Kline, B.C. (1989b) The F plasmid *ccd* autorepressor is a complex of CcdA and CcdB proteins. *Mol Gen Genet* **219:** 26-32.

Thisted, T., Nielsen A.K., and Gerdes, K. (1994a) Mechanism of post-segregational killing: Translation of Hok, SrnB and Pnd mRNAs of plasmids R1, F, and R483 is activated by 3`-end processing. *EMBO J* **13:** 1950-1959.

Thisted, T., Sorensen, N.S., Wagner, E.G.H., and Gerdes, K. (1994b) Mechanism of post-segregational killing: Sok antisense RNA interacts with Hok mRNA *via* its 5`-end single-stranded leader and competes with the 3`-end of Hok mRNA for binding to the *mok* translational initiation region. *EMBO J* **13:** 1960-1968.

Thisted, T., Sorensen, N.S., and Gerdes, K. (1995) Mechanism of post-segregational killing: Secondary structure analysis of the entire hok mRNA from plasmid R1 suggests a

fold-back structure that prevents translation and antisense binding. *J Mol Biol* **247:** 859-873.

Thomas, C.M., Theophilus, B.D.M., Johnston, L.J., Jagura-Burdzy, G., Schilf, W., Lurz, R., and Lanka, E. (1990) Identification of a seventh operon on plasmid RK2 regulated by the *korA* gene product. *Gene* **89:** 29-35.

Thomson, V.J., Jovanovic, O.S., Pohlman, R.F., Chang, C-H., and Figurski, D.H. (1993) Structure, function, and regulation of the *kilB* locus of the promiscuous plasmid RK2. *J Bacteriol* **175:** 2423-2435.

Tian, Q.B., Ohnishi, M., Tabuchi, A., and Terawaki, Y. (1996) A new plasmid-encoded proteic killer gene system: Cloning, sequencing, and analysing *hig* locus of plasmid Rts1. *Biochemical and Biophysical Res. Comm.* **220:** 280-284.

Tsuchimoto, S., Ohtsubo, H., and Ohtsubo, E. (1988) Two genes, *pemK* and *pemI*, responsible for stable maintenance of resistance plasmid R100. *J Bacteriol* **170:** 1461-1466.

Tsuchimoto, S., and Ohtsubo, E. (1989) Effect of the *pem* system on stable maintenance of plasmid R100 in various *Escherichia coli* hosts. *Mol Gen Genet* **215:** 463-468.

Tsuchimoto, S., Nishimura, Y., and Ohtsubo, E. (1992) The stable maintenance system *pem* of plasmid R100: Degradation of PemI protein may allow PemK protein to inhibit cell growth. *J Bacteriol* **174:** 4205-4211.

Tsuchimoto, S., and Ohtsubo, E. (1993) Autoregulation by co-operative binding of the PemI and PemK protein to the promoter region of the *pem* operon. *Mol Gen Genet* **237:** 81-88.

Tucker, W.T., Miller, C.A., and Cohen, S.N. (1984) Structural and functional analysis of the *par* region of the pSC101 plasmid. *Cell* **38:** 191-201.

van Dyk, T.K., Gatenby, A.A., and LaRossa, R.A. (1989) Demonstration by genetic suppression of the interaction of *groE* products with many proteins. *Nature* **342:** 451-453.

van Melderen, L., Bernard, P., and Couturier, M. (1994) Lon-Dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol Microbiol* **11:** 1151-1157.

van Melderen, L., Thi, M.H.D., Lecchi, P., Gottesman, S., Couturier, M., and Maurizi, M.R. (1996) ATP-dependent degradation of CcdA by Lon Protease *J Biol Chem* **271:** 27730-27738.

Wahle, E., and Kornberg, A. (1988) The partition locus of plasmid pSC101 is a specific binding site for DNA gyrase. *EMBO J* **7:** 1889-1895.

Walter, E.G., Thomas, C.M., Ibbotson, J.P., and Taylor, D.E. (1991) Transcriptional analysis, translational analysis, and sequence of the *kilA*-tellurite resistance region of plasmid RK2Te^R. *J Bacteriol* **173:** 1111-1119.

Watanabe, E., Inamoto, S., Lee, M-H., Kim, S.U., Ogura, T., Mori, H., Hitaga, S., Yamasaki, M., and Nagai, K. (1989) Purification and characterisation of the *sopB* gene product which is responsible for stable maintenance of mini-F plasmid. *Mol Gen Genet* **218:** 431-435.

Watanabe, E., Wachi, M., Yamasaki, M., and Nagai, K. (1992) ATPase activity of SopA, a protein essential for active partitioning of the F plasmid. *Mol Gen Genet* **234:** 346-352.

Williams, R.D., and Thomas, C.M. (1992) Active partitioning of bacterial plasmids. *J Gen Micro* **138:** 1-16.

Willmott, C.J.R., Critchlow, S.E., Eperon, I.C., and Maxwell, A. (1994) The complex of DNA gyrase and quinolone drugs with DNA forms a barrier to transcription by RNA polymerase. *J Mol Biol* **242:** 351-363.

Wilson, J.W., Sia, E.A., and Figurski, D.H. (1996) Abstracts of Plasmid Biology 96: International symposium on plasmid molecular biology. p72

Yang, Y., Ames, G.F-L. (1988) DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. *Proc Natl Acad Sci USA* **85:** 8850-8854.

Young, C., Prince, A.S., and Figurski, D.H. (1985) *korA* function of the promiscuous plasmid RK2: an autorepressor that inhibits expression of the host-lethal gene *kilA* and replication gene *trfA*. *Proc Natl Acad Sci USA*. **82:** 7374-7378.

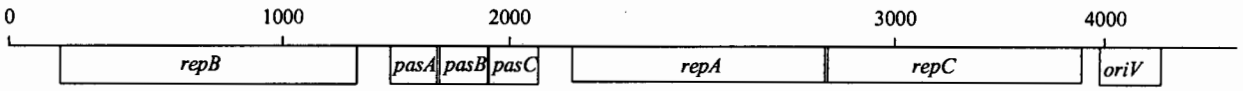
Appendix 1

Bacterial strains used in this study

Strain or plasmid	Source or reference
Strains	
JM109	Yanisch-Perron *et al.* (1985)
JM107	Yanisch-Perron *et al.* (1985)
JM105	Yanisch-Perron *et al.* (1985)
P. putida	Franklin *et al.*, 1981
E. coli GW125a	Dorrington and Rawlings, (1989)
SG22025	S. Gottesman
SG22093	S. Gottesman
SG22095	S. Gottesman
Plasmids	
pMC1403	Casadaban *et al.*, (1983)
pACYC184	Chang and Cohen, (1978)
pUC19	Yanisch-Perron *et al.* (1985)
pMc	Stanssens *et al.*, (1989)
pMa	Stanssens *et al.*, (1989)
pKK223-3	Brosius and Holy, (1984)
pSUP106	Puhler *et al.*, (1985)
pOU82	Gerdes *et al.*, (1985)
pKG339	Jensen *et al.*, (1995)

Appendix 2

Plasmid constructs containing portions of pTF-FC2 generated or used by this study



Plasmid	Position relative to *Cla*I-*Pst*I fragment of pTV100 (Dorrington and Rawlings, 1991) (see Appendix 4)	Vector	Antibiotic resistance marker	Reference
pTF200	1-4910	-	Cm	Dorrington R.A., PhD Thesis
pTV101	1-4910	pUC19	Ap Cm	Dorrington and Rawlings, 1989
pTV151	1-1296	Bluescript SK	Ap	Dorrington R.A., PhD Thesis
pTV400	1240-4441	pUC19	Ap	Dorrington and Rawlings, 1989
pTV4101	1854-4441	pUC19	Ap	Dorrington and Rawlings, 1989
pTV4293	1240-1559	pUC19	Ap	Dorrington R.A., PhD Thesis

pTV151	1-1296	Bluescript SK	Ap	Dorrington R.A., PhD Thesis
pP1H	1-257	pMC1403	Ap	This work
pP2H	1240-1362	pMC1403	Ap	This work
pP3H	1854-2162	pMC1403	Ap	This work
pTac-pasA	1316-1559	pKK223-3	Ap	This work
pTac-pasB	1518-1816	pKK223-3	Ap	This work
pTac-pasC	1789-2028	pKK223-3	Ap	This work
pTac-pasAB	1316-1816	pKK223-3	Ap	This work
pTac-pasBC	1518-2028	pKK223-3	Ap	This work
pTac-pasABC	1316-2028	pKK223-3	Ap	This work
pTac-pasABC*	1316-2028 T insertion at 1833	pKK223-3	Ap	This work
pTac-pasA-pACYC	1316-1559	pACYC184	Cm	This work
pTac-pasB-pACYC	1518-1816	pACYC184	Cm	This work
pTac-pasC-pACYC	1789-2028	pACYC184	Cm	This work
pTac-pasAB-pACYC	1316-1816	pACYC184	Cm	This work
pTac-pasABC-pACYC	1316-2028	pACYC184	Cm	This work
pMRepM0	1-4910	pMc	Cm	This work
pMRepM1	1-4911 T insertion at 1402	pMc	Cm	This work
pMRepM2	1-4911 G insertion at 1563	pMc	Cm	This work
pMRepM3	1-4911 T insertion at 1833	pMc	Cm	This work
pMRepM4	1-4911 A insertion at 2074	pMc	Cm	This work
pKmM0	1-4910	-	Km	This work
pKmM1	1-4911 T insertion at 1402	-	Km	This work

pKmM2	1-4911 G insertion at 1563	-	Km	This work
pKmM3	1-4911 T insertion at 1833	-	Km	This work
pKmM1del1	Δ1232-1943	-	Km	This work
pKmM1del2	Δ1217-1321	-	Km	This work
pKGCm	-	pACYC	Cm	This work
pORF 3	1240-1559	pSUP106	Cm	This work
pOU-PTF	1158-2027	pOU82	Ap	This work
pOU-RSF	5344-5879 *	pOU82	Ap	This work

* Nucleotide co-ordinates of the RSF1010 insert are given as published by Scholz *et al.*, 1989

Appendix 3

Nucleotide sequence of mutagenesis and PCR and sequencing oligonucleotides used in this study

NAME	SEQUENCE
MP1	5′-CCACAGGGCGGATCCAAGACTTTC T-3′
MP2	5′-GGTTGAACTCGAGCCCAGCCGCCGA-3′
MP3	5′-CTAGGCAAGATCTGAC CAGCAGACC-3′
-1212F	5′-CGCCAGGGTTTTCCCAGTCACGAC-3′
FP1	5′-AGTAGGGATCCGCGTACCGCTCGAACTGC-3′
FP2	5′-AGTAGGGATCCACTTCGGCGGGCAGTCGG-3′
FP3	5′-AGTAGGGATCCAAGAATCCGGGCCAGATG-3′
PTF-ECO	5′-ATATGAATTCTTGGCAAGTACCGGCAGC-3′
PTF-BAM	5′-TATTAGGATCCGCAATCCCCTGCCGCGCCGC-3′
RSF-ECO	5′-ATATGAATTCGGCATGGACAGGGGCGGGCC-3′
RSF-BAM	5′-AAATATAAGCTTGGATCCCCTTTGCAGGCAGTTGG-3′
3F	5′-TATTGAATTCGAACAGGAGCCGAAACATGC-3′
3R	5′-TAATGAAGCTTAGTTCAACCTTCCATGCCATAGC-3′
4F	5′-ATACGAATTCCTCGAAGAAGTGATGAAACGC-3′
4R	5′-TCGCCAAGCTTCGATCATTGGCGGTAAACCTCC-3′
5F	5′-TCCGGAATTCCGTAAGGAGGTTTACCGCC-3′
5R	5′-TTTCTAAGCTTCGCTCAATTTTCGCCCCTCG-3′
-1212F	5′-CGCCAGGGTTTTCCCAGTCACGAC-3′
LACZPRI	5′-CGCCAGCTGGCGAAAGGGGG-3′
TACPRI	5′-GACAATTAATCATCGGCTCG-3′

Appendix 4

Complete nucleotide sequence of the pTF-FC2 origin of replication starting at the *Cla*I site (position 1) and ending at the *Pst*I site of pTV100 (position 4434)

```
   1 ATCGATATTCACGACTGGCGCAACCGGGACAGCACGCTTGCCGCCTTGCA
  51 GCTATCGGCGCAGAAGTGGGGCAGCTTCACCGTAACCGGGAACGACGAAT
 101 ACAAGGCCATGTGCGGGAAGTTAGCAGCAGAGCACGGCTTCAAGATCACA
 151 AATGCAGAGCTTCAGGAGAGCATCCAGCAGGAGCGGCAGCGGATACAGCA
                     RepB»
 201 GGAGAGGGCGCAGGCGATGAAATCGGAGCAGCTAAAGCAGTTCGAGCGGT
 251 ACGCGGAAGCGGTAGGCGCGGAGCGCTACCGGGTAACGTCCATCAAGATG
 301 CAGGCAGACGGAAGGAAGCAAACCTTCATCCTCGACAAGAAGGACGGCAT
 351 CACGCGGGGTTTCACACCGCAGGAGATCGAGCAGCGCACGCCGGAGATGC
 401 AGCGCCTACAGCGCCGGGGCGAAAACCTCTACTACACGCCGCTATCGGAC
 451 AAGAAGCATCACATCCTCATCGACGACATGAACCGGGAGAAGCTGGAGCG
 501 GCTTATCAAAGACGGCTACCAGCCCGCCGCCGTGCTGGAATCCAGCCCCG
 551 GCAACTATCAGGCCATCATCACCGTGTCGAAGCTGGGGACGGCCCACGAT
 601 AAGGACGTGGGCAACCGCCTGAGCGATGCCCTGAACCGTGAATACGGCGA
 651 CCCGAAGCTATCGGGAGCCATCCACCCGCACCGCGCACCCGGCTACGAGA
 701 ACCGCAAGCCCAAGCACCAGCGGGAGGACGGCAGCTATCCAGAGGTGCGC
 751 TTGCTCAAGGCCGAGCGCCGGGAGTGCGTCAAGGCGCTGGCCTTGTCCAG
 801 CCAGATCGACGCCGAGTATCAGCGGCAAGCGGCCTTGAAGGCGCAGCAGC
 851 CCGAGCGCACGAAAGCCAAGCCCGCCTTGGAGCTTGCAGCGGCCAGCGGC
 901 AGCGCGATCGACGCCTACCAGCGGCATTACCGCGACGTGCTCAAGCGGCA
 951 GCGTGGCGGCGAGGTGGACTTGTCCCGCGTGGATTCCATGATTGCCGTGC
1001 GTATGCGCGTCACCGGCCACGATCAAGCGGCCATCGAGGGCGCTATCCGC
1051 CAGTGCGCACCGGCCACCCGGCAGAAAGACGAGGGCCGCGATTGGAACGA
1101 CTACGCGCAGCGCACCGCCCGCTATGCCTACAGCGCCGCACAGCCGCAAG
1151 CCGCCGATCTTGGCAAGTACCGGCAGCAGTGGGAGAAGCTGGAAGGGCGC
1201 GAGCCTGTACGACAGCAGGAGCAGGCAAAGGCGCAGAAGATCGAGCGCGA
1251 CAACTCGCCGGGAATGAGTCTCTAGCGTTGCGTGGTGGTTGTGATATACT
                                  PasA»
1301 TGTATAGCGTTTTCAGAACAGGAGCCGAAACATGCTTGCAATCCGACTGC
1351 CCGCCGAAGTGGAAACCCGCCTTGAAGCACTGGCGCAGGCCACAGGGCGG
1401 ACCAAGACTTTCTATGCCCGCGAAGCCATCCTTGAGCACTTGGATGACCT
```

1451 CGAAGATTTGTACCTTGCAGAGCAACGCCTGATCGACATTCGCGCAGGCA

PasB»

1501 AAACCCAAACCGTGCCACTCGAAGAAGTG**ATG**AAACGCTATGGCATGGAA

1551 GGTTGAACTCGACCCAGCCGCCGAGCGCGAGCTAGGCAAGATCGACCAGC

1601 AGACCGCCCGCCGCATCCTCGCTTTTTTGCATGGCCGTGTCGCCCAGCTC

1651 GACGACCCGCGCAGCATTGGCGAAGCCCTCAAAGGCTCCAAACTGGGAGC

1701 CTTCTGGAAATACCGCGTTGGGGATTGGCGAATCATCGCCAGCATCGAGG

1751 ACGGTGCTTTGCGCATCCTCGTTATGCGCATCGGCAATCGTAAGGAGGTT

PasC»

1801 TACCGCCA**ATG**ATCGAATACAGCTACCAGATCGACCCGCGCCCCTCCGAC

1851 CTTGGCGGCGGCTGGCGGTTGCGCCTGTTGGAAAGCGGCGAGGAAGTCGG

1901 CGGCGGAGTGTTCCCGTTGTCCGAGTACGCCACAGCAGAGAACGCAGAAG

1951 AAGCGGCCACGTACGCCTATGAGGACGCCTTGGCCGAGGCTTCGGCGTGG

2001 CTGGCATCGAGGGGCGAAAATTGAGCGGCGCGGCAGGGGATTGCGGCCCC

RepA»

2051 GGCAGCGCCTAACCACAACTGTCTGAAAAGGAGACAAGC**ATG**GCTTTAGA

2101 CATTATGGCGGCCTTCACCAATGAGCCGCCAGAACTTGATTTCATCTGGC

2151 CCGGATTCTTGGCCGGAACCGTGGGCGCACTTGTCGCACCTGGCGCAACT

2201 GGCAAGAGCTTTTTTGCTCTTGAAGCGGCCATGTCAATCGCTTGCAGTGT

2251 GGCAGGCGGCGACCTTGTGGGACTAACCCCGGCGCACACCGGGCGCGTGG

2301 TTTATCTCGCTGGCGAAGATCCACAGCCCGCCCTTGTGCGACGTGTCCAC

2351 GCCATCGGCCAGCACCTCAACCAGTCGGCCCGCGAAGCCATCGCTGAGAA

2401 CCTGATGCTTGAGCCGATCATGGGCAAGCGGCTAAACGTGATGGACGACG

2451 CGCACTTGCGCCGCGTCATCGACTACAGCGCAGGGGCCCGGCTGATTGTG

2501 CTGGACACCCTGAGCCGGATTCACATCCTCGACGAGAACAGCAATGGCGA

2551 CATGGCCCACCTTGTTTCCGTGTTGGAACACATCGCGGCGACCACCGGGG

2601 CGGCTGTCCTGTACCTGCACCACGTCAACAAGGGCAGCGCCCGCGACGGC

2651 CAGACCGACCAGCAGCAGGCAGCGCGGGGCGCGTCTGCCCTGATCGACAA

2701 CGCCAGATGGTGCGGCTATGTCGCCAAAATGACGGAGCAGGAAGCCGAGC

2751 GCATGAGTGACCGGGGCTTTGATCGTTCGCCATCCGGCAACGAGCGGCGC

2801 GGCCTTTTTGTCCGCTTTGGCGTGAGCAAGCAGAACTACGACGCGACCCC

2851 GCTAGACCGCTGGTATCAGCGGCACAGCGGCGGGGTGTTGTTGCCCGTTG

RepC»

2901 AACTACAGGAGGCAATCAGCA**ATG**GAGCAGGAAAAAAAGGGGGAAAGCGC

2951 AATGAGCTATGACCTCACCCATGCGCGGCACGACCCCGCGCATTGCCTCA

3001 CGCCGGGGCTTTTCCGCAGTCTCAAGCGCGGAGAACGAAAGAGGCTCAAG

3051 CTCGATGTGACCTACAACTACGGAGATGACTCAATCCGTTTTTGGGGGCC

3101 TGAACCACTTGGCGGCGATGACTTGCGCGTATTGCAAGGGCTGGTGGCAA

```
3151 TGGCTGCAATTTCCGGAGATAACGGGCGCGGCATCGTGCTACGGCACGAA
3201 ACGGAATCAGAAGCAGGCCAGCAACTCCGCCTATGGCTTGATATGCGGTG
3251 GGACGCCATAGAGAAAGATACGATGGTAGCCAAGGGCAGCTTCCGCCAGT
3301 TGGCCCGAGAACTTGGCTACGCCGAAGATGGAGGAAGTCAGTTTAAAACC
3351 ATCCGGGAAAGCATCGAACGGCTTTGGGCGGTATCGGTGATTGTCGAAAG
3401 AGGTGGTAAGCGGCAAGGGTTCCGCATTCTGTCCGAGTACGCGAGCGACG
3451 AGCAAGAAGGCAAGTTATTTGTTGCGCTTAATCCCCGGCTGGCGGATGCG
3501 GTTATGGGAGAGCGCCCGCACACCCGCATCAACATGGCAGAAGTTCGCAA
3551 GCTGGAAACAGACCCGGCACGGCTGCTACACCAGCGGCTATGTGGCTGGA
3601 TTGACCCCGGAAAGTCTGGCAAAGCTGAAATCGACACGCTGTGCGGTTAT
3651 GTATGGCCAGACGCAGCCAACGATGAAGCAATGAAAAAGCGCCGCCAGAC
3701 CGCGCGCAAGGCGCTTGTCGAGCTTGCCGCCGTTGGTTGGACGGTGAACG
3751 AGTACGCCAAGGGCAAGTGGGAAATCAGCAGGCCCAACCCCCGGCGTAAC
3801 GTTCCCCAACCCCCGGCGTAACGTTCCCCAACCCCCGGCGTAACGTTCCC
3851 CGCCGAAATCTGAAAAACCTAGCAACGGCGCGGGTTTGCGGGCGATTCGC
3901 AAAATCCCTCCATGATCTATCCAAGATCATCCACTAGGCGCGGTACTTTT
3951 CGCCGCCCTTGAGGGCGTCAAAAGTCCTTACCAAAAACCCCAAAAGAAAC
4001 GGCCGGGACAAGCCCGGCCAGACACACCCCGCCCCGCCCCGCTCGCCTTC
4051 ATTCTTCCACCGGGACAATGGACACCATCACCCGGTAGCGTTTGGCCTCT
4101 CCGGCAGGTAGCGCAGCGGCCAGCTTGGCGAGCGTTTCGGCTGGCTTGTC
4151 GGTCGTGCTTGTGGAGCACATCGCCTCATACCCGAACAGAAGCCATCAGA
4201 ATCGCCTACAGCGGATTTTTGGATGTTCTGGCTGCCTTGAGCTAGGGTTG
4251 GTAAAGAAAACGCCTATGGCTGTTTGCGGGGCTTCTGCGAGCATTGCCGG
4301 GACGGTCTTGGGCTTGCTTGTGCCGTTGAGGCGAAAAACGCCACCGCCAG
4351 GACAAGCAGGGGTCGTCTCAGAAAACGGAAAATAAAGCACGCTAAGCCGG
4401 TTGCAGCCGTAGCGGCCTGAACTCGCCCGCGCCGATC
```