CHARACTERISATION OF A REPLICON OF THE CONJUGATIVE, MULTIPLE DRUG RESISTANCE, MODERATELY PROMISCUOUS, PLASMID pGSH500

Fernanda Maria Palma Ribeiro da Silva Tatley

A dissertation submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy in the Faculty of Medicine, University of Cape Town.

Cape Town, August 1992

To my father, for whom no task was ever daunting

CONTENTS

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Dr. L. M. Steyn, for his support and encouragement. I am also most grateful to Professor A. A. Forder for arranging financial support and for giving me the opportunity of having been a member of the Department of Medical Microbiology.

I wish to thank Professor D. Rawlings for interesting and useful discussions, as well as for his critique of my work. In addition, I would also like to thank Dr. B. Elisha and Mrs. N. Douglass for our delightful "after hours" discussions. I would also like to thank all the other members of the Department of Medical Microbiology for their interest.

My deepest gratitude goes to my mother, my late father and my sister for all the love, prayers, encouragement and understanding I received from them. Lastly, but not least, I owe much to my husband for his understanding, unfailing support and for waiting for me after all the lonely nights.

I gratefully acknowledge the PhD Bursary/Scholarship received from the Medical Research Council, in addition to their grant supporting this reseach project.

ABBREVIATIONS

A	adenine
Ap^r	ampicillin resistance
ATP	adenosine triphosphate
bp	base pair(s)
C	cytosine
cop	copy number control locus
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
EDTA	ethylenediaminetetra-acetic acid
G	guanine
Gm^r	gentamicin resistance
IHF	integration host factor
inc	incompatibility control locus
IR	inverted repeat
Kb	kilobase pair(s)
KDa	KiloDalton
LA	Luria agar
LB	Luria broth
mg	milligram
ml	millilitre
M_r	relative molecular mass
mRNA	messenger RNA
MWM	molecular weight marker
ORF(s)	open reading frame(s)
*ori*C	chromosomal origin of replication
*ori*V	vegetative origin of replication
polA1	DNA polymerase I gene

rep	replicon or autonomous replication locus
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
spp	species
ssDNA	single stranded DNA
T	thymine
Tn	transposon
ts	temperature sensitive
μgm	microgram
μl	microlitre
UV	ultra-violet light

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

pGSH500 is a large plasmid (107 Kb), obtained from a nosocomial isolate of *Klebsiella pneumoniae* (APP). The plasmid is conjugative, encodes multiple drug-resistance and its copy number is estimated to be between 3 and 5 copies per host chromosome equivalent. pGSH500 has a moderately promiscuous host-range, being able to transfer, both naturally andt experimentally to members of the *Enterobacteriaceae* family. Its incompatibility group could not be identified by classical incompatibility assays or replicon typing. Thus, the incompatibility group of pGSH500 is presumed to be novel. The fact that pGSH500 was transmissible to a *polA1* deficient *E. coli* showed that its replication was *polA1* independent. Two *polA1* independent, autonomously replicating regions, repα and repß, were isolated from pGSH500 after shot-gun cloning of this plasmid in pUC19. repα was contained within a 3.8 Kb *Hind*III fragment (pFDT100), whereas repß was identified within a 3.3 Kb *Pst*I fragment (pFDT200). Restriction mapping analysis and hybridisation studies demonstrated that repα and repß were distinct. In addition, incompatibility assays confirmed that repα and repß are fully compatible with each other. Copy number determinations showed that repß (pFDT200) was dominant in *Escherichia coli*, since its copy number corresponded to that of pGSH500 in the same host. The 3.3 Kb *Pst*I fragment in pFDT200 (repß) was shown to function as a true mini-replicon when ligated to an antibiotic marker, in the absence of any vector sequences (pFDT200Gm^r). Functional analysis of repß suggested that all the necessary elements for replication, incompatibility and stable plasmid maintenance are contained within a 2.7 Kb DNA fragment (pFDT210). Sequence analysis of pFDT210 led to the delineation of the region containing the basic replicon of repß. This was shown to correspond to a 1.8 Kb sequence (nucleotides 478-2281) within a 2.2 Kb *Eco*RV-*Ava*II fragment of pFDT210. The copy number, incompatibility and host-range properties of repß are similar to those of pGSH500.

Like other basic replicons, repß consists of 3 domains. The first is a *cis*-acting 245 bp region, *cis-ori*V. This region contains an 90% AT-rich 70 bp sequence, 2 overlapping putative *Escherichia coli* DnaA protein-binding sequences (DnaA boxes) and several direct and inverted repeated sequences. These repeated sequences have the potential to form energetically favourable stem-loop structures. The second domain, *repB*, includes an open reading frame, ORF295, which has several in frame initiation sites and was shown to encode a 30.8 KDa protein (RepB) required for plasmid replication. The third domain is present immediately downstream from *repB* and was designated *inc/cop*. This domain consists of eighteen, 30 to 36 bp iterons. When cloned in pUC19, each iteron exerts incompatibility towards pFDT200, but not to pGSH500. Of the eighteen iterons present in repß, the fourteen most proximal to *repB* are involved in copy number control. Two open reading frames (ORF175 and ORF212) were found to overlap the *inc/cop* domain. No expression products were, however, obtained from these ORFs and their role in repß is unknown.

The molecular organisation of the repß is similar to that of the basic replicons of the narrow host-range plasmids F and P1 and to that of the broad-host range plasmid pCU1. The DNA sequence of repß shows a 53% identity to that of the mini-replicon of pCU1. Analysis of the predicted amino-acid sequence of the RepB protein shows a 58% identity with that of the Rep protein of pCU1 and a 31% identity with the RepE protein of the mini-F plasmid. The similarity of RepB to RepE of mini-F was unexpected as this is not reflected at the DNA level. Derivative plasmids with deletions of RepB can be complemented by both the Rep of pCU1 and the RepE of mini-F *in trans*. In addition, pGSH500 and repß are fully compatible with both pCU1 and mini-F. An 80% identity over 200 bp was also observed between the *cis*-acting *ori*V region of repß and the equivalent region of *ori-2* of mini-F. This identity stops abruptly before the direct repeats of *incB* (*ori-2*) and this may explain the compatibility of repß with mini-F.

The organisation of the repß in pGSH500 suggests that it is either a natural hybrid between a pCU1-like and an F-like replicon or an ancestral replicon from which the others have evolved. pCU1 and F are frequently encountered within the *Enterobacteriaceae* which appear to be the favoured hosts for pGSH500. Regardless of the evolutionary origin of repß, it appears to confer at least two advantages on pGSH500. Firstly, it is compatible with other plasmids occurring in the same strains. Secondly, in the event of *repB* being disrupted, pGSH500 can be complemented, at least, by some IncN and IncF plasmids. The success of pGSH500 may be further attributed to the presence of an additional replicon (repα), especially if this replicon is able to support replication of pGSH500, when repß cannot. Thus, it is possible that this plasmid has adapted to replicate within the *Enterobacteriaceae* flora, especially in the hospital setting. The prevalent use of antibiotics in the hospital environment may be sufficient selective pressure to ensure that pGSH500 is retained by its bacterial hosts.

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1

GENERAL INTRODUCTION

In the Gram-negative bacteria, plasmids are usually covalently closed circular double-stranded DNA molecules, capable of autonomous replication without integration into the host chromosome (Thomas, 1988; Scott, 1984). Plasmids are wide-spread among bacteria and may vary in length from 1.7 to 700 Kb (Yoon & Pack, 1987; Hogrefe & Friedrich, 1984). They may encode a number of non-essential functions and may confer an energy burden on their hosts (Cheah *et al.*, 1987; Zünd & Lebek, 1980). Some of the functions, however, such as drug resistance, degradation of organic substrates and virulence factors, may also impart selective advantage on the bacterial hosts by extending their versatility and adaptability (Thomas, 1987a; Womble & Rownd, 1988; Datta, 1985; Foster, 1983; Kües & Stahl, 1989). Plasmids are also important contributors to the evolution of prokaryotes by providing a mechanism for gene amplification and transfer within or between bacterial populations (Kües & Stahl, 1989).

Most naturally occurring plasmids are stably maintained in a bacterial population, even under non-selective conditions. This is especially true of the low copy number plasmids F, P1 and R1 (de Haan & Stouthamer, 1963; Jacob *et al.*, 1963; Prentki *et al.*, 1977; Engberg & Nordström, 1975), which offer no selective advantage to their host, yet, are lost at rates of less than one plasmid per million cell divisions (Nordström & Austin, 1989). This stability is achieved by two independent processes, replication and partition. Efficient plasmid-encoded replication ensures that the number of plasmids per cell is kept

constant (Kolter & Helinski, 1979; Nordström, 1985) whereas, accurate partitioning distributes plasmid copies to daughter cells (Nordström *et al.*, 1980; Novick, 1987; Nordström & Austin, 1989; Austin & Norsdtröm, 1990). In addition, other processes may also contribute to stable plasmid maintenance in a bacterial population. These include, the resolution of oligomers resulting from plasmid-plasmid recombination; the killing of plasmid free segregants; and infectious conjugal transfer (Austin *et al.*, 1981; Summers & Sherratt, 1984; Ogura & Higara, 1983a and 1983b ; Miki *et al.*, 1984, 1985b; Higara *et al.*, 1986; Gerdes *et al.*, 1986; Jaffe *et al.*, 1985; Lundquist & Levin, 1986).

Each plasmid is characterised by a specific copy number per host chromosome. Copy number is controlled by a plasmid-encoded system that determines the rate of initiation of replication (Novick, 1987) and may differ for the same plasmid in different hosts or under different growth conditions (Norsdtröm, 1990). Incompatibility, which is defined as the inability of two co-resident plasmids to be stably maintained in the absence of selective pressure (Novick *et al.*, 1976), results from plasmids competing for one or more elements of either the replication or partitioning systems.

Plasmids replicate in synchrony with the host chromosome (Nordström & Austin, 1989). In Gram-negative bacteria, plasmid replication depends on both host- and plasmid-encoded enzymes. Narrow host-range plasmids such as F, pSC101, P1, NR1 and R1, encode a single initiator protein that is essential for plasmid replication (Tokino *et al.*, 1986; Linder *et al.*, 1985; Chattoraj *et al.*, 1985a and 1985b; Miki *et al.*, 1980; Kollek *et al.*, 1978; Masai & Arai, 1987). In addition, these plasmids are largely dependent on host-encoded enzymes (Scott, 1984; Novick, 1987; Womble & Rownd, 1988) and can only replicate in a few related hosts. In contrast, promiscuous plasmids such as those belonging to the IncC, IncN, IncJ, IncP, IncQ and IncW groups, are capable of replication in a wide range of unrelated bacteria. These plasmids are largely independent of host-encoded enzymes and carry functions for initiation and regulation of replication, which are active in most Gram-

negative bacteria (Krishnapillai, 1986; Kües & Stahl, 1989). Promiscuous replication, rather than promiscuous conjugation, may be the crucial difference between narrow and broad host-range plasmids, as many of the former plasmids are capable of promoting promiscuous conjugation between members of unrelated bacterial genera (Krishnapillai, 1988). This has been demonstrated for some narrow host-range conjugative plasmids belonging to IncF, IncFII, $IncI_1$, IncM, IncT and IncX groups, which are transmissible from *Escherichia coli* to *Pseudomonas aeruginosa*, but are unable to replicate in the latter (Tardif & Grant, 1983; Guiney, 1982).

This review addresses two aspects of plasmid replication in Gram-negative bacteria. The first, is the functional organisation of the "basic replicon" of both narrow- and broad-host range plasmids. The second, is the manner in which functional organisation contributes to control of replication and host range. Models of plasmid replication for F, P1, RK2 and pCU1 are reviewed.

CONTROL OF PLASMID REPLICATION

THE BASIC REPLICON

The minimal region required for replication of a plasmid is designated the "basic replicon" (Kollek *et al.*, 1978). This definition was modified by Norsdtröm (1985) who proposed that the basic replicon is the smallest segment of a plasmid, which when ligated *in vitro* to a suitable selective marker, can form a mini-plasmid, or mini-replicon, which replicates autonomously with the same characteristics as the parental plasmid. Basic replicons are usually continuous segments of DNA of 1.0 to 2.5 Kb in length (Norsdtröm, 1985). Notable exceptions to this are the basic replicons of RK2 (IncP) and RSF1010 (IncQ). In RK2, three major areas, dispersed over 20 Kb, are involved in initiation and regulation of

replication (Barth, 1979; Thomas *et al.*, 1979; Thomas *et al.*, 1980), whereas the basic replicon of RSF1010 spans a 5 Kb fragment (Frey and Bagdasarian, 1989).

All basic replicons described to date consist of a *cis*-acting origin of replication (*ori*V) and a negative-feedback control system (*rep, cop/inc*; *Table 1.1*). An exception to this, however, is the basic replicon of ColE1, which lacks a *rep* gene (Donoghue & Sharp, 1978; Tomizawa *et al.*, 1977). The control system includes genes essential for replication (*rep*) as well as genes which regulate the initiation of replication and control copy number (*cop*). Thus, the interaction between the *rep* genes and *cop* loci, determines the plasmid copy number (Norsdtröm & Austin, 1989). In addition, the *cop* region is associated with plasmid incompatibility and is often designated *cop/inc*.

Finally, although most plasmids contain a single basic replicon, multiple replicons have been found in the large IncF plasmids (Bergquist *et al.*, 1986) and in some R plasmids (Grant *et al.*, 1980). A brief overview of salient features of the *ori*V, *rep* and *cop* loci is given below.

Table 1.1. Functions encoded by some basic replicons (adapted from Nordström, 1985).

Plasmid	Size of basic replicon (Kb)	Origin	Cop functions	*rep* genes	Source
ColE1	1-1.5	1	RNAI Rop Protein	None	1, 2
R1 (FII)	2.5	1	CopA-RNA CopB Protein	*repA*	3, 4 5
F	2.2	*ori-1* *ori-2*	 iterons	 *repE*	6 7, 8
R6K	4.0	α, ß,	iterons	*pir*	9
P1	2.1	1	iterons (*incA/ori* pairing)	*repA*	10, 11
RK2	5/20*	1	iterons TrfA	*trfA*	12

1, Tomizawa *et al.* (1977); 2, Tomizawa (1990); 3, Light & Molin (1983); 4, Womble *et al.* (1984); 5, Nordström & Nordström (1985); 6, Murotsu *et al.* (1981); 7, Tolun & Helinski (1981); 8, Tsutsui *et al.* (1983); 9, McEachern *et al.* (1989); 10, Abeles *et al.* (1984); 11, Abeles & Austin (1991); 12, Kittell & Helinski (1991). * Indicates the extent of dispersal of some regulatory elements in RK2.

*ori*V, Origin of replication

The origin of replication of a plasmid, *ori*V, is the site within the basic replicon where the DNA helix melts, allowing for initiation of daughter strand synthesis to occur. Replication from this site may proceed uni- or bi-directionally (Eichenlaub *et al.*, 1977 and 1981; Guyer *et al.*, 1976; Tomizawa *et al.*, 1974). The plasmids F and R6K have several origins of replication and although it is assumed that only one *ori* site is active in F, all three origins (α, ß and γ) are equally functional in R6K (Lane, 1981; Kolter, 1981).

The *ori*Vs of F, pSC101, RK2, P1 and R6K, show several characteristic features. The *oriV* tends to be AT-rich and may have repeated sequences, which are often the sites of action of an initiator protein (Murotsu *et al.*, 1984; Masson & Ray, 1988; Yamaguchi & Yamaguchi, 1984a and 1984b; Smith & Thomas, 1985; Abeles *et al.*, 1984 and 1989; Shon *et al.*, 1982). In addition, binding sites for host DnaA protein (DnaA boxes) have been identified in the plasmids F, P1, RK2, pSC101 and pCU1 (Kline *et al.*, 1986; Murakami *et al.*, 1987; Hansen & Yarmolinsky, 1986; Abeles *et al.*, 1990; Gaylo *et al.*, 1987; Vocke & Bastia, 1983; Krishnan & Iyer, 1990).

rep locus

The narrow host-range plasmid ColE1 does not encode any replicative functions except for a 0.6 Kb *cis*-acting origin (*ori*) (Backman *et al.*, 1978; Inselburg, 1974; Oka *et al.*, 1979; Tomizawa *et al.*, 1974 and 1977). Thus, this plasmid is entirely dependent on host-encoded proteins for its replication, including DNA polymerase I (Pol I). Plasmids such as ColE1, that require DNA polymerase I for initiation at the origin, are designated *polA1* dependent (Donoghue & Sharp, 1978; Tomizawa *et al.*, 1977).

Unlike ColE1 and related plasmids, the basic replicons of F, P1, pSC101, R6K and RK2 are *polA1* independent since they encode an replication initiator protein (Rep). Rep is responsible for the initiation of replication and is integrally involved in copy number control (Murotsu *et al.*, 1981; Trawick & Kline, 1985; Abeles *et al.*, 1984 and 1989; Linder *et al.* 1985; Thomas *et al.* 1979 and 1980; Durland *et al.*, 1990). The Rep proteins of F (Masson & Ray, 1986 and 1988), Rts1 (Kamio *et al.*, 1988), R6K (Filutowicz *et al.*, 1985; Kelley & Bastia, 1985), pSC101 (Linder *et al.*, 1985) and P1 (Abeles *et al.*, 1989) were shown to bind the direct repeat sequences (iterons) at the origin. Furthermore, the Rep proteins of F, P1, RK2 and R6K can function *in trans* (Tolun & Helinski, 1982; Chattoraj *et al.*, 1985a; Figurski & Helinski, 1979; Shafferman *et al.*, 1982).

inc/cop locus

Deletion and cloning analysis of several plasmids have led to the identification of plasmid regions which can act as negative control elements of plasmid replication. These include: genes for repressor proteins, antisense RNA; and DNA iterons. Examples of plasmids in which these elements have been recognised are listed in *Table 1.2.* These elements are involved in copy number control (*cop*) and incompatibility properties (*inc*).

Table 1.2. Negative control elements involved in control of plasmid replication (adapted from Nordström, 1990).

Negative control element	Function controlled	Plasmid
Repressor protein	Synthesis of Rep protein	λ*dv*[1,2]
Antisense RNA	Synthesis of Rep protein	IncFII[3, 4]
	Processing of preprimer	ColE1[5,6], p15[7]
DNA iterons	Synthesis of Rep protein or looping of replicon	P1[8], F[9], RK2[10], RSF1010[11], pSC101[12] R6K[13]

1, Matsubara (1976); 2, Murotsu & Matsubara, 1980; 3, Light & Molin (1983); 4, Öhman & Wagner (1991); 5, Tomizawa (1986); 6, Tomizawa *et al.*, 1981; 7, Chang & Cohen (1978); 8, Abeles & Austin (1991); 9, Rokeach *et al.* (1985); 10, Kittell & Helinski (1991); 11, Haring & Scherzinger (1989); 12, Vocke & Bastia (1985); ; 13, McEachern *et al.* (1989).

REPLICATION OF NARROW HOST RANGE PLASMIDS

F PLASMID

The F plasmid is the prototype sex factor of *E. coli*. It is a large (94.5 Kb), conjugative, low copy number, *polA1* independent plasmid (Sharp *et al.*, 1972; Lane, 1981). One of the most remarkable features of the F plasmid is its ability to be stably maintained within a bacterial population despite its low copy number of 1 to 2 per cell equivalent (Frame & Bishop, 1971; Collins & Pritchard, 1973). This stability is comparable to that of the *E. coli* chromosome and is of the same order of that of the plasmids P1, R1 and NR1 (Frame & Bishop, 1971; Prentki *et al.*, 1977; Austin *et al.*, 1981; Nordström *et al.*, 1984). Recently, Keasling *et al.* (1991) demonstrated that F replicates at a specific time during cell division and thus exhibits cell-cycle-specific replication. Cell-cycle-specific replication has been described for the *E. coli* chromosome and occurs only when a constant cell mass amount per origin is attained (Cooper & Hemlstetter, 1968). Similar rules were shown to apply for F replication, except that the initiation mass for the F plasmid is different from that for the chromosome.

The stability of F results from the interaction of its tightly controlled replication with other plasmids-encoded systems (Kline, 1988). These include an accurate partition mechanism mediated by 2 *sop* (stability of plasmid) proteins (*sopA* & *sopB*) and a *cis*-acting *sopC* (*incD*) site (Austin & Abeles, 1983; Mori *et al.*, 1986). In addition, a *ccd* locus (coupled cell division) that is involved in the killing of plasmid free cells (Ogura & Higara, 1983a and 1983b; Jaffe *et al.*, 1985) and an effective conjugative transfer (Loh *et al.*, 1988) also play a role in the stability of F. The presence of at least 2 complete basic replicons, repFIA and repFIB, as well as part of a third, repFIC (Bergquist *et al.*, 1986; *Figure 1.1.*) have also been proposed to contribute to the stability of F (Norsdtröm & Austin, 1989).

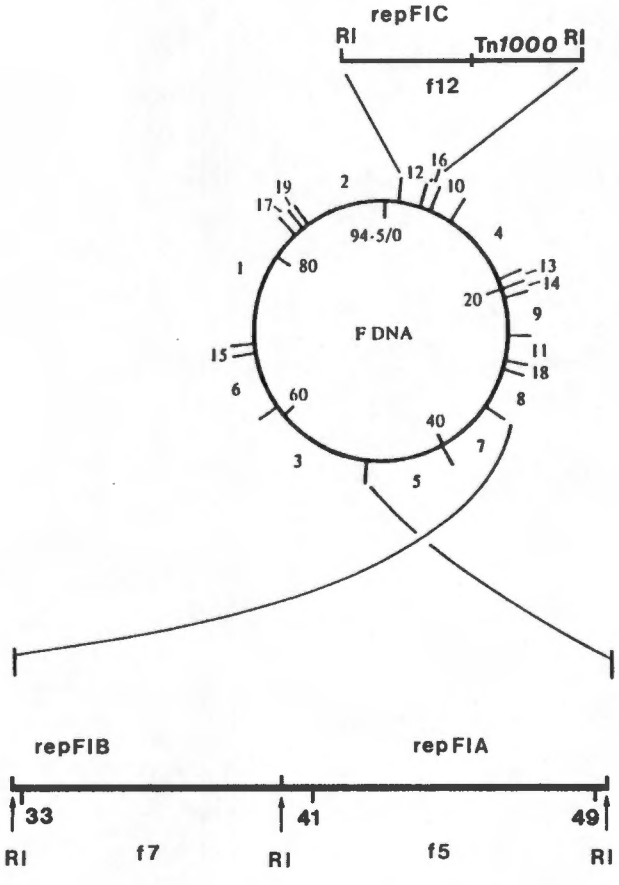


Figure 1.1. Restriction map of F (Skurray *et al.*, 1976). The numbers inside the plasmid refer to the F kilobase coordinates, whereas the numbers outside indicate the fragments generated by *Eco*RI digestion. The expanded regions show the fragments (f5, f7 and f12) containing the replicons of F (repFIA, repFIB and repFIC, respectively) (Bergquist *et al.*, 1982 and 1986). repFIC in F is incomplete and is interrupted by the insertion of Tn*1000* (Guyer, 1978). The abbreviation RI indicates *Eco*RI restriction sites.

Host factors required for F replication

The host encoded genes *dnaA* (Kline *et al.*, 1986), *dnaB* (Goebel & Schrempf, 1972), *dnaC* (Van Brunt *et al.*, 1977; Zeuthen & Pato, 1971) and *dnaE* (Pol III) (Thompson & Broda, 1973) are essential for F replication initiation.

Identification of the primary replicon of F

Heteroduplex analysis of deletion mutants of F plasmids (Davidson *et al.*, 1975) showed that an essential region for replication was present between the kilobase coordinates 43 to 53. This region was conserved in all of the F deletions capable of autonomous replication. Following the construction of an *Eco*RI restriction map of F (Skurray *et al.*, 1976), an autonomously replicating fragment, f5, corresponding to the coordinates 40.3 to 49.3 Kb was cloned independently by two laboratories (Timmis *et al.*, 1975; Lovett & Helinski, 1976) (*Figure 1.1.*). The resulting mini-plasmid of F, or mini-F, formed a true replicon when ligated to an antibiotic marker. It displayed the same copy number, *polA1* independence, incompatibility, stability of inheritance and sensitivity to acridine orange, as native F (Lovett & Helinski, 1976; Timmis *et al.*, 1975; Guyer *et al.*, 1976). Since the properties of mini-F mimic those of F, it is accepted that mini-F is the primary replication system of F under normal circumstances (Lane, 1981).

In addition to mini-F (f5), a second autonomously replicating fragment, *Eco*RI f7 (32.8-40.3 F Kb coordinates; *Figure 1.1*), was identified in F (Lane & Gardner, 1979). These fragments (f5 and f7) were subsequently designated repFIA and repFIB, respectively (Picken *et al.*, 1984; *Figure 1.1.*). Replication of f7 plasmids differs from that of f5 mini-plasmids in several respects. Unlike f5, the replication of f7 is *polA1* dependent. In addition, the f7 fragment is less stable than f5, being easily lost from fast growing bacterial populations. Furthermore, the maintenance of f7 is insensitive to acridine orange. Although

these fragments can replicate independently of each other and there is no evidence that repFIA and repFIB interact during the normal maintenance of F, this possibility cannot be excluded (Lane, 1981).

A third replicon, repFIC, was identified as a second basic replicon of EntP307 (Picken *et al.*, 1984; Saadi *et al.*, 1984). This replicon was later shown to be present in F (*Eco*RI f12, 5.15 Kb) (Bergquist *et al.*, 1986). In F, repFIC is incomplete, interrupted by the insertion of Tn*1000* (Guyer, 1978) and shows sequence homology to one of the replicons of plasmids belonging to a different incompatibility group, repFIIA (Saadi *et al.*, 1984).

The functional organisation and replication control of the basic replicon of mini-F are discussed below.

Functional organisation of the basic replicon of mini-F

The mini-F plasmid (Timmis *et al.*, 1975; Lovett & Helinski, 1976) was shown to contain 2 different replicating origins, which were designated *ori*V and *ori*S (Eichenlaub *et al.*, 1977; Manis & Kline, 1977). Of these, only *ori*S forms part of the basic mini-F replicon (see later) (Murotsu *et al.*, 1981). A brief description of the characterisation of these origins is included.

*ori*V (*ori-1*) and *ori*S (*ori-2*) of mini-F

The *oriV* of mini-F was mapped by Eichenlaub *et al.* (1977) to the 42.6 F Kb coordinate. Since this origin was the first to be recognised in mini-F, it is also known as *ori-1*. The *ori*V (*ori-1*) region, however, could be deleted without impairment of autonomous replication (Manis & Kline, 1977) and a second origin, *oriS* (*ori-2*), was identified at the 45.07 Kb F coordinate (Eichenlaub *et al.*, 1981; Bergquist *et al.*, 1981). Replication from

ori-1 is bi-directional, whereas that from *ori-2* is unidirectional to the left (Eichenlaub *et al.*, 1977 and 1981).

Originally it was proposed that the origin of choice for repFIA replication was *ori-1,* since *ori-2*-initiated replication was only observed when *ori-1* was deleted from mini-F (Eichenlaub *et al.*, 1981). This was supported by experimental evidence showing that blocking of replication from the *ori-1*, did not result in replication from *ori-2* (Eichenlaub & Wehlmann, 1980). Despite this, *ori-2*-initiated replication of the primary replicon mimics F replication and is currently accepted as the favoured origin (Willetts & Skurray, 1987; Kline, 1988).

The basic replicon

The smallest mini-F (f5) segment satisfying the definition of a basic replicon (repFIA) is contained within the 44.1 to 46.5 Kb coordinates of F (Murotsu *et al.*, 1981; Tsutsui *et al.*, 1983; Murotsu *et al.*, 1984; Bergquist *et al.*, 1986). The molecular organisation of repFIA (2.25 Kb) is illustrated in *Figure 1.2.*

repFIA includes a 217 bp *cis*-acting origin of replication or minimal replication region, which corresponds to *ori-2* (Murotsu *et al.*, 1981; Murotsu *et al.*, 1984) (*Figures 1.2. & 1.3.*). The *ori-2* consists of two DnaA boxes (Fuller *et al.*, 1984) and a 90% A-T rich 50 bp sequence followed by four 19 bp iterons (*Figures 1.2 & 1.3.*). These iterons are responsible for the incompatibility properties of *ori-2* and are designated *incB* (Tsutsui *et al.*, 1983). Immediately downstream of the *ori-2* lies an open reading frame (*repE*) (Murotsu *et al.*, 1981), also designated *copA* (Kline & Seelke, 1982; Seelke *et al.*, 1982). The *repE* gene was shown to encode a 29 KDa protein which is essential for mini-F maintenance (Ebbers & Eichenlaub, 1981; Watson *et al.*, 1982; Tolun & Helinski, 1982; Maki *et al.*, 1983; Murotsu *et al.*, 1984). Five 19 bp iterons (*incC*) , similar to those in

incB, are also found downstream of *repE*. These are oriented in the opposite orientation to the *ori-2* iterons and are involved in incompatibility and copy number control (Tolun & Helinski, 1981; Tsutsui *et al.*, 1983). This iteron region was designated *incC/copB* and is associated with the phenotypic sensitivity to acridine orange (*aos*) (Wechsler & Kline, 1980). The consensus sequence for the iterons within *ori-2* (*incB*) and those of *incC (copB*) is CTGTGACAAATTACCCTCA (Murotsu *et al.*, 1981). The minimal autonomously replicating region of mini-F includes *ori-2* (*incB*) and *repE* (*copA*) and spans approximately 1.2 Kb (Tsutsui *et al.*, 1983). The 19 bp repetas of *incB* and *incC*, as well as the *ori-2* region are the targets of RepE (Masson & Ray, 1986; Tokino *et al.*, 1986).

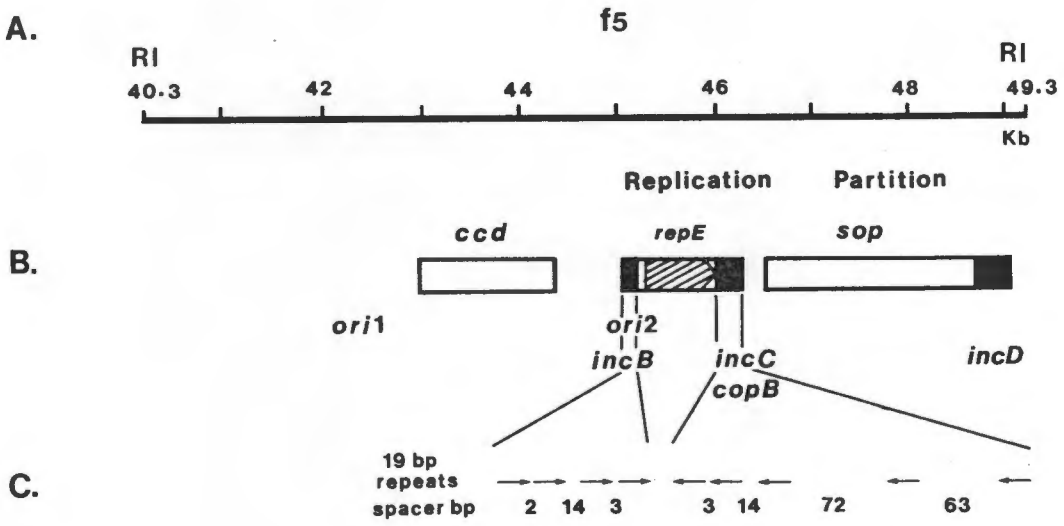


Figure 1.2. Molecular organisation repFIA within mini-F (f5). **A.** The F kilobase coordinates of f5 are shown. **B.** The basic replicon, repFIA, includes the *repE* gene (thick, dashed arrow) flanked by the origin of replication (*ori-2/incB,* narrow black box) and the copy number control locus (*incC/copB*, intermediate size black box) (Eichenlaub *et al.*, 1981; Murotsu *et al.*, 1981; Komai *et al.*, 1982; Murotsu *et al,* 1984). The *repE* gene specifies a 29 KDa initiator protein RepE. Regions of f5 not involved in replication control include the coupled cell division locus (*ccd*) (open box; Bex *et al.*, 1983; Ogura & Higara, 1983a; Jaffe *et al.*, 1985) and the partition locus (*sop*, *incD*) (mixed open/black box; Ogura & Higara, 1983b and Mori *et al.*, 1986). The *incB (ori-2)* and *incC (copB*) loci represent iteron-specified incompatibility determinants (Tsutsui *et al.* 1983), whereas *incD* represents iteron-specified partitioning incompatibility (Helsberg & Eichenlaub, 1986; Lane *et al.*, 1987). The position of *ori-1* (additional replication origin outside of the basic replicon) is also indicated (Eichenlaub *et al.*, 1977 and 1981). **C.** Organisation of the repFIA iterons involved in incompatibility (*incB*, *incC*) and copy number control (*incC/copB*). Numbers between repeats represent spacer nucleotides (Murotsu *et al.*, 1981; Murotsu *et al.*, 1984).

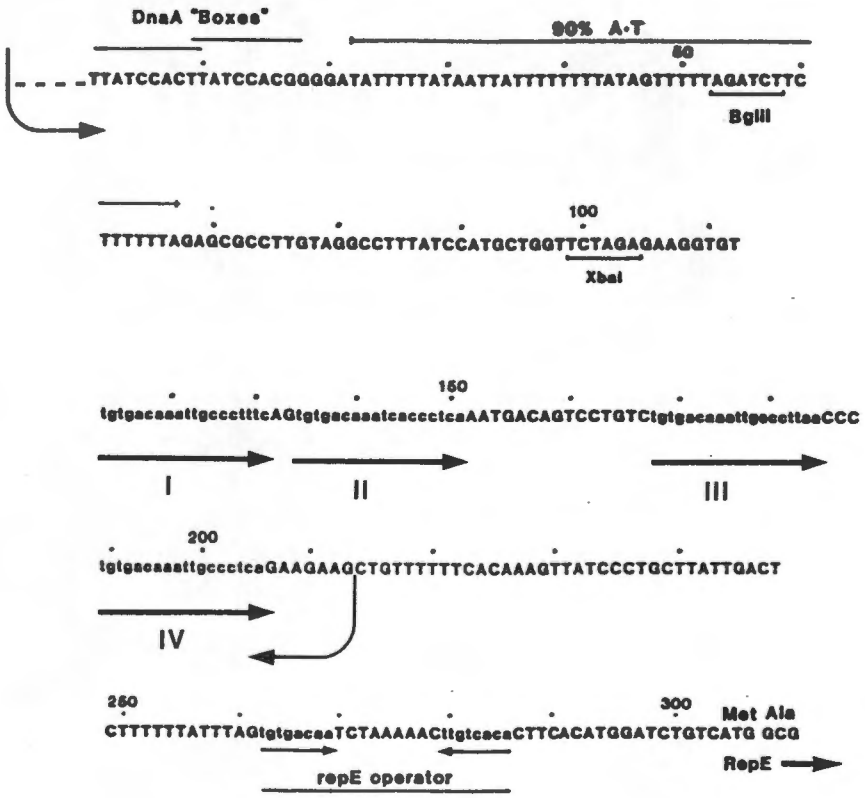


Figure 1.3. Sequence of the F *ori-2* (217 bp; between curved arrows) (Murotsu *et al.*, 1984). DnaA boxes are DnaA protein binding sites and are located at the left end of *ori-2*. A 90% AT-rich region (50 bp) and a spacer (approximately 50 bp) precede four 19 bp direct repeats (**I, II, III, IV**; lower case letters). *repE* operator (promoter) and start of *repE* are also shown (modified after Kline, 1988).

Models for control of mini-F replication

Model 1. ***The initiator-titration model.*** The replication of F is negatively controlled by a "switch-off" mechanism (Tsutsui & Matsubara, 1981; Kline & Seelke, 1982), which prevents replication when the plasmid copy number doubles. A model to explain this observation was proposed by Tsutsui *et al.* (1983). They proposed that the repeated sequences within *incB/incC* regions "titrate" a positive acting element that is essential for replication. According to this model, replication will be turned off when all titrating sites are occupied by the positive element. A fixed plasmid copy number could be obtained if the positive factor would be supplied at a constant rate. This would be possible if the positive factor were autoregulated.

The positive factor was postulated to correspond to the RepE protein whereas the titrating repeats correspond to the *incC/copB* and the *ori-2* (*incB*) regions (Tsutsui *et al.*, 1983; Kline & Trawick, 1983). RepE was thus proposed to be rate-limiting for initiation of replication (Tsutsui *et al.*, 1983).

These proposals were consistent with results previously obtained by Bergquist *et al.* (1981) and Seelke *et al.* (1982) who showed that *copA* mutations within *repE*, led to loss of copy number control. Furthermore, since the *copA* mutations of *repE* prevented suppression of the *repE* promoter, it suggested that RepE may be an autorepressor as well as being essential for mini-F replication (Trawick & Kline, 1985). This is in agreement with the autoregulatory concept of Sϕgaard-Andersen *et al.* (1984) and was also supported by the findings of Rokeach *et al.* (1985) and Bex *et al.* (1986) on the binding of RepE to *incB* and *incC* as well as on the autoregulatory properties of *repE*. Indeed, Rokeach *et al.* (1985) proposed that a model for the control of replication of F has to satisfy 3 conditions: 1. that RepE is a rate-limiting initiator; 2. that the expression of *repE* is autoregulated; 3. that RepE binds to *incB/ori-2* and to *incC/copB*.

By definition, autoregulation of *repE* expression would be expected to maintain low levels of free RepE in the cell, since synthesis should only occur if more RepE binding sites become available. Thus, a set steady-state equilibrium between free and bound RepE would be maintained. Sequestering of RepE by both the *ori-2* and *copB* loci would lead to a decrease in the free RepE concentration. To compensate for this shift in equilibrium, the RepE operator would initiate another round of synthesis. This would result in another round of replication with consequent sequestration of RepE, once more at the *ori-2* and *copB* loci, to set a new steady-state equilibrium. This chain of events would inevitably lead to run-way rather than controlled replication as is observed. To resolve this paradox, Trawick & Kline (1985) proposed two possible variations to the titration model, based on the autorepressor/initiator operon model of Sompayrac & Maalφe (1973) (*Figure 1.4A.*). These variations are discussed in the next two sections (**Model 2.** and **3.**) and are illustrated in *Figure 1.4B. & C.*

Model 2. ***Independent initiator-repressor model.*** This model is represented in *Figure 1.4B.* There are two additional open reading frames within *repE* which could encode two 9 KDa polypeptides (Murotsu *et al.*, 1981). If one of these polypeptides acted as the autorepressor, then RepE was free to bind at to *ori* and to *copB*. Thus the autorepressor function would be independent of sequestration. However, since there is no evidence for the expression of these polypeptides, this model is not accepted.

Model 3. ***The two-stage molecular model.*** An alternative model was proposed (Trawick & Kline, 1985), whereby the RepE protein can exist in two forms, Ei and Er (*Figure 1.4C.*). The RepE is translated in its repressor form (Er) and is converted irreversibly to the initiator form (Ei). This model presumes that the formation of Ei is dependent on both the Er concentration and on the rate constant for its conversion to Ei. This implies that binding of Ei to the target sites has no effect on the synthesis of Er with corresponding conversion

to Ei. As a consequence, autoregulation becomes a process distinct from sequestration and would prevent an increase in *repE* expression. The concentration of RepE would be rate-limiting for plasmid replication. Furthermore, this would ensure maintenance of the correct copy number. A computer simulation of this model faithfully reproduced the behaviour of F (Womble & Rownd 1987).

Although this model is attractive, no direct evidence for irreversible conversion of RepE forms from Er to Ei has been demonstrated (Tokino *et al.*, 1986; Masson & Ray, 1988). Recently, however, Kline *et al.* (1992) purified two naturally occurring proteolytic forms of RepE (Δ1- and Δ17-RepE). In the presence of *E. coli* chromosome, the Δ17-RepE forms bound *ori-2* and *incC* with greater affinity than Δ1-RepE. This resulted in an inhibitory effect to replication. In addition, one of the modified proteins (Δ17-RepE) also functions as an incompatibility determinant. These results suggest that proteolytic forms the of RepE differ from Er and Ei postulated in the two-stage molecular model (Trawick & Kline, 1985). Consequently, Kline *et al.* (1992) proposed that RepE processing may be a novel mechanism involved in DNA replication control of F.

In summary, RepE is a bi-functional protein, acting both as an autorepressor of its own transcription (Sϕgaard-Andersen *et al.*, 1984; Trawick & Kline, 1985) and as an initiator of replication (Helsberg *et al.*, 1985; Muraiso *et al.*, 1987; Masson & Ray, 1988). Autorepression is achieved by binding to the *repE* gene, whereas initiation of replication follows binding to the iterons of *ori-2* and *incC*. Alternative models of iteron regulation of replication have been described (Nordström, 1990), although none has been tested for F. Clearly the replication control of F requires further investigation and may yet prove to be similar to that currently accepted for mini-P1 (see later).

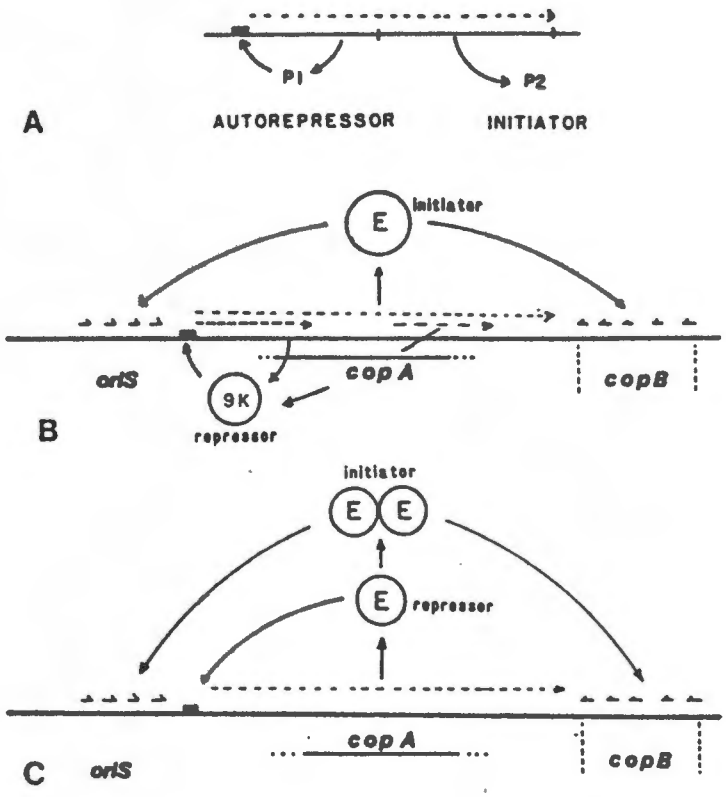


Figure 1.4. Models for control of chromosomal and plasmid DNA replication. **A.** The autoregulatory model control of Sompayrac & Maalϕe (1973). In this model, P1 is the autorepressor and P2 is the initiator. P1 and P2 are transcribed from the same open reading frame. Transcription is regulated only by P1 binding to the promoter-operator region of P1 (black box). **B.** Model for the control of mini-F replication involving RepE (E) and a putative 9 KDa protein. Transcription of both proteins is regulated from the same promoter, that of RepE (black box). In this model, the 9 KDa protein is the promoter repressor and RepE the initiator of DNA replication (Kline, 1985; Trawick & Kline, 1985). **C.** The same model as in (**B.**), though RepE is both the repressor (monomer unit) and the initiator (dimer unit). This monomer /dimer relationship is a symbolic, not a quantitative representation of the irreversible transition from repressor to initiator. The (⟶) represent iterons, which are the binding sites for RepE. *oriS* is an alternative designation for *ori-2*. *copA* is an alternative designation for the replication domain with copy number involvement. *copB* represents the domain containing the iterons associated with plasmid copy number control. The dashed lines represent mRNA transcription (Kline, 1985; Trawick & Kline, 1985).

Incompatibility: a consequence of replication control

The repFIA iterons express incompatibility in direct proportion to their numbers (Tolun & Helinski, 1981; Tsutsui *et al.*, 1983). This is related to the finding that deletions in *copB* lead to an increase in copy number of mini-F that is proportional to the number of deleted iterons (Tolun & Helinski, 1981; Tsutsui *et al.*, 1983; Kline & Trawick, 1983). Because the rate of formation of the initiator is assumed to be constant (Tsutsui *et al.*, 1983), excess iteron sequences retard or inhibit the activation of the origin. In contrast, a deficit in the number of iterons diminishes the interval between successive rounds of replication. Incompatibility between two related plasmids would result from the competition of each others initiator for the iterons on either of the plasmids. This could result in saturation of the binding sites of one plasmid by the Rep proteins of both plasmids, leaving available sites in the other plasmid. Furthermore, this problem might be compounded since RepE has been postulated to assume two forms (initiator and repressor) (Trawick & Kline, 1985). Thus competition might also arise between the RepE forms for binding on the two plasmids.

P1 PLASMID

P1 is a large plasmid of about 90 Kb and is the prophage form of a temperate bacteriophage (Bachi & Arber, 1977; Yun & Vapnek, 1977). It belongs to the incompatibility group Y (Hedges *et al.*, 1975). It is non-conjugative, but can be mobilised by helper plasmids (Boice & Luria, 1963).

P1 is stably maintained as a single copy plasmid (Prentki *et al.*, 1977), being lost from its hosts at a frequency of less than 10^{-5} per generation (Rosner, 1972; Austin *et al.*, 1981). This is suggestive of tightly regulated replication as well as partition systems (Austin &

Abeles, 1983) and is reflected by its incompatibility properties towards related plasmids (Sternberg & Austin, 1981; Abeles *et al.*, 1984). Other functions of P1 which contribute to this stability include a partition locus which comprises two open reading frames (*parA, parB*) and their target, a *cis*-acting par site (*parS/incB*) (Abeles *et al.*, 1985; Davis & Austin, 1988). In addition, a *ccd* locus (coupled cell division locus) with kill properties similar to those of F is also present in P1 (Capage & Scott, 1983). Furthermore, the *cre* gene and the *loxP* sites are involved in site specific recombination to promote stability by resolving P1 multimers (Austin *et al.*, 1981).

Host functions required for P1 replication

P1 replication is *polA1* independent (Abeles *et al.*, 1984), but requires other host factors. Among these are the *dnaA* (Hansen & Yarmolinsky, 1986; Wickner & Chattoraj, 1987), *dnaB, dnaC and dnaG* (Scott & Vapnek, 1980), *dnaJ, dnaK and grpE* gene products (Tilly & Yarmolinsky, 1989; Wickner, 1990).

Isolation of the mini-P1 replicon

Although P1 exhibits typical plasmid properties in the prophage state, its dual role as a phage must impose certain constrains on its maintenance. Thus, complex plasmid replication and lytic cycle controls are present in P1 (Sternberg & Austin, 1981). Independent controls for these processes had been proposed by Novick (1969) and were confirmed by Austin *et al.* (1978) who demonstrated that maintenance of the plasmid form of P1 did not require any of the factors involved in the lytic cycle.

A physical map of P1 displaying plasmid maintenance functions as well as *Eco*RI restriction sites is shown in *Figure 1.5*. Deletions of P1 were used to locate these functions. Of the three deleted forms of P1 shown in *Figure 1.5*, P1ΔN19 (Austin *et al.*, 1978) and pIH1972 (Shafferman *et al.*, 1979) are able to replicate autonomously. Since these two deletion plasmids overlapped a region fully contained within the *Eco*RI-5 fragment, it seemed logical that cloning this fragment into a vector would generate a P1 mini-replicon. However, this deduction proved to be incorrect, since when cloned into a lambda vector, the *Eco*RI-5 fragment could not support replication (Sternberg, 1979; Sternberg & Austin, 1981). The paradox was resolved when it was realised that P1ΔN19 and pIH1972 are not analogous plasmids and that each contained a distinct replicating region of P1 present at either end of the *Eco*RI-5 fragment. Therefore, essential functions for replication occurred to the right (lambda-P1:5R) and to the left (lambda-P1:5L) of the *Eco*RI-5 fragment (Sternberg, 1979). Since it became evident that P1ΔN19 and lambda-P1:5R contained the same region they were classified as "right-side" replication mode plasmids, whereas lambda-P1:5L and pIH1972 were designated "left-side" replication mode plasmids (Sternberg & Austin, 1981).

The "right-side" replication mode plasmids define a region that contains 2 adjacent *Eco*RI fragments, 5 and 8 (*Figure 1.5.*), the latter of which contain all the elements required for the "right-side" mode of replication. Furthermore, since this mode of replication closely resembles that of P1 itself and the copy number of the "right-side" plasmids is one per host chromosome equivalent (Austin *et al.*, 1978; Sternberg & Austin, 1981), the "right-side" replication mode plasmids contain the regions that control the replication of P1 (Abeles *et al.*, 1984). In addition to the replication determinants, a partition locus of P1 was also identified within lambda-P1:5R (Austin *et al.*, 1982). In contrast, the "left-side" replication mode plasmids lambda-P1:5L and pIH1972 have a high copy number and are unstable in Rec+ hosts (Sternberg & Austin, 1981).

The lambda-P1:5R plasmid became known as the P1 replicon or mini-P1.

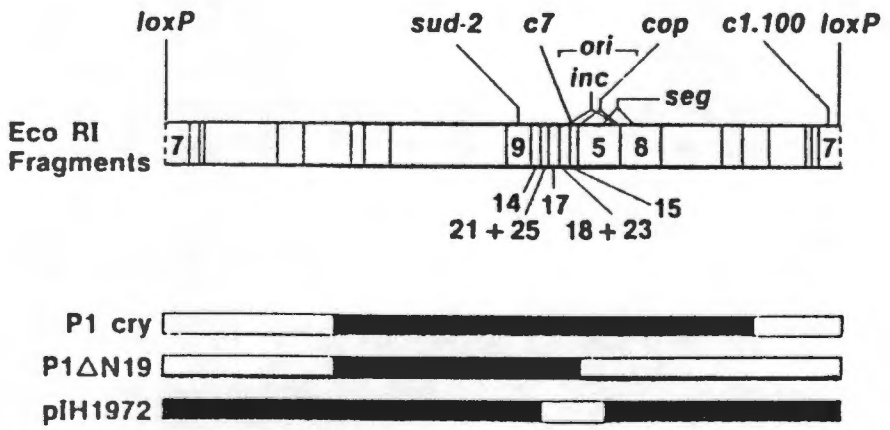


Figure 1.5. The physical map of P1. The numbered regions indicate the *Eco*RI digest fragments as described by Bachi & Arber (1977) and Sternberg *et al.* (1979). The regions associated with replication functions include: *ori*, origin(s) of replication; *inc* (*incA* and *incB*), incompatibility domain; *cop*, site of copy number mutations. Other loci not directly involved in replication but that contribute to stable plasmid maintenance are also indicated: *loxP*, site-specific recombination, promotes plasmid retention; *sud2*, weak virulent mutation and low level of plasmid loss; *c7* (c7.765), temperature-dependent plasmid loss in P7 or P1-P7 hybrids; *seg*, site of replication or partition mutants causing incorrect segregation; *c*1.100, temperature-sensitive lytic repressor. Deletion maps: black bars indicate deleted regions (Sternberg & Austin, 1981).

The basic replicon

The mini-P1 replicon contains 11 Kb of P1 DNA (*Eco*RI-5 and -8 fragments), of which only 2.1 Kb are required to promote P1 specified replication (basic replicon) (Abeles *et al.*, 1984) (*Figure 1.6.*).

The P1 basic replicon consists of three contiguous elements. These are the a 250 bp *cis*-acting origin, *oriR* (*incC*), the gene encoding the 32 KDa *repA* protein and a copy-number control locus, *incA*, that is also involved in the expression of incompatibility (Abeles *et al.*, 1984; Chattoraj *et al.*, 1985b). The *oriR* region is sufficient for replication when RepA is supplied *in trans* (Chattoraj *et al.*, 1985b). *incC* is the alternative designation for *ori*R, since it is an effective incompatibility determinant when presented *in trans* in a high copy number vector (Abeles *et al.*, 1984).

The structural organisation of the *oriR* is shown in *Figure 1.7.* (Brendler *et al.*, 1991). The *oriR* contains two DnaA boxes that are binding sites for DnaA protein, which is essential for P1 origin function (Hansen & Yarmolinsky, 1986). Recently it has been demonstrated that only one of these DnaA boxes is required for *oriR* function (Abeles *et al.*, 1990). Downstream from these boxes are five tandem repeats of 7 bp. Four of these contain GATC sites that are substrates for the host DNA adenine methylase (Hattman *et al.*, 1978). A fifth GATC sequence is also present outside and to the right of the last repeat. Activation of the origin requires methylation both *in vivo* and *in vitro* (Abeles & Austin, 1987). Adjacent to the *dam* methylation sites is a 39-bp GC-rich region, which is immediately followed by five 19 bp repeats, separated from each other by two helix turns, over a stretch of 104 bp (Abeles *et al.*, 1984). The repeats, which overlap the promoter for the *repA* gene, are the site for RepA autoregulation (Chattoraj *et al.*, 1985b; Abeles, 1986). Between these repeats and immediately before the start of the *repA* gene are three more

potential DnaA boxes (Brendler *et al.*, 1991). These are outside the boundary of the *oriR* (Chattoraj *et al.*, 1985b).

Downstream from the *repA* gene, the *incA* region contains nine 19 bp direct repeats spaced by three to four helix turns over a 285 nucleotide sequence. This region is not essential for replication, though plasmids lacking *incA* have an eight-fold higher copy number than the wild-type (Chattoraj *et al.*, 1984; Pal *et al.*, 1986). This *incA* copy-control locus was also shown to be a binding site for RepA (Abeles, 1986) and to negatively regulate the replication of P1 by titrating the essential replication protein, RepA (Chattoraj *et al.*, 1984; Chattoraj *et al.*, 1985b).

In Summary, the two incompatibility loci, *incA* and *incC*, which are the targets of RepA, flank the *repA* gene. The RepA protein has three roles (Abeles *et al.*, 1989). Firstly, it acts as a replication initiator by binding to the origin (*incC)*. Secondly, it acts as a repressor of its own transcription (Chattoraj *et al.*, 1985b) by excluding RNA polymerase from the promoter (Swack *et al.*, 1987). Thirdly, RepA interacts with the copy-control locus resulting in a negative control of initiation (Abeles, 1986; Pal *et al.*, 1986).

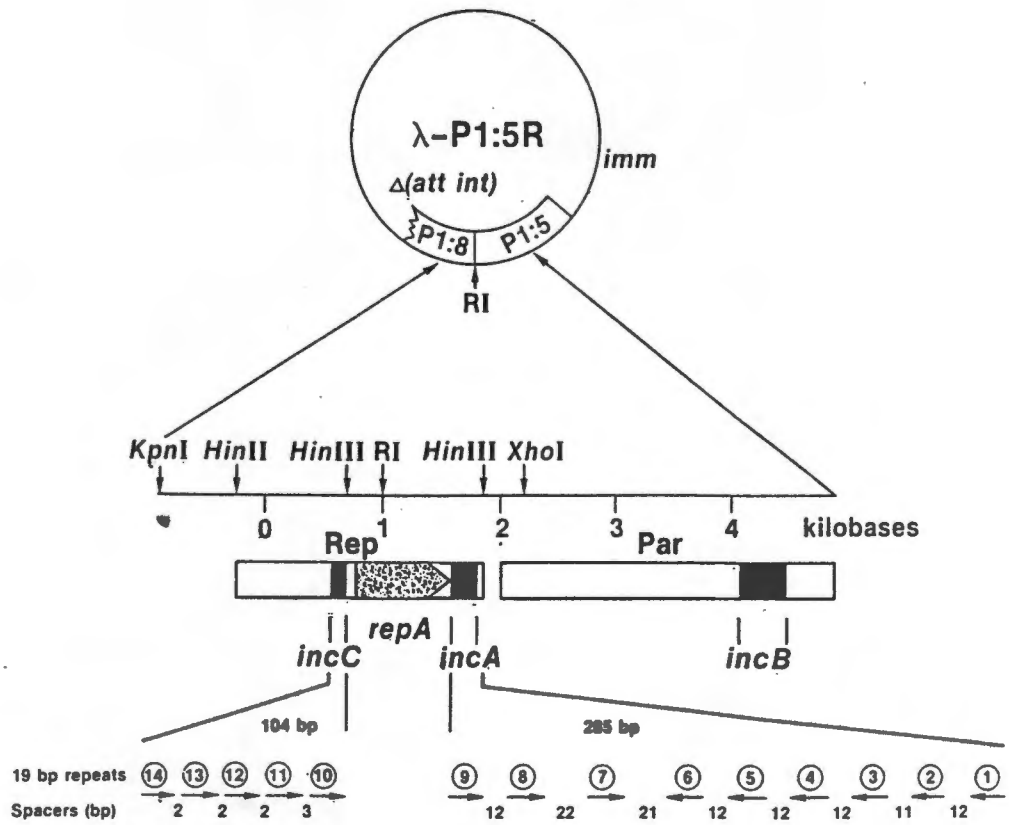


Figure 1.6. Functional organisation of the basic replicon of mini-P1 plasmid in _-P1:5R phage. A 2.1 Kb *Hin*II-*Hin*III fragment, marked Rep, is sufficient to allow for unit-copy number replication of the plasmid. The Rep region contains two sets of 19 bp repeats (*incC* and *incA*). These repeats are identified by circled numbers, whereas the numbers below the repeats indicate nucleotide spacers. *incC* comprises five repeated sequences (104 bp), *incA* includes nine repeats (285 bp). *incC* and *incA* are associated with expression of incompatibility and copy number control. These loci flank the gene for the initiator RepA (stippled area). Adjacent to the basic replicon is a region labelled Par. This is involved in equipartition and also includes an incompatibility determinant (*incB*) (Chattoraj *et al.*, 1984).

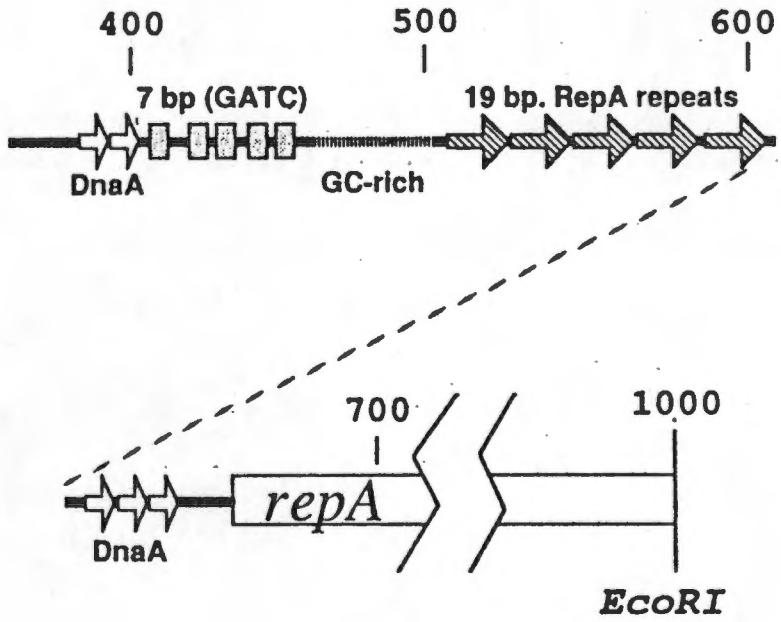


Figure 1.7. P1 plasmid replication origin, *ori*R. The base pair coordinates are as described in Abeles *et al.* (1984). The boxes represent the five 7-bp imperfect repeats containing GATC adenine methylation sites (GATC motif absent in second box). Open arrows are potential DnaA boxes. Shaded arrows are 19-bp RepA-binding repeats. Vertical shading marks the location of the 39-bp GC-rich sequence. The upper part of the diagram represents the minimal origin (Chattoraj *et al.*, 1985b), whereas the lower part shows associated sequences, including the rightward set of DnaA boxes and the start of *repA* gene (Brendler *et al.*, 1991).

Assembly of the replication complex at the P1 origin

A model has been proposed for the initiation of DNA synthesis at the P1 origin (Brendler *et al.*, 1991). This is represented in *Figure 1.8.* Two major elements are involved in this process. The first is the binding site (five closely spaced 19 bp repeats, *incC*; *Figure 1.7.*) for the initiator (RepA). This site is situated on the right side of *oriR*. The second is the strand opening region containing the five 7 bp imperfect repeats with their *dam* sites (melting site). Unlike the typical melting site within the *oriC* of *E. coli* (Bramhill & Kornberg, 1988), this site in P1 is not AT-rich. An additional region, the GC-spacer is crucial for the correct physical interaction between the RepA recognition site and the strand melting site.

This model proposes that RepA binds to its DNA target, wrapping it around a protein core. Specific contacts between the wrapped complex and the 7 bp region induce strand melting. The GC-rich spacer ensures that strand opening does not spread towards the wrapped complex. This clamp effect might be responsible for directing the replication fork unidirectionally to the left (Wickner and Chattoraj, 1987). The DnaA boxes, which are not part of the complex and whose positions are not critical (Abeles *et al.*, 1990), direct the DnaA proteins to the melted region, where they play a role in pre-primer formation on the exposed single strands. The reason for the methylation dependence of the 7 bp sequences is unclear. It has been suggested that this region may be the recognition site for an additional factor which stabilises the RepA-DNA complex before and during strand melting (Brendler *et al.*, 1991).

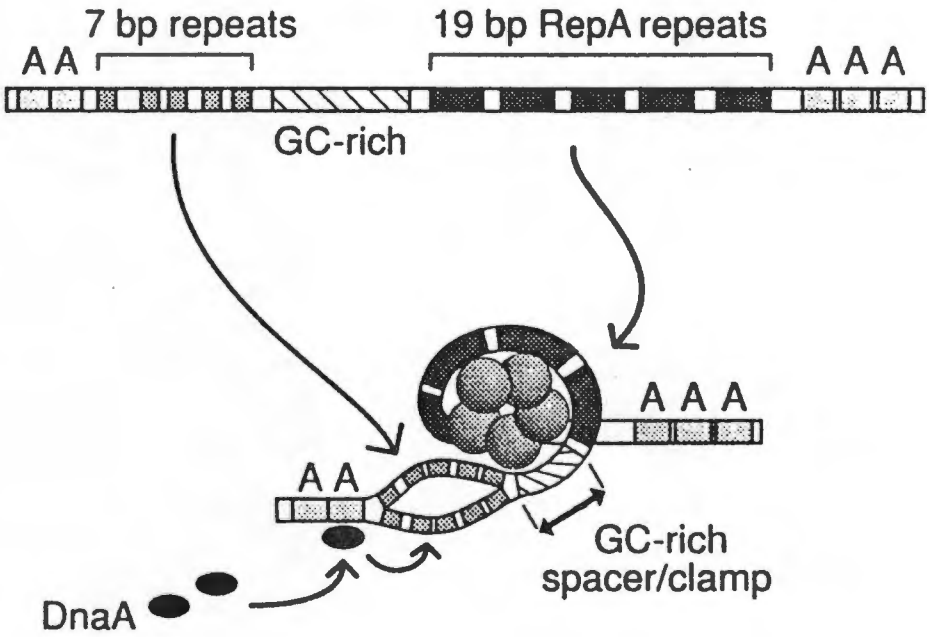


Figure 1.8. Model for assembly of the replication complex at the P1 origin. This model is an adaptation of the general model of Bramhill & Kornberg (1988). DnaA boxes (A), the 7-bp repeats, the GC-rich spacer and the 19-bp RepA-binding repeats can be identified as boxes of different sizes and shades. Binding of P1 RepA initiator protein (shaded spheres) to the 19-bp repeats results in DNA bending, by wrapping of the double strand around a protein core. Specific contact between the wrapped complex and the 7-bp repeat region results in strand melting. The size of the GC-rich spacer is critical for the correct positioning of the two repeat regions. The stability of the GC-rich spacer duplex ensures strand opening away from the RepA complex. The DnaA boxes remain outside of the complex to capture DnaA protein (shaded spheroids) essential to the melted region. DnaA may be important in preprimer formation on the melted single strands (Brendler *et al.*, 1991).

Models for control of mini-P1 replication

Model 1. ***The initiator-titration model.*** It has been proposed that the binding of the initiator (RepA) to the origin repeats is essential for replication. Yet, RepA can also bind to the repeats of the copy number control locus (*incA*), reducing the availability of the initiator. Since the concentration of RepA does not exceed 20 dimers per mini-P1 replicon, the amount of free initiator may be rate limiting for replication (Swack *et al.*, 1987). Therefore, the low copy number of the plasmids which contain an intact *incA* locus may result from the sequestration of RepA at the *incA* site, away from the *ori* (Chattoraj *et al.*, 1984). A similar mechanism has been proposed for F (Tsutsui *et al.*, 1983, see above). Control of replication by sequestration is inconsistent with the autoregulation of *repA* expression since the loss of RepA protein to the *incA* locus should be compensated by new synthesis of RepA (Chattoraj *et al.*, 1985a and 1985b). Furthermore, as it had been demonstrated with mini-F (Tsutsui *et al.*, 1983), addition of extra *incA* repeats did not increase the RepA rate of synthesis, instead it destabilised the mini-P1 plasmid (Pal *et al.*, 1986). This suggests that the autoregulated promoter is not sensitive to the sequestration of RepA at the *incA* repeats. Thus this model can not explain the control of replication of mini-P1 and alternative models have been proposed.

Model 2. ***The two-stage molecular model.*** The model of Trawick & Kline (1985) that was originally developed for the F plasmid, was tested for mini-P1 (Chattoraj *et al.*, 1988). This is represented in *Figure 1.9.* In this model, the initiator and the repressor are different molecular species, the initiator being irreversibly produced from the repressor. The initiator form, not the repressor, binds *incA* and *incC*. Thus, the concentration of the repressor is independent of that of the initiator and the autoregulation-sequestration is resolved. However, Chattoraj *et al.* (1988) were unable to demonstrate two different forms of RepA

and proposed an alternative model whereby DNA looping is responsible for the repressor activity of RepA.

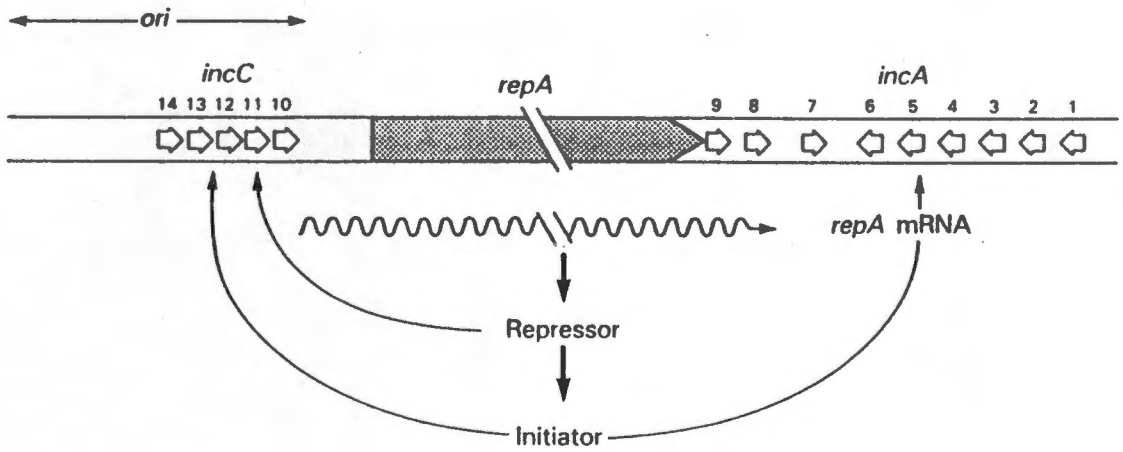


Figure 1.9. Depiction of the two-stage molecular model of Trawick & Kline (1985) for the replication of mini-P1 (Chattoraj *et al.*, 1988). The general organisation of the mini-P1 replicon (***Figure 1.6.***) is shown. The *cis*-acting *ori* (*incC*) includes five 19-bp repeats (arrows 10-14). The *repA* gene (stippled) codes for the initiation protein (RepA), whose promoter maps within repeats number 10-12. The control locus (*incA*), contains nine 19-bp repeats (arrows 1-9). According to the model, the repressor form of RepA binds to the repeats of the *repA* promoter. This results in autorepression of the *repA* gene. The RepA repressor form is irreversibly converted to RepA initiator. Only RepA initiator is competed by the repeats of *incC* and *incA*.

Model 3. ***Replication control by DNA looping.*** The model of Chattoraj *et al.* (1988) proposes that the autoregulation-sequestration paradox can be explained by a DNA looping mechanism (*Figure 1.10*). This model was developed following the observation that both the *incA* and *incC* repeats are able to compete with the *repA* promoter for repressor binding. RepA dimers that bind to the *incA* locus repeats also repress the *repA* promoter by simultaneously binding to the origin repeats. This would result in looping of the intervening DNA. RepA binding would exclude RNA polymerase by steric hindrance from the origin. No suggestions have been made as to the nature of the binding of the RepA dimmer to the two sites. DNA looping has been demonstrated in both prokaryotes and eukaryotes (Ptashne, 1986, Martin *et al.*, 1986; Griffith *et al.*, 1986; Milman & Hwang, 1987; Krämer *et al.*, 1987). In addition, looping has also been implicated in the control of other replication systems (Fuller *et al,* 1984; Mukherjee *et al.*, 1985 and 1988).

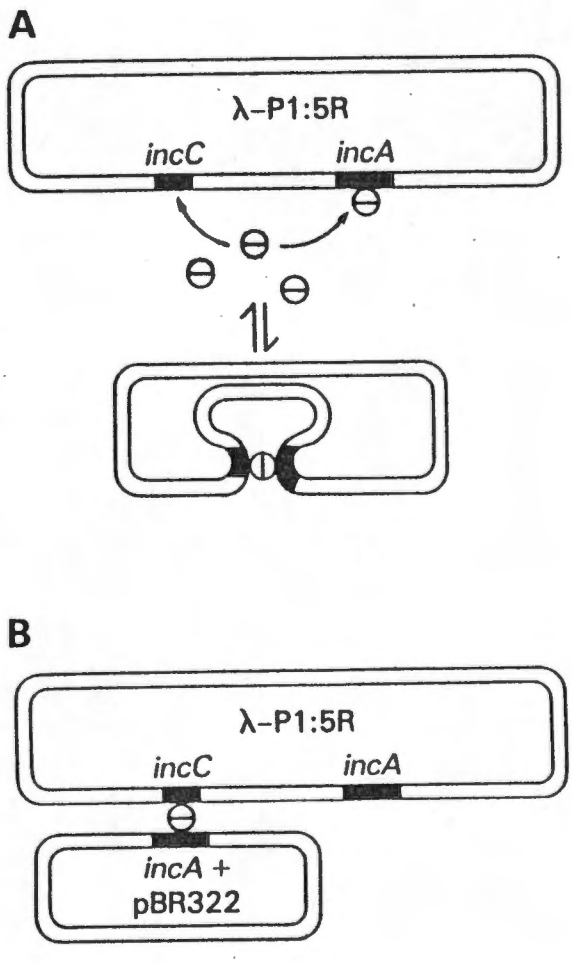


Figure 1.10. Replication control in P1 by DNA looping. The RepA protein dimer is represented by small circles. RepA binds simultaneously to the *incA* and *incC* repeats. When the *incA* and *incC* sites are *in cis* the intervening DNA loops out (**A**). When these sites are *in trans*, a dimeric DNA structure (**B**) is formed (Chattoraj *et al.*, 1988).

Model 4. ***Control of replication by steric hindrance.*** This model is essentially an extended version of the previous model. It was proposed by Pal & Chattoraj (1988) as a result of electron microscopic observations and analysis of the effect of *incA* on mini-P1 copy number. This model is shown in *Figure 1.11.*

To test whether the rate of replication was determined by the initiator concentration the authors constructed 2 plasmids. In one of these plasmids, the *repA* gene was cloned under the control of an inducible *lac* promoter inserted in pBR322. This construct allowed variation of RepA concentration within the host. Another construct contained the mini-P1 origin, but lacked the *repA* gene, so that RepA could be supplied *in trans*. When both plasmids were co-resident and the RepA concentration was raised four-fold beyond the autoregulated level, the copy number of the mini-P1 increased eight-fold. However, if the mini-P1 origin and *incA* were present *in cis*, the effect on copy number was slight, with less than a 2 fold increase. These results suggested that RepA alone can affect the rate of replication but that the copy number control could not be due exclusively to the autoregulation and sequestration of RepA.

The model proposes that the *incA* locus has two functions, one sequestrates RepA from the *ori* and the other is to directly control initiation by steric hindrance of the origin activity. By electron microscopy it was shown that the repeated sequences within the *incA* and origin loci could interact with each other in the presence of purified RepA (Chattoraj *et al.*, 1988). When the two sites are *in cis*, the interposed sequence loops. RepA is also able to pair these sites *in trans*. Either mechanism would prevent initiation of replication. This model is consistent with the failure of excess RepA to initiate replication due to steric hindrance by *incA* and with preferential *cis* action of *incA* (Pal *et al.*, 1986).

Control of replication by steric hindrance (or direct interference) occurs preferentially in the presence of a complete *incA* locus, though Pal & Chattoraj (1988) showed that a single

incA repeat was able to induce loop formation. It is likely then that a single site is also able to interfere with the origin function by DNA looping.

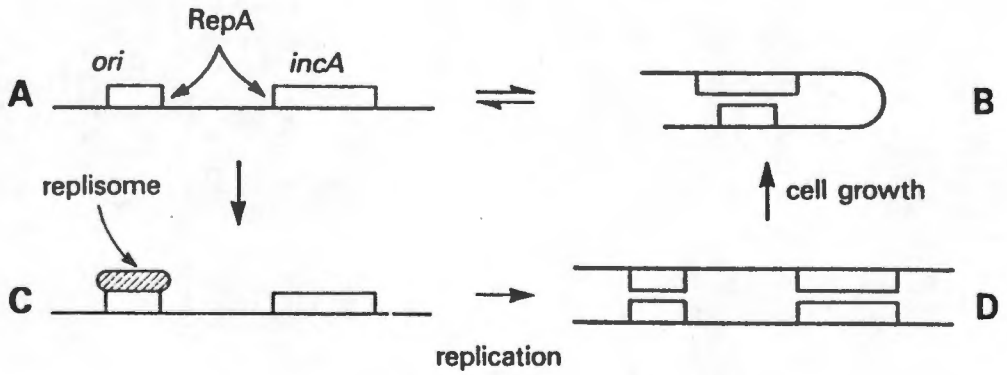


Figure 1.11. Control of replication by steric hindrance at high RepA concentration. Pal & Chattoraj (1988) propose that most RepA-binding sites are occupied when RepA dimers are not limiting (saturated sites represented by open rectangles). At low DNA concentration, species **A** and **B** are in equilibrium. Replisome assembly at *ori* (**C**) blocks pairing *in cis*. After replication, the concentration of DNA increases and pairing *in trans* is favoured (**D**). As the cell grows, the DNA concentration decreases and species **A** and **B** are favoured. At low RepA concentrations, the rate-limiting step is the filling up of the binding sites (**A**). While RepA is limiting, species **D** is least likely to occur.

Model 5. ***Antiparallel plasmid-plasmid pairing replication control.*** The previous two models proposed that origin regulation was achieved either by regulation of RepA synthesis by occlusion of the *repA* promoter (Pal & Chattoraj, 1988; Chattoraj *et al.*, 1988) or by direct interference with the origin function (steric hindrance) (Pal & Chattoraj, 1988). These models favoured *cis* contacts between the origin and *incA*, though the possibility of replication control by *trans* contacts was not excluded. Abeles & Austin (1991) decided to investigate whether or not origin interference by *trans* contacts is as effective as *in cis*. Their results led to the development of the antiparallel plasmid-plasmid pairing model represented in *Figure 1.12.*

Their system was designed to exclude control of replication by autoregulation of RepA synthesis or titration by providing excess RepA *in trans*. Moreover, to eliminate any *cis* contacts, the target sequence (the origin) and the effector sequence (*incA*), were inserted in different plasmids. They showed that when excess RepA was added to an equimolar ratio of these plasmids, no replication occurred. This demonstrated that DNA contacts were able to block the origin function. But since in the wild-type plasmid, the origin and *incA* are *in cis*, their contacts could take place either by looping (*cis*) or by "DNA handcuffing" (*trans* contacts). Abeles & Austin (1991) suggest that DNA handcuffing is favoured. DNA Handcuffing was independently proposed for the control of replication of R6K (McEachern *et al.*, 1989) and for RK2 (Kittell & Helinski, 1991). How would the DNA handcuffing model by applied to the control of replication of P1?

In the wild-type plasmid, the *ori* and the *incA* loci are *in cis*. Following replication these sites are *in trans* in the two daughter molecules (*Figure 1.12.A & B*). Soon after replication the daughter plasmids pair in an antiparallel configuration so that direct contact between *incA* of one daughter plasmid would be in contact with the *ori* of the other (*Figure 1.12.C*). The data of Abeles & Austin (1991) demonstrated that these contacts are formed with little interference from looping. Handcuffing of the two molecules results in blocking of

replication by origin occlusion. Initiation of another round of replication can only occur after plasmid-partition to individual daughter cells (*Figure 1.12.D*).

This model correctly predicts that when a single plasmid is present in a cell, the probability of replication is maximal. But after one round of replication the two plasmids couple and block each other at the *ori* until partition resets the cycle.

Therefore the control of RepA synthesis is not crucial as replication control is achieved by direct inhibition of the origin by plasmid-plasmid contact. Likewise, the inhibition by *incA* is insensitive to the RepA concentration, since the DNA is the inhibitor molecule.

The model accurately mirrors the tight replication control observed *in vivo*. As stated by Abeles & Austin (1991), "it predicts a machine-like, 1-to-2-to-1 cycling of the copy number with each cell generation". In addition, this model describes a powerful negative feed-back loop. If a second round of replication occurred before the plasmids coupled, 3 plasmids would occur at cell division. Two of those plasmids would pair into one daughter cell and would not replicate until the next cell division. Thus, within two generations the "error" would be rectified. This model stresses the importance of partitioning in the control of plasmid copy number.

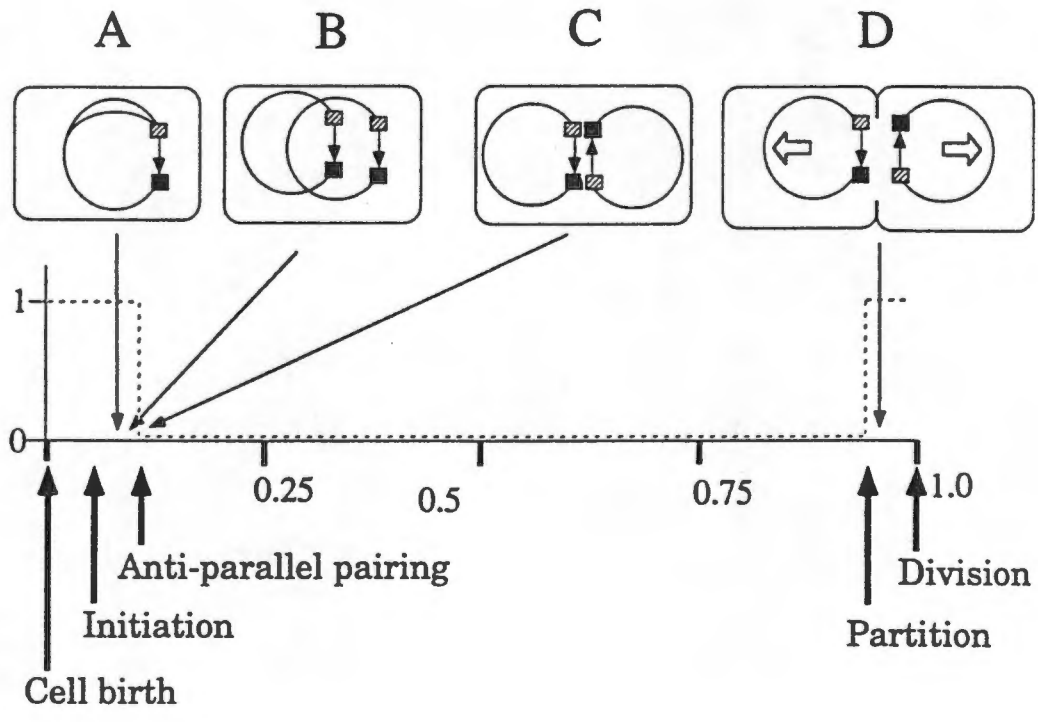


Figure 1.12. The antiparallel plasmid-plasmid pairing model for P1 replication (Abeles & Austin, 1991). The single copy plasmid replicates (**A**), giving rise to two copies of self (**B**). The newly synthesised plasmids pair in an antiparallel configuration (**C**). This event blocks origin function until partition (**D**). The probability of replication initiation (ordinate) is plotted against fractional divisions of the cell cycle (abscissa). Since replication is self-inhibiting, initiation can only occur once in each cell cycle. The *incA* locus and the origin are represented by solid and hatched boxes, respectively.

Replication based incompatibility

As is evident from the replication models proposed, the *incA* locus is involved in plasmid incompatibility (Abeles *et al.*, 1984). By taking the last model into consideration, however, competition will arise by direct binding of the *ori* to *incA*. The previous models suggest that incompatibility arises from RepA binding to the iterons (Abeles *et al.*, 1984), although incompatibility results could not be clearly reconciled with the autorepressor function of RepA. The antiparallel plasmid-plasmid model illustrates why addition of excess cloned iterons do not cause a detectable increase in the RepA concentration, as expected (Chattoraj *et al.*, 1985b).

REPLICATION OF BROAD HOST-RANGE PLASMIDS

IncP PLASMIDS

The IncP plasmids can be divided into 3 major subgroups: IncPα, IncPß and atypical IncP plasmids. Subgroups IncPα and IncPß have a single replicon (Chikami *et al.*, 1985). Representatives of IncPα include RK2, R702, R839, R935, R938, R1033 and pUZ8, whereas R751, R772, R906 and pJP4 belong to IncPß (Villarroel *et al.*, 1983; Yakobson & Guiney, 1983; Smith & Thomas, 1985; Chikami *et al.*, 1985; Smith & Thomas, 1987). The third subgroup includes plasmids which do not fit into either the IncPα or ß subgroups and have been designated "atypical IncP plasmids". These plasmids have a restricted host-range, exhibit asymmetric incompatibility and contain at least two replication system (Thomas & Helinski, 1989). The plasmids pHH502.1 (Nugent *et al.*, 1982), pAV1 (Hinchliffe & Vivian, 1980), pMU700-pMU707 (Grant *et al.*, 1980) and pTM89 (Monti-Bragadin & Samer, 1975) are examples of the "atypical IncP plasmids". Despite their different functional organisation and host range, there is evidence that all of the IncP subgroups evolved from a common ancestor (Smith & Thomas, 1987).

The best studied IncP plasmid, RK2, is identical to RP4, RP1, R18 and R68 (Burkardt *et al.*, 1979; Currier & Morgan, 1981; Stokes *et al.*, 1981; Villarroel *et al.*, 1983; Smith & Thomas, 1989). Its replicon is discussed here.

RK2 PLASMID

The plasmid RK2 is a 60 Kb representative of the IncPα plasmids (Lanka *et al.*, 1983; Pansegrau & Lanka, 1987) and has a copy number of 4 to 7 copies per chromosome equivalent in *E. coli* (Figurski *et al.*, 1979; Grinter, 1984) and of 3 in *Ps. aeruginosa* (Itoh *et al.*, 1984). The overall organisation of RK2 is far more complex than that of narrow host-range plasmids. There are large segments coding for replication functions interspersed with plasmid stability functions, namely the *kil* genes, the multimer resolution system (*tnpR*) of Tn*1* and the *par* locus, as well as of resistance genes (Tc^r, Te^r, Ap^r) (*Figure 1.13.*) (Bradley & Taylor, 1987; Pansegrau & Lanka, 1987; Thomas & Helinski, 1989).

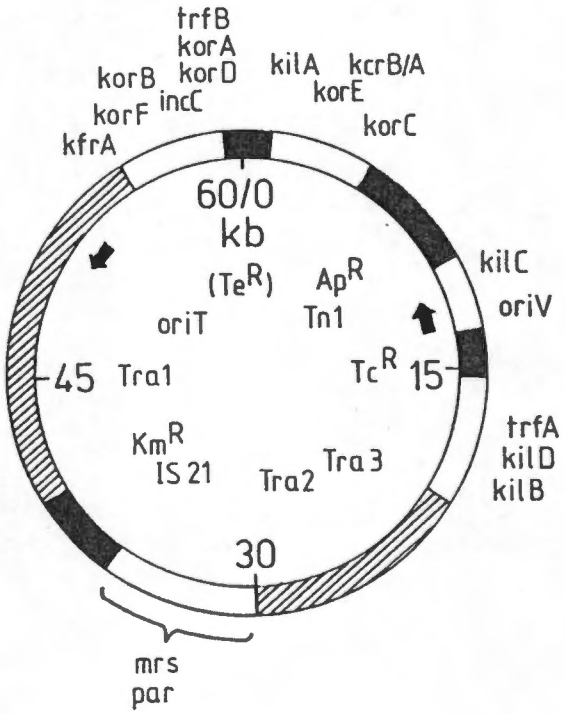


Figure 1.13. Functional organisation of RK2. The loci shown are described in the text. Regions involved in stable inheritance are shown as open segments, whereas those encoding conjugation functions are indicated by hatched sections. Resistance determinants and transposable elements are shown as solid segments (Thomas & Helinski, 1989)

Host proteins involved in RK2 replication

Replication of RK2 in an *E. coli*-derived, *in vitro* system, requires DnaA, DnaB helicase, DnaG primase and DNA Pol III (Pinkney *et al.*, 1988). In addition RK2 appears to have different requirements for DnaA in various hosts. Replication in *E. coli* and *Ps. putida*, but not in *Ps. aeruginosa* is dependent on the presence of DnaA boxes. The correct positioning and spacing of the DnaA boxes within the *oriV* is also important (Gaylo *et al.*, 1987; Cross *et al.*, 1986; Krishnapillai, 1986). This suggests that there may be different specificities of DnaA proteins in different hosts (Ogasawara *et al.*, 1985; Skovgaard & Hansen, 1987). The presence of an IHF site within the *oriV* suggests that this factor may be required for replication, but this remains to be confirmed.

The minimal replication system

The minimal essential regions for replication of RK2 include the *oriV* and its activating locus, the *trfA* gene (trans-activating replication function A) (Figurski & Helinski, 1979; Shingler & Thomas, 1984; Thomas, 1981). The *trfA* locus is regulated by the *trfB* operon (Schreiner *et al.*, 1985; Theophilus *et al.*, 1985; Thomas & Hussain, 1984; Young *et al.*, 1985). The latter is, however, not essential for RK2 replication in *E. coli, Ps. aeruginosa, Ps. putida, Azotobacter vinelandii, Alcaligenes eutrophus* (Schmidhauser *et al.*, 1983; Schmidhauser & Helinski, 1985) though it may contribute to stable plasmid maintenance in these hosts (Schmidthauser & Helinski, 1985; Durland & Helinski, 1987; Thomas & Helinski, 1989).

Unlike the basic replicons of the narrow host range-plasmids which are contiguous within a small segment of DNA (up to 3 Kb), the *oriV, trfA* and *trfB* loci are spread over a 20 Kb region of RK2 (Barth, 1979; Thomas *et al.*, 1980; Thomas *et al.*, 1979) (*Figure 1.14.*).

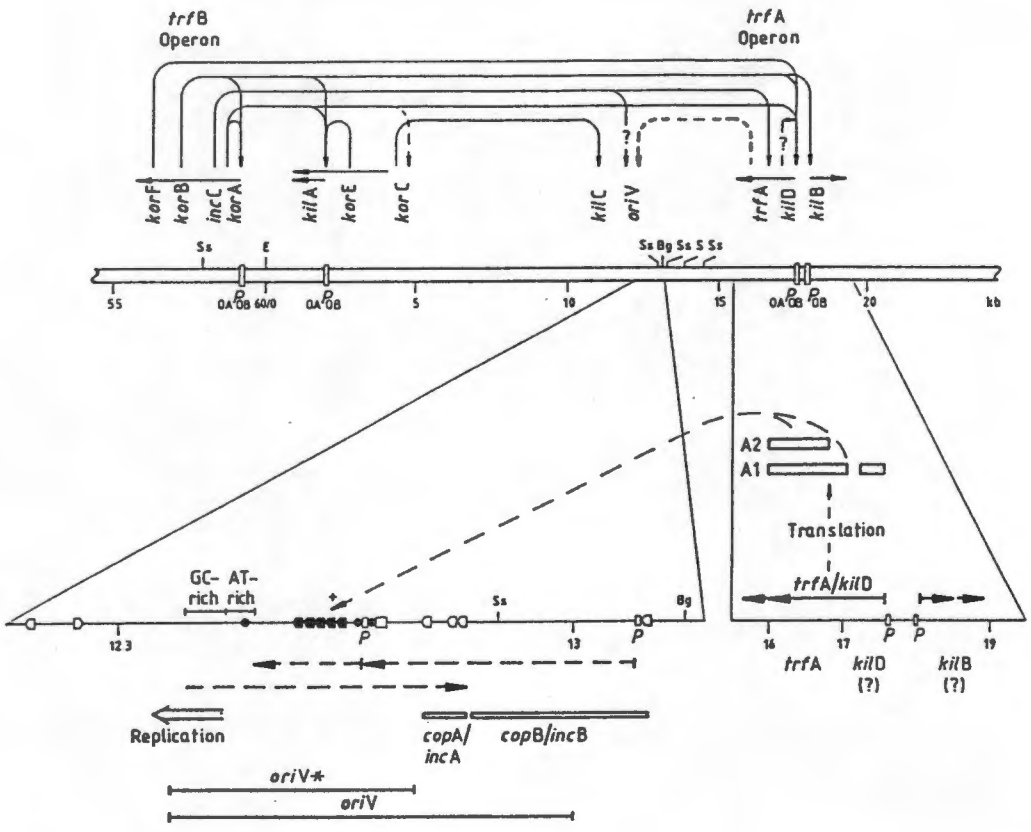


Figure 1.14. Regions of RK2 involved in replication initiation and regulation (Kües & Stahl, 1989). **A.** Organisation of the *trfA* and *trfB* operons (Pansegrau & Lanka, 1987), whose regulation is described in the text. $_{OA}P_{OB}$ indicate the promoters of the KorA and KorB operators. Horizontal arrows below operons indicate the direction of transcription. Interaction between loci is represented by curve arrows (dashed, positive effect; solid, negative). **B.** Some restriction enzyme sites (Bg, *Bgl*II; E, *Eco*RI; S, *Sal*I; Ss, *Sst*II) are shown. Kilobase coordinates of RK2 are also shown. **C.** Molecular organisation of the *ori*V region as well as of the *trfA* operon. The arrows (◀ , ◁) indicate eight 17-bp direct and inverted repeats. These are binding sites for TrfA (Pinkney *et al.*, 1988). In addition, DnaA boxes (●), putative IHF binding site (□), putative promoters (*P*), *ori*V* in *E. coli*, *ori*V in most bacteria, direction of replication (⇐), copy number control regions (*copA/incA*; *copB/incB*), GC-rich and AT-rich sequences are shown. Putative transcripts within the *ori*V are marked by dashed arrows.

The structure of the *ori*V

The molecular organisation of the RK2 *oriV* is similar to that of many narrow host-rage plasmids and is depicted in *Figure 1.14.* (Filutowicz *et al.*, 1987; Linder *et al.*, 1985; Scott, 1984; Stalker *et al.*, 1981; Thomas *et al.*, 1984). The 617 bp *oriV* includes the following features: **1)** eight, 17 bp iterons in two clusters (three and five); **2)** a putative promoter surrounded by DnaA boxes and a single putative IHF-binding site; **3)** a 49 bp AT-rich sequence (74% A+T) containing an additional DnaA box; and **4)** a 67 bp GC-rich sequence (79% G+C) (Gaylo *et al.*, 1987; Smith & Thomas, 1985; Stalker *et al.*, 1981). An additional 17 bp iteron occurs upstream from the *oriV* (Thomas *et al.*, 1984). This iteron overlaps the promoter for two putative transcripts. Downstream from the *ori*V, 2 degenerate repeats are present in an inverted orientation (Thomas & Smith, 1987). Three open reading frames are found within the *oriV* region, two of which are associated with copy number control and incompatibility (*copA/incA* and *copB/incB, Figure 1.14.*) (Thomas *et al.*, 1984).

Host specificity of *ori*V elements

Although the same origin fragment is used in most bacteria, in some hosts certain *oriV* elements are redundant. Replication in *E. coli,* requires only a 393 bp *Hpa*II fragment (*oriV**, *Figure 1.14*), which includes the DnaA boxes, five of the iterons as well as the AT- and GC-rich region (Stalker *et al.*, 1981; Thomas *et al.*, 1981). In the *Pseudomonas* species, however, the entire 617 bp *oriV* fragment is necessary (Schmidthauser *et al.*, 1983). Interestingly, disruption of the single iteron outside the *oriV* prevents replication from occurring in *Ps. stutzeri*, but not in *E. coli, Ps. aeruginosa* and *Ps. putida* (Krishnapillai, 1986; Nash & Krishnapillai, 1987). This suggests that the promoter overlapping this iteron may be required for *Ps. stutzeri* replication.

The *trfA* operon

The *trfA* operon includes two cistrons. The first, *kilD,* encodes the regulator KilD and the second, *trfA*, encodes two proteins of 382 amino acids (43 KDa, TrfA-43) and 285 amino acids (32 KDa, TrfA-32) respectively. The TrfA proteins are translated from alternative translation start sites within the same open reading frame (Shingler & Thomas, 1984; Smith & Thomas, 1984) and exhibit specific binding to the *oriV* at low concentrations (Pinkney *et al.*, 1988). The protein TrfA-32 is essential for RK2 replication in *E. coli* (Kornacki *et al.*, 1984; Shingler & Thomas, 1984), as well as in all other hosts tested (Kornacki *et al.*, 1984; Schmidhauser *et al*, 1983; Schmidhauser & Helinski, 1985; Shingler & Thomas, 1984). The role of TrfA-43 is unclear, though it may be required to support replication in *P. aeruginosa* (Durland & Helinski, 1987; Thomas & Helinski, 1989).

Regulation of replication: Expression and regulation of the *trfA* operon

The control of replication initiation of RK2 remains poorly understood (Durland *et al.*, 1990). Replication is initiated by the binding of TrfA to the iterons of the *ori*V and proceeds unidirectionally to the left (Figurski & Helinski, 1979; Stalker *et al.*, 1981; Thomas, 1981; Thomas *et al.*, 1980; Thomas *et al.*, 1981; Meyer & Helinski, 1977). This implies that the control of *trfA* expression is vital to replication initiation.

The *trfA* gene is transcribed from a strong promoter that was identified by polymerase footprinting and reverse transcriptase mapping of the 5' end of the mRNA transcript (Pinkney *et al.*, 1987; Smith *et al.*, 1984). The 5' end of the transcript is identical in *E. coli. Ps. aeruginosa* and *Ps. putida*, showing its promiscuous character (Pinkney *et al.*, 1988). Transcription of this gene is under control of a complex regulatory network of operons (Thomas *et al.*, 1984; Schreiner *et al.*, 1985; Kornacki *et al.*, 1984). These are the *kor* (*kil*-override) genes, *korA/korD, korB*, *korC, korE* and *korF*, that prevent host killing

by the host-lethal *kil* genes, *kilA*, *kilB*1/*kilD*, *kilB* and *kilC*, and stabilise the plasmid (Bechhofer & Figurski, 1983; Young *et al.*, 1987). These genes form part of a regulatory system which enables RK2 to replicate in a wide range of hosts (Schreiner *et al.*, 1985).

The *trfA* expression is negatively regulated by the products of the *trfB* operon (*korA*/*korD*/*trfB*, *korB*, *korF* and *incC*) (*Figure 1.14.*) (Thomas, 1986; Thomas & Smith, 1986; Bechhofer *et al.* 1986; Kornacki *et al.*, 1987; Theophilus & Thomas, 1987) (*Figure 1.14*). The products of *korA* and *korB* repress transcription from the *trfA* operon promoter (P*trfA*) by binding to a pair of inverted repeats overlapping P*trfA* (Shingler & Thomas, 1984; Smith *et al.*, 1984; Schreiner *et al.*, 1985; Theophilus *et al.*, 1985; Young *et al.*, 1985; Bechhofer *et al.*, 1986). In addition, *korA* may modulate the expression of TrfA post-transcriptionally (Durland & Helinski, 1990). The *kilD* determinant, associated with the first open reading frame of the *trfA* operon appears to modulate the activity of *korA* and *korB* to prevent over-repression of P*trfA* (Schreiner *et al.*, 1985).

KorA also has an inhibitory effect on the expression of the *kilA* locus as well as autorepression of the *trfB* operon (Theophilus *et al.*, 1985; Young *et al.*, 1985). The product of the *korC* gene (KorC) is an 88 amino acid protein (9.15 KDa) with an α-helix-turn-α-helix motif which together with KorA, acts as a repressor of the *kilC* and *kilE* genes (Kornacki *et al.*, 1990).

KorB is 39 KDa protein (Kornacki *et al.*, 1987; Theophilus & Thomas, 1987) that represses the *trfA* operon (Schreiner *et al.*, 1985; Shingler & Thomas, 1984; Thomas & Hussain, 1984). KorB also represses the *trfB* operon (autorepression) and the *kilB* locus (Figurski *et al.*, 1982; Shingler & Thomas, 1984; Smith & Thomas, 1989). In the presence of KorA, KorB acts as a repressor of *kilA* together with KorE (Young *et al.*, 1987). KorB can be substituted by KorE as corepressor of the *kilA* gene (Young *et al.*, 1987). KorF has an inhibitory effect on the expression of the *trfA* and *trfB* operons, that is similar to that of KorB (Thomas, 1988).

The net result of the complex regulation of this replication system is the maintenance of the TrfA protein(s) at rate-limiting concentration (Durland & Helinski, 1990). Thus, the *trfA* and *trfB* operons together with the *kilA* operon form a co-regulated network which permits replication in a wide range of unrelated hosts. This property is characteristic of promiscuous plasmids. Rearrangements and deletions of some of these operons lead to a reduction of host-range or plasmid instability (Barth *et al.*, 1984; Schmidhauser & Helinski, 1985; Schreiner *et al.*, 1985; Theophilus *et al.*, 1985; Thomas, 1983; Thomas *et al.*, 1982). The atypical IncP plasmids may be examples of such deleted IncP plasmids (Smith & Thomas, 1989).

The evolutionary development of *kil/kor* system is uncertain. It has been suggested the *kil/kor* regulation of *trfA* might have evolved to modulate promiscuous plasmid replication following entry into a new host (Shingler & Thomas, 1984). This system could lead to a initial round of replication while arresting the host cellular division until the correct plasmid copy number is attained. It has also been suggested that it could have evolved to enhance the plasmids capacity to synthesise the correct levels of TrfA proteins for replication and stable maintenance in genetically unrelated hosts (Schreiner *et al.*, 1985). Either possibility would be to the advantage of the promiscuous plasmid (Krishnapillai, 1988).

Copy number control in RK2

It was originally proposed that the copy number of RK2 was determined primarily by regulating the synthesis the TrfA protein (Pohlman & Figurski, 1983; Schreiner *et al.*, 1985; Thomas & Hussain, 1984). Durland & Helinski (1990) showed, however, that the mechanism controlling the replication of RK2 is independent of TrfA concentration, since raising the TrfA levels two or three fold above the normal physiological concentrations leads to a 30% increase in copy number. Furthermore, 170-fold increase in the concentration of TrfA had no effect on copy number, showing that the concentration of TrfA is not the only factor determining the replication frequency of RK2 (Durland & Helinski, 1990). The TrfA protein(s) must be involved in copy number control, since copy-up mutants with base changes mapping within the TrfA gene have been isolated (Durland *et al.*, 1990). This information, together with the demonstration that the copy number of derivatives of RK2 consisting of only the origin and *trfA* gene under the control of a constitutive promoter, was similar to that of wild-type, suggested that the copy control function is an intrinsic property of the origin or of the *trfA* gene or of both (Schmidhauser *et al.*, 1983; Schmidhauser & Helinski, 1985).

Copy number control has also been associated with the origin iterons, as deletion of three iterons upstream of the *ori*V raises the plasmid copy number (Thomas *et al.*, 1984).

Incompatibility in RK2

In addition to copy number control, the *oriV* iterons express incompatibility towards intact RK2 origins (Thomas *et al.*, 1981). The magnitude of this incompatibility is related to the number of iterons present, indicating that iterons are important in the regulation of plasmid replication (Thomas *et al.*, 1984; Thomas & Hussain, 1984). This is also true for F (Tolun & Helinski, 1981; Tsutsui *et al.*, 1983), P1 (Abeles *et al.*, 1984; Chattoraj *et al.*, 1984),

pSC101 (Yamaguchi & Yamaguchi, 1984a and 1984b), R1162 (Lin & Meyer, 1986), RSF1010 (Persson & Norsdtröm, 1986), Rts1 (Kamio *et al.*, 1984 and 1988) and R6K (McEachern *et al.*, 1986).

From the work of Kittell & Helinski (1991) a replication control model was suggested which is able to integrate the above findings in a unified model. This is the intermolecular coupling model of replication origins and is discussed below.

Model for control of RK2 replication

The intermolecular coupling model. The work of Kittell & Helinski (1991) confirmed that the iteron region of RK2 is responsible for expression of incompatibility of this replicon. This was demonstrated by cloning the iteron region in pUC19 and providing these *in trans* to RK2 replicons. They also established that at least 2 correctly positioned iterons are necessary to elicit incompatibility. A similar finding has been reported for R1162 (Lin *et al.*, 1987). Raising the concentration of the RK2 iterons leads to further inhibition of replication. Iteron inhibition is specific for RK2 replicons, suggesting that no host factors are involved. This inhibition is also not reversed by raising the levels of TrfA to saturate all the existing replicons. These findings are inconsistent with the titration models developed for other iteron containing plasmids (Tsutsui *et al.*, 1983; Chattoraj *et al.*, 1984; Trawick & Kline, 1985). These results can be explained by a model that proposes that inhibition of replication results from the intermolecular coupling of the TrfA bound iterons from two different plasmids. Replication cannot take place because the initiation complex is unable to assemble (Kittell & Helinski, 1991).

This model predicts that the concentration of TrfA determines whether the TrfA-bound origin interacts with other host initiation factors or with another TrfA-bound origin to regulate copy number. Since the model allows for *in cis* and *in trans* intermolecular

coupling of the iterons, this may also explain why deletion of some iterons leads to a raised copy number of the plasmid.

RK2 is the first broad-host-range iteron containing plasmid for which an intermolecular replication inhibition mechanism of copy number control has been proposed. Similar models have been proposed, namely "hand-cuffing" for R6K (McEachern *et al.*, 1989) and antiparallel plasmid-plasmid coupling for P1 (Abeles & Austin, 1991).

IncN PLASMIDS

All IncN plasmids identified to date are conjugative and promiscuous, though not as promiscuous as those of the IncP group, such as RK2 (Krishnan & Iyer, 1988). IncN plasmids have been isolated from a wide range of bacterial species, including *Aeromonas* spp., *Enterobacter* spp., *E. coli* spp., *Klebsiella aerogenes, Proteus* spp., *Providencia* spp., *Salmonella* spp., *Shigella* spp. and *Vibrio cholera* (Aoki *et al.*, 1977; Jacob *et al.*, 1977; Arai *et al.*, 1980). The overall organisation of the IncN plasmids N3, pCU1, pKm101, R46 and R15 is shown in *Figure 1.15*. There are four major domains in these plasmids (see Iyer, 1989): **1)** a contiguous region containing the essential replication features, *rep*; **2)** accessory systems which are not essential for replication, but contribute to plasmid stability and maintenance. These include the killing of *K. pneumoniae* (*kik*), the killing of *E. coli* (*kil*), the inhibitor of *kil* (*kor*), the fertility inhibition of P-group plasmids (*fip*), the stability of plasmid locus (*stb*), nuclease (*nuc)* and the mutability by UV or by chemicals (*muc).* **3)** the region containing the conjugative apparatus (*tra*). **4)** insertion sequences (IS46) and antibiotic resistance genes.

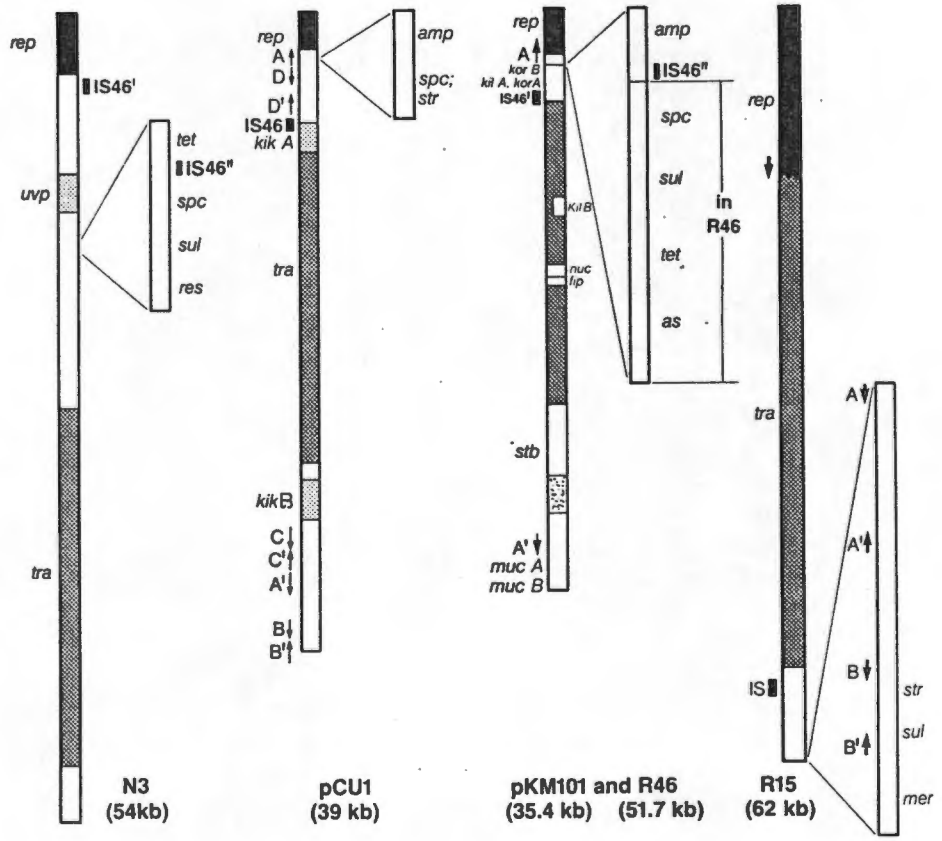


Figure 1.15. Comparison of the structural and functional organisation of four IncN group plasmids. The circular maps were converted to linear representation for clarity. The solid arrows (A,A', B,B', C,C', D,D') indicate regions of extensive sequence similarity. Abbreviations *amp, as, mer, spc, sul, str*, indicate resistance to ampicillin, arsenate, mercuric chloride, spectinomycin, sulfonamide and streptomycin, respectively (Iyer, 1989). Other features are described in the text.

pCU1 PLASMID

The 39 Kb, *polA1* independent, pCU1 plasmid was identified in a hospital isolate of *Salmonella typhimurium* (Konarska-Kozlowska & Iyer, 1981a). Its physical properties, functional organisation (Konarska-Kozlowska & Iyer, 1981 and 1983; Konarska-Kozlowska *et al.*, 1983; Thatte *et al.*, 1985), conjugative properties (Konarska-Kozlowska & Iyer, 1981; Thatte *et al.*, 1985), host-range (Krishnan & Iyer, 1988) and replicon organisation (Kozlowski *et al.*, 1987; Krishnan & Iyer, 1990; Krishnan *et al.*, 1990) have been characterised.

Recently, it has been shown that although the *tra* system of pCU1 exhibits similar broad-host-range character to that of RK2, the plasmid could not be maintained in *Acinetobacter calcoaceticus, Proteus mirabilis, Ps. cepacia, Ps. stutzeri* and *Rhizobium leguminosarum* (Krishnan & Iyer, 1988). This suggested that pCU1 exhibits conjugational but not replicative promiscuity and its host-range is thus limited to the hosts in which it can replicate. Indeed, the molecular organisation of the pCU1 replicon differs from that of the typical promiscuous plasmids in that all the essential replication and maintenance functions are contained within a 2 Kb region (Krishnan & Iyer, 1990). The molecular organisation of this region is described below.

Host functions required for replication

Although no *in vitro* replication assays have been carried out, it is inferred that pCU1 requires the host genes *dnaA* and *dam* for replication (Krishnan & Iyer, 1990).

Molecular organisation of the mini-replicon of pCU1

The DNA segment containing the minimal region essential for replication of pCU1 (*Figure 1.16)* was identified after deletion analysis (Konarska-Kozlowska & Iyer, 1981b) and mapping of the transposon-like structures of the plasmid. The mini-replicon is contained within a 2 Kb *Pvu*II fragment and includes 3 domains (Kozlowski *et al.*, 1987; Krishnan & Iyer, 1990; Krishnan *et al.*, 1990). The first is a 368 bp *cis*-acting origin of replication (*ori*V), containing 4 GATC methylation sites and one DnaA box. The second domain is the *rep* that is contained within a 1 Kb region. This domain includes two open reading frames (ORF239 and ORF46), whose expression products are essential for replication (27 and 5.5 KDa Rep polypeptides, respectively) (Krishnan *et al.*, 1990). The 3' end of ORF239 overlaps the *ori*. In addition, the product of ORF239 is responsible for the *polA1*-independence of the replicon. Two 30 bp iterons (Group II iterons) are present immediately upstream from ORF46. The third domain consists of thirteen 37 bp tandem iterons (Group I iterons), which are involved in expression of incompatibility and copy number control. Group I iterons are situated downstream from ORF239. Several other open reading frames have been identified on sequence analysis of the replicon, but no corresponding expression products have been demonstrated (Krishnan *et al.*, 1990).

The most remarkable feature of the basic replicon of pCU1 is its compact nature. For instance the overlap of the *ori* and *rep* functions. At present the significance of this overlap is unknown. The compact molecular organisation of the pCU1 replicon resembles more closely that of the F and P1 basic replicons (*Figures 1.2 & 1.6*). Despite this, pCU1 has a broader host-range than either of these plasmids, although it is narrower than the host-range of RK2.

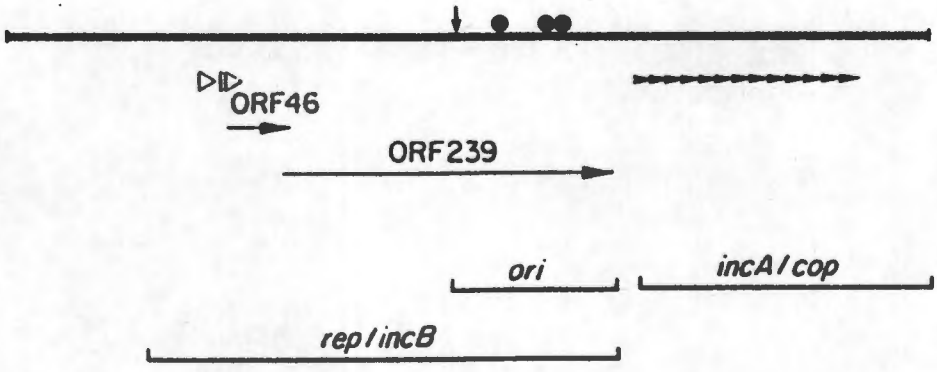


Figure 1.16. Functional organisation of mini-pCU1. The *cis*-acting region (*ori*) includes an essential DnaA box (arrow) and 3 GATC sequences (filled circles). *rep* indicates the minimal autonomously replicating region in a (PolA$^+$) *E. coli*. *incA* and *incB* designate incompatibility determinants. *cop* corresponds to the *trans*-acting copy control region. ORF46 and ORF239, encode essential proteins for replication. Open arrows indicate the location of two 30-bp iteron family (Group II), whereas horizontal filled arrows display the thirteen, 37-bp tandem repeat family (Group I) (Krishnan & Iyer, 1990; Krishnan *et al.*, 1990).

Control of copy number and incompatibility

Deletion of the Group I iterons leads to a 2 two-fold increase in pCU1 copy number. This can be reversed by re-introduction of the Group I iterons *in trans* (Krishnan & Iyer, 1990). Thus, the Group I iterons are a copy number control element. When Group I iterons are cloned in pBR322, they express incompatibility towards pCU1. As few as 6 iterons of Group I suffice to reproduce the wild-type incompatibility and copy number control of pCU1. It is not clear if the strength of the incompatibility reaction and plasmid copy number are proportional to the number of Group I iterons presented in pCU1 (Krishnan & Iyer, 1990).

Model for replication control of pCU1

Because of the similarities in the functional organisation of the pCU1 mini-replicon with those of F and P1, Krishnan & Iyer (1990) suggested that the replication of pCU1 may be regulated by the initiator-titration mechanism as has been proposed for F (Tsutsui *et al.*, 1983). However, it may be premature to assign such replication control models to pCU1, since details regarding the role of the 27 and 5.5 KDa polypeptides are lacking. DNA foot-printing and retardation assays are necessary to identify the target of these proteins. Furthermore, the role of Group I and II iterons needs definition. The physiological concentration of the "Rep" proteins has to be determined, so that molar ratios of "Rep"-DNA complexes can be calculated. Only then can a suitable model for replication control of pCU1 be proposed.

CONCLUDING REMARKS

Plasmid replication in Gram-negative bacteria is dependent on both host- and plasmid-encoded elements. Some plasmids encode elements which can be recognised in a wide range of bacteria enabling correct assembly of plasmid replisome complexes. These are the broad-host range plasmids. In contrast, the narrow host-range plasmids lack this capability (Krishnapillai, 1988; Kües & Stahl, 1989).

Despite the intense structural and functional analysis of basic replicons, the factors involved in control of plasmid replication and host-range are still poorly understood, with the possible exception of F and P1 (Kline, 1988; Abeles & Austin, 1991). The basic replicons of these narrow host-range, low copy number plasmids encode a *rep* gene and thus are *polA1* independent. They are nevertheless dependent on other host-encoded components, such as DnaA, DnaB, DnaC, DnaE and DnaG amongst others (see above), which limit their host range.

The molecular organisation of the basic replicon of the plasmid pCU1 is similar to that of F and P1. pCU1, however, exhibits a broader host-range than that of either of the former two plasmids (Krishnan & Iyer, 1990). Like F and P1, it is also *polA1* independent, with similar requirements for host factors. It is therefore difficult to understand the differences in host range of these plasmids.

The promiscuous replication properties of RK2 are also difficult to reconcile exclusively with replicon organisation. RK2 contains a single, but extended replicon, of which only the *ori*V and *trfA* operon are used in most hosts (Thomas & Helinski, 1989). Indeed, this replicon organisation does not differ extensively from that of narrow-host-range plasmids (Kües and Stahl, 1989). The promoter of the *trfA* gene is structurally different from those of the *rep* genes of F and P1, and exhibits a promiscuous character. This promoter promiscuity was demonstrated by showing that the 5' end of mRNA transcripts is identical

in *E. coli, Ps. aeruginosa* and *Ps. putida* (Pinkney *et al.*, 1987; Smith *et al.*, 1984). Therefore, this promoter must be recognised by host-specific-enzymes in a wide range of bacteria. The presence of a promiscuous promoter together with differential protein requirements in different hosts may determine the extent of plasmid promiscuity (Krishnapillai, 1988).

Kües & Stahl (1989), however, reject that promoter recognition alone, is sufficient for promiscuous replication. Instead they suggested that the promiscuous promoters such as that of the *trfA* gene, can be modulated by host-specific regulators, to ensure the correct levels of gene expression required for replication in each host. This process may be facilitated or stabilised by product(s) of the complex network of genes known to control replication in RK2 (see above). This may be one of the reasons why promiscuous replicons tend to be larger and more complex than narrow-host range ones. In consequence, the promoters of narrow-host-range plasmids are probably as efficient as those of promiscuous plasmids. Promoters of narrow-host-range replicons may however be incorrectly regulated by host-specific factors, as these replicons lack some of the auxiliary stabilising factors that broad-host-range replicons seem to encode. Evidence in favour of this hypothesis was presented by Lodge *et al.* (1990) who demonstrated that promoters tested in *E. coli* and *Ps. aeruginosa* were equally active in both hosts. In addition, dependence of activity on sequence was the same in *E. coli* and *Ps. aeruginosa*.

The moderately promiscuous properties of pCU1 may be due to the compact size of its replicon, although the presence of a promiscuous promoter in its *rep* gene has not been excluded. No evidence for the latter is available, however.

In conclusion, narrow-host range plasmids contain small basic replicons which are responsible for tight replication control in a restricted number of hosts. In contrast, broad-host range plasmids seem to be characterised by larger replicons with promiscuous promoters that enable replication in a variety of unrelated hosts. Plasmids such as pCU1 are

moderately promiscuous and are characterised by small stably maintained replicons. This plasmid was originally isolated from a hospital acquired *S. typhimurium* and its host range seems to be restricted mainly to the *Enterobacteriaceae* of the gamma subdivision (Konarska-Kozlowska & Iyer, 1981a; Krishnan & Iyer, 1990). Since the *Enterobacteriaceae* are commonly found within the hospital environment, it is possible that pCU1 has adapted or specialised to replicate in members of this family. Thus, in these moderately promiscuous plasmids the limitation of a restricted host range may be off-set by replicon stability.

PURPOSE OF THIS STUDY

This study has two major objectives. First, to describe the functional organisation of the **repß**, of the moderately promiscuous plasmid pGSH500. Second, to compare this replicon with those of similar plasmids.

CHAPTER 2

DELINEATION AND FUNCTIONAL CHARACTERISATION OF repß

CHAPTER 2

DELINEATION AND FUNCTIONAL CHARACTERISATION OF repß

SUMMARY

The conjugative, multiple-resistance plasmid, pGSH500, was shown to contain two unique *polA1* independent replicons, repα (pFDT100) and repß (pFDT200). The incompatibility group of pGSH500 and its replicons could not be established and it is likely that pGSH500 belongs to a novel incompatibility group. pGSH500 is maintained at 3-5 copies per chromosome equivalent. In *E. coli*, repß is the dominant replicon of pGSH500. This replicon, whose replication, incompatibility and copy number properties resemble those of pGSH500, was confined to a contiguous 2.2 Kb DNA region.

INTRODUCTION

The basic replicon contains the essential genetic information for plasmid replication and generally consists of 3 domains. The first, is the site of replication initiation (*ori*V). The second, encodes a replication initiator protein (*rep*). The third, controls incompatibility and copy number (*inc/cop*). Thus, to investigate a replicon of a plasmid, it is necessary to functionally define and characterise these domains (Thomas, 1987b).

The identification of a basic replicon involves testing of plasmid deletions to determine their ability to replicate autonomously. Ideally these deletions should be contained in small

fragments which are free of host-lethal genes (Thomas, 1987b). The methods used to construct and characterise these deletions are largely dictated by the properties of the replicon itself.

The first mini-replicons to be isolated were those of the large plasmids F (Lovett & Helinski, 1976) and R6-5 (Cohen & Chang, 1973). To isolate the mini-replicons, the plasmids were digested with the restriction enzyme *Eco*RI. Specific digest fragments were then ligated to other *Eco*RI fragments which lacked autonomously replicating functions, but contained selectable markers such as antibiotic resistance genes. These mini-replicons, which faithfully mimicked the replicative behaviour of the parent plasmid, were then isolated following selection in *E. coli* hosts.

If the replicons of interest are known to be *polA1* independent, a more specific approach can be used. This involves the construction of a library of plasmid DNA in *polA1* dependent vectors, such as pMB1/ColE1 plasmids, followed by selection in *E. coli polA1* mutants (Kingsbury & Helinski, 1970, 1973), such as *E. coli* 3478 (F^- *polA thy*) (de Lucia & Cairns, 1969), *E. coli* W3310 (*polA1*) (Saadi *et al.*, 1987) and *E. coli* C2110 (*polA1, rha, his*) (Thomas, 1987b). Since the pMB1/ColE1 plasmids are unable to replicate in these host strains, survival of the recombinants depends entirely on the *polA1* independent insert (Thomas, 1987b). This approach has been successfully used for replicon characterisation of the plasmids F, R6K, pSC101, RK2, R1, R6-5 and R6 (Lane, 1981; Kontomichalou *et al.*, 1970; Cabello *et al.*, 1976; Thomas *et al.*, 1980; Kolleck *et al.*, 1978; Timmis *et al.*, 1978).

Temperature sensitive *polA1 E. coli* mutants (Grindley & Kelley, 1976) have been useful to distinguish plasmid establishment and maintenance functions as well as for the determination of plasmid copy number. Some of these strains have been used in the characterisation of plasmids belonging to the F incompatibility group (Saadi *et al.*, 1987;

Maas *et al.*, 1989). In these strains, the *polA1* dependent vectors can only function while the *polA1* allele is active, that is, only at the permissive temperature (30°C). A standard assay thus involves the establishment of the *polA1* dependent vector construct in the host at 30°C. This is followed by transfer of the culture to the non-permissive temperature (42°C). At 42°C, plasmid replication is controlled from the *polA1* independent replicon. Since the vector is unable to replicate at 42°C, the copy number of the plasmid will also be controled from the *polA1* independent replicon. Inactivation of *polA1* dependent replication at the non-permissive temperature, however, is not absolute and care must be taken in interpreting results.

The isolation of large or non-contiguous replicons from plasmids such as RK2 required a different approach (Figurski & Helinski, 1979). RK2 was partially digested with a restriction enzyme (*Hae*II) which cut the plasmid frequently. Once digested, the fragments were circularised by self-ligation. The mini-replicon containing fragments were isolated by selection for one of the encoded antibiotic resistance genes of RK2 (Thomas, 1987b). Characterisation of RK2 has also been carried out after transposon mutagenesis which was used to alter or inactivate specific replication domains (Shingler & Thomas, 1984) or to introduce new restriction sites to facilitate cloning of specific regions (Thomas, 1987b).

If the plasmid of interest is *polA1* dependent and cannot be established in *E. coli polA1* mutants, its replicon can be cloned in a temperature sensitive, *polA1* dependent vectors such as pSH1 (Hashimoto-Gotoh & Sekiguchi, 1977). pSH1 carries a point mutation in its essential replication gene, rendering it inactive at 42°C. The recombinants are first established at the permissible temperature for the vector (30°C), after which the temperature is increased to 42°C. At this temperature, the vector replication system is inoperative and replication is sustained by the cloned *polA1* dependent replicon.

A property which is a natural consequence of plasmid replication control and maintenance is incompatibility (Couturier *et al.*, 1988). Incompatibility is the inability of two plasmids to be stably inherited in the same host and results from plasmid relatedness (Novick, 1987). This relatedness reflects the sharing of common replication or partitioning determinants (Ishii *et al.*, 1978; Novick, 1987; Nordström & Austin, 1989). Historically, plasmid incompatibility was demonstrated for the F plasmids in the 1960s (Maas & Maas, 1962; Scaife & Gross, 1962), although a formal scheme of plasmid classification based on incompatibility was only developed later by Datta & Hedges (1971 and 1972).

Testing for incompatibility involves introduction of a plasmid by conjugation or transformation into a host strain carrying another plasmid (Bergquist, 1987). In the case of prophages, introduction is mediated via transduction (see Sternberg & Austin, 1981). The segregation of the plasmids is then monitored. Selection is mainly for the incoming plasmid, while the progeny are examined for maintenance of the resident plasmid. If the resident plasmid is displaced, then the two plasmids are incompatible and are assigned to the same incompatibility group.

Classical incompatibility testing has limitations. Firstly, the plasmid to be tested may not contain suitable markers or may not be transmissible. Secondly, the plasmid may be unable to transfer to the new host due to surface exclusion. Some of these limitations may be resolved by using mini-gal plasmids (Davey *et al.*, 1984) or "replicon specific probes" (Couturier *et al.*, 1988). The latter were constructed after identification of conserved regulatory regions within the basic replicons of plasmids which belong to known incompatibility groups. These regions were subcloned in multi-copy vectors for use as probes in hybridisation studies with plasmids of unknown incompatibility groups. Although "replicon typing" is not without its problems, it does circumvent the difficulties arising from surface exclusion and lack of conjugative properties. Replicon probes may be useful in plasmid classification if invariant regions with low potential for recombination can be identified.

This chapter deals with the isolation and functional characterisation of the dominant replicon of pGSH500 (repß).

MATERIALS AND METHODS

Bacterial strains, plasmids and bacteriophages

The bacterial strains, plasmids and bacteriophages used in this study are listed in *Appendix 1*. They include, *E. coli* strain LKIII (Zabeau & Stanley, 1982) which was used for plasmid maintenance. Strain GW125a (obtained from Professor D. Rawlings, Microbiology, UCT) is a *polA1*$^-$ derivative of *E. coli* AB1157 (*Appendix 1*) carrying the *polA1* allele, which was used to test for Pol I-independent plasmid replication. Plasmid pNZ116 (mini-F, f5) was obtained from Professor P. Bergquist and Dr. A. Spiers (University of Auckland, N. Zealand). *E. coli* strain J53-2(RM98) (IncN) as well as strains containing plasmids belonging to known incompatibility groups used in classical incompatibility assays were obtained from Dr. K. Towner (PHLS Laboratory, Nottingham, U.K.). pGSH500 was isolated from a *Klebsiella pneumoniae* strain (APP), cultured from the urine of a leukaemic patient at Groote Schuur Hospital (GSH). This plasmid, its mutant derivatives (pGSH510-pGSH540), as well as the recombinant plasmids constructed in this study are described in *Appendix 2* and *3*.

Media and antibiotics

Antibiotic sensitivity testing was done according to the method of Barber & Stokes (1966). Media composition is listed in *Appendix 4*. Luria broth cultures were supplemented with the following antibiotics (μgml^{-1}) when required: amikacin (10); ampicillin (50); gentamicin (10); kanamycin (30); naladixic acid (100); tetracycline (15); tobramycin (10). See also *Appendix 4*.

Enzymes and chemicals

Restriction and DNA modifying enzymes were obtained from Boehringer Mannheim, Amersham International, Bethesda Research Laboratories (GIBCO/BRL) or Promega. Nick-translation kits, Hybond-N membranes and [^{32}P]dCTP (3000 Ci/mmol), [^{35}S]dATP (> 1000 Ci/mmol) and were obtained from Amersham International. [methyl ^{3}H]Thymine (60 Ci/mmol) was purchased from Dupont, New England Nuclear. All reagents were of analytical grade.

Molecular biological techniques

Unless otherwise indicated, all techniques were followed as described in Sambrook *et al* (1989).

Bacterial transformations

E. coli DK-1 and LKIII cells were made competent by treatment with calcium chloride (Mandel & Higa, 1970). A modified method of Chung & Miller (1988), which included a heat shock step (2 minutes at 42^{o}C) was used to prepare competent *E. coli* GW125a

(*polA1*) cells. Transformation of GW125a was carried out within 1 hour of competent cell preparation. The cells were cultured at 37°C in L-broth for 3 hours prior to platting on selective media.

Bacterial conjugation

Conjugation of *K. pneumoniae* strain APP with other Gram-negative bacteria was carried out on L-agar. Donor and recipient strains were cultured to the exponential growth phase in L- broth at 37°C. Aliquots of donor and naladixic acid resistant recipient cells were diluted 100 fold and cultured in L-broth with vigorous aeration for a further 4 hours prior to conjugation. Donor and recipient cells were mixed (1:1) on L-agar plates and were left undisturbed at room-temperature for 2 hours to allow conjugation to take place. This was followed by incubation at 37°C overnight. The cells were washed off the plates with physiological saline, concentrated by centrifugation (4000 rpm for 15 min.) at room temperature and resuspended in 1 ml of physiological saline. Aliquots (100 μl) of the cell suspension (200-250 cells) were spread onto selective plates containing tobramycin (10 μgml^{-1}) and naladixic acid (100 μgml^{-1}) and were incubated overnight at 37°C. *E. coli* GW125a transconjugants were identified on MacConkey agar plates. In addition, the transconjugants were tested for the presence of antibiotic markers and sensitivity to bacteriophages. Transconjugant plasmid DNA was isolated by the method of Birnboim & Doly (1979).

Bacteriophage typing

The method of Khatoon and Iyer (1971) was used to determine if any of the phages listed in *Appendix 1* attached to the conjugative pili of *K. pneumoniae* APP(pGSH500).

Restriction enzyme analysis, DNA transfer, radiolabelling and hybridisation

Plasmid DNA was prepared by the method of Ish-Horowicz & Burke (1981) and was purified by two sequential caesium chloride density gradients. Plasmid DNA was restricted in the appropriate enzyme buffers and analysed by electrophoresis in horizontal agarose gels (0.7%), in a 40 mM Tris-acetate buffer, pH8.0, containing 1 mM EDTA. Electrophoresis was performed overnight at room temperature (1 volt/cm). When appropriate, the DNA fragments were transferred to Hybond-N membranes by the alkali transfer method (Amersham blotting and hybridisation protocols for Hybond-N membranes). DNA hybridisation procedures were carried out in the presence of 50% formamide, 5x Denharts solution and salmon sperm DNA (1mg/ml) at 42^{o}C as described in the Amersham protocols. The temperature of hybridisation and subsequent washes was adjusted depending on the stringency of hybridisation required. DNA was extracted from agarose gels by the method of Seth (1984) and was radiolabelled with [^{32}P]dCTP by nick-translation for use as probes.

Isolation of the replicons in pGSH500

The restriction enzymes *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Pst*I, *Sac*I, *Sau*3AI, *Sma*I and *Sph*I were used to construct pUC19 libraries of pGSH500. These were screened for recombinants containing essential replicative functions (*polA1* independence) by transformation into *E. coli* GW125a(*polA1*). Two distinct *polA1* independent clones were identified (pFDT100 and pFDT200).

Construction of deletion derivatives of pFT200

The plasmid pFDT210, which was obtained by deleting 0.4 Kb of pFDT200 with exonuclease III (Henikoff, 1984), maintained all the replicative properties of the original clone (pFDT200) and was used in the generation of all other derivatives used in this study (*Appendix 2.*). DNA fragments were excised from pFDT210, blunt-ended by treatment with Klenow enzyme (5 Uml^{-1}) and cloned in the *Sma*I site of pUC19 to generate pFDT211 (0.2 Kb *Hpa*I-*Pst*I); pFDT212 (0.80 Kb *Bst*XI-*Pst*I); pFDT213 (1.00 Kb *Hpa*I); pFDT214 (1.03 Kb *Hin*cII); pFDT215 (0.66 Kb *Eco*RI-*Hin*cII); pFDT216 (1.80 Kb *Eco*RI-*Bst*XI) and pFDT220 (2.4 Kb *Eco*RI-*Ava*II). Ordered deletions of *Eco*RV/*Kpn*I digested pFDT220 (pFDT2201- pFDT2208) were obtained by treatment with exonuclease III. Similarly, sequential deletions from the opposite end of pFDT220 (pFDT220201-pFDT220210) were generated with exonuclease III, after prior digestion of the plasmid with *Pst*I and *Bam*HI.

The mini-replicons pFDT200Gm^r and pFDT220204Gm^r were constructed to exclude any pUC19 vector sequences other than those present in the multiple cloning site. pFDT200Gm^r was constructed by ligating the gel purified 3.30 Kb *Eco*RI-*Hin*dIII fragment from pFDT200, to a 2.1 Kb AAC(3) gene (Elisha & Steyn, 1991). Similarly, pFDT220204Gm^r was constructed by ligating the gel purified 1.30 Kb *Eco*RI-*Hin*dIII fragment of pFDT220204 to the AAC(3) gene.

Complementation assays

The strain GW125a(pFDT200Gm^r) was made competent by the modified method of Chung & Miller (1988). This strain was used to determine if any of the plasmids that were unable

to replicate autonomously (pFDT220206-pFDT2202210), could be rescued by pFDT200Gm^r.

Plasmid stability and incompatibility tests

The stability properties of the plasmids pGSH500, pFDT100 and pFDT200 were assessed in GW125a(*polA1*), essentially as described by Meacock & Cohen (1980). Single colonies of plasmid containing bacterial strains to be tested for stability of plasmid maintenance were inoculated in selective L-broth and cultured 37^{o}C until the cell density was $5x10^9$ /ml. The cultures were diluted 10^6 fold with fresh, non-selective L-broth and incubated at 37^{o}C for up to 180 generations, with 10^6 fold dilutions into fresh non-selective medium at every 20 generations or 12 hours. Cell viability was assessed at 8, 20, 60, 120 and 180 generations, by plating 100 μl aliquots (diluted 10^5 fold) on non-selective L-agar. From each plate 36 colonies were screened for the presence of the plasmid. A plasmid was considered to be stably maintained if 90% of the cell population still harboured the plasmid after 120 generations under non-selective conditions.

The ability of the plasmids pFDT210-pFDT220210 to displace pFDT200Gm^r was used to locate the incompatibility determinants of repß. The incompatibility properties of pFDT200Gm^r towards pGSH500, pFDT100 and pUC19 were investigated as described (Meacock & Cohen, 1980). Bacterial cells containing a resident plasmid were transformed with a second plasmid (incoming plasmid). Large incoming plasmids were introduced in the bacterial cells by conjugation. Hosts carrying both plasmids were maintained on double selective media. These hosts were subcultured as described above in the absence of selection or with selection for either the resident or the incoming plasmid. Aliquots of these cultures were screened for the presence of both plasmids at the generations indicated above. The criteria of Meacock & Cohen (1980) were used to determine the degree of incompatibility. The ratio of the number of transformants carrying both plasmids to that of

the transformants carrying only the incoming plasmid was determined. A value of 0 (zero) denotes incompatibility, whereas a value of 1 (one) represents compatibility. Incompatibility was considered high at a ratio of <0.4.

Determination of the incompatibility group of pGSH500

Classical incompatibility studies were carried out as described by Bergquist (1987). The plasmid pGSH500 was conjugated with strains containing standard plasmids of known incompatibility groups, followed by establishment and segregation tests.

Identification of the incompatibility group(s) of pGSH500 was also attempted by molecular hybridisation studies using cloned *inc/rep* probes (pULB2154 to pULB2410) as described by Couturier *et al.* (1988). These probes contain the incompatibility determinants of representative plasmids of IncFI, IncFII, com9, IncI1, IncB/O, IncK, IncHI1, IncHI12, IncL/M, IncN, IncP, IncQ, IncT, IncU, IncW, IncX, and IncY.

Plasmid copy number

Two methods were used to determine plasmid copy number in GW125a(*pol*A1) (Thomas, 1987b; Nordström *et al.*, 1980). In the first method, the cells containing the relevant plasmids were grown in the presence of [methyl-^{3}H]thymine. Plasmid copy number was estimated by separating tritiated, covalently closed circular, plasmid DNA from total linear DNA on ethidium bromide-caesium chloride density gradients. The ratio of radioactivity in the CCC DNA fraction to that of the total DNA fraction was used to calculate the plasmid copy number (Thomas, 1987b). In addition, the plasmid copy number relative to that of pFDT100 was assessed by determining the host cell resistance to ampicillin at antibiotic concentrations ranging from 50-3000 $\mu g ml^{-1}$ (Nordström *et al*, 1980). Bacterial cells were

grown overnight in L-broth containing 50 μgml^{-1} of ampicillin. The cultures were diluted 10^5, 10^6 and 10^7 fold and an aliquot of each dilution was plated onto L-agar plates containing ampicillin concentrations of 50 to 3000 μgml^{-1}, and were incubated overnight at 37°C. The degree of resistance to ampicillin was proportional to the antibiotic gene dosage and therefore to plasmid copy number.

RESULTS

Properties of pGSH500

The plasmid pGSH500, which was isolated from a clinical strain of *K. pneumoniae* (APP), was found to encode resistance to several antibiotics (*Appendix 2A*). pGSH500 is a large plasmid with a molecular weight of 107 Kb (*Appendix 3*). In addition, pGSH500 can undergo spontaneous deletions to give rise to deletion mutants (pGSH510-pGSH540) ranging in size from 76 Kb to 10 Kb (*Appendix 2.*; *Figure 2.1.*). Unlike the deletion mutants, which are non-conjugative, pGSH500 could be transferred to *Acinetobacter baumannii*, *Citrobacter diversus*, *Proteus mirabilis*, *Morganella morganii* and to *E. coli,* with a frequency of transfer ranging from $2.5x10^{-3}$ to $8.6x10^{-2}$. In general, it is stably maintained in these strains. An exception is the *E. coli* strains containing an F', from which pGSH500 is lost after prolonged storage (more than 1 year) at 4°C, under non-selective conditions. Conjugation of pGSH500 with either *Ps. aeruginosa* or *Ps. putida* was unsuccessful. Thus, pGSH500 appears to have a moderately promiscuous host range, predominantly within the *Enterobacteriaceae*.

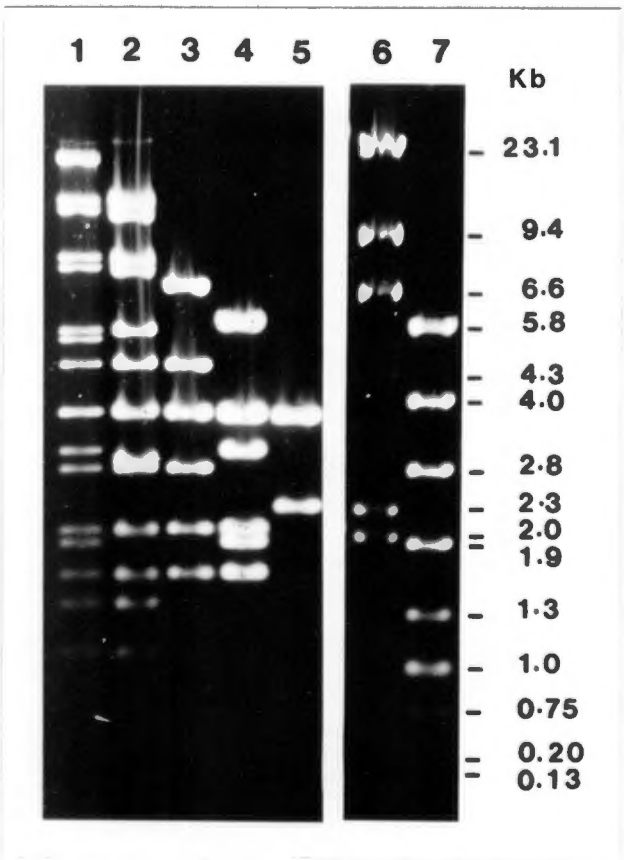


Figure 2.1. *Hin*dIII restriction digest of pGSH500 and its spontaneous deletions. Lane 1, pGSH500; Lane 2, pGSH510; Lane 3, pGSH520; Lane 4, pGSH530; Lane 5, pGSH540; Lane 6, molecular weight marker (Kb), *Hin*dIII digest of lambda DNA; Lane 7, molecular weight marker (Kb), *Eco*RI digest of pAT15 containing a fragment of pox DNA (K. R. Dumbell, UCT). Electrophoretic conditions are described in the text.

Since pGSH500 exhibited a moderately promiscuous host range, it was decided to determine if its transfer system was similar to that of known conjugative systems. Bacteriophage typing of strains APP(pGSH500) and J53(pGSH500) was carried out with the IncN specific bacteriophage Ike, as well as with bacteriophage PR772, which is specific for the pili of the promiscuous plamids IncN, IncP, IncW, and with bacteriophage X, which has a broader specificity, namely, IncM, IncN, IncP-1, IncU, IncW, IncX. None of the bacteriophages were able to infect host cells containing pGSH500, whereas the positive controls J53(RP4) and GW125a(pCU1) were infected. Strains GW125a(pNZ116), GW125a and J53 were used as a negative controls. It was concluded that the pili encoded in pGSH500 are distinct from those of typical promiscuous plasmids.

pGSH500 could not be assigned to a known incompatibility group and results suggest that the incompatibility group of this plasmid may be novel. Classical incompatibility studies demonstrated that pGSH500 was compatible with all the plasmids tested, namely R386 (IncFI), R1 (IncFII), R27 (IncH1), R64 (IncIL), R391 (IncJ), R496b (IncL), N3 (IncN), RP4 (IncP), S-a (IncW), R6K (IncX) and P1 (IncY). The incompatibility of pGSH500 was also investigated by the replicon specific DNA probes (pULB2154 to pULB2410) of Couturier *et al.* (1988). These contain the *inc/rep* regions of the plasmids F (IncFI), P307 (IncFI), R1*drd-19* (IncFII), pIP71 (com9), R64*drd-11* (IncI1), pMU700 (IncB/O), R387 (IncK), TR6 (IncHI1), TP116 (IncH12), pMU407.1 (IncL/M), R46 (IncN), RK2 (IncP), R1162 (IncQ), Rst1 (IncT), RA3 (IncU), RSa (IncW), R6K (IncX), P1 (IncY). The results of these replicon typing assays for pGSH500 were ambiguous (*Figure 2.2*). Under standard conditions of hybridisation which allows for a 10% mismatch, the only hybridisation signal obtained was that between the IncN probe (pULB2432, R46) and pGSH500 (*Figure 2.2A, B & C*). When the hybridisation stringency was decreased to 30% mismatch, a signal was obtained between the IncP probe (pULB2420) and pGSH500 (*Figure 2.2D*). At this level of stringency, a signal was also detected between this probe and pFDT200 (an autonomously replicating fragment in pGSH500, see next section), as well as with pUC19

(*Figure 2.2D*). In addition, hybridisation signals were obtained between the IncP probe and other replicons. Decreasing the stringency of hybridisation to 50% mismatch resulted in the detection of a signal between the IncN probe and pFDT200, but a signal was also obtained between this probe and pUC19 (*Figure 2.2E*). No signal was obtained between any of the probes and pFDT100 (another autonomously replicating fragment in pGSH500, see next section). Therefore, it was concluded that pGSH500 contains a region with sequence similarity to the IncN probe (pULB2432, R46). This region may be part of a replicon which is distinct from the two cloned replicons pFDT100 and pFDT200.

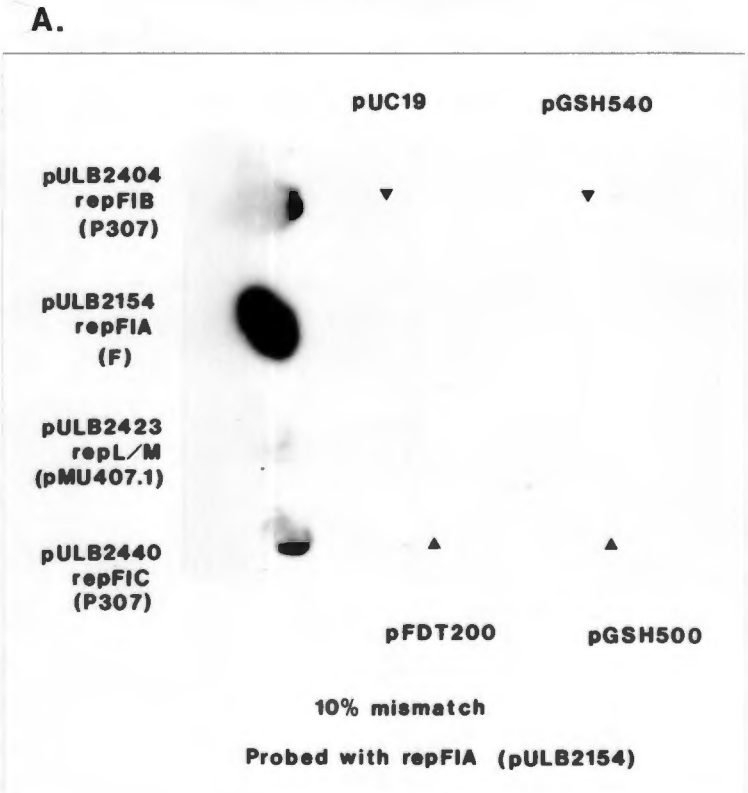


Figure 2.2. Autoradiographs of colony hybridisation blots of strains tested during replicon typing. Strains to be tested were grown (over-night, at 37^{o}C) on Hybond-N membranes (Amersham, U.K.), laid over L-agar with antibiotic(s) corresponding to the plasmid(s) resistance markers. The membranes were treated as suggested in the Amersham blotting and hybridisation protocols for Hybond-N membranes. Hybridisation conditions are described in the text. Strains containing replicon probes, pGSH500 and derivative plasmids (pFDT100 and pFDT200) were probed as described in the panels (**A**-**E**). Panel **A**. Probing of blots with repFIA (pULB2154), 10% mismatch; Panel **B**. Probing with repP (pULB2420), 10% mismatch; Panel **C**. Probing with repN (pULB2432), 10% mismatch; Panel **D**. Probing with repP (pULB2420), 30% mismatch; Panel **E**. Probing with repN (pULB2432), 50% mismatch (see next page also).

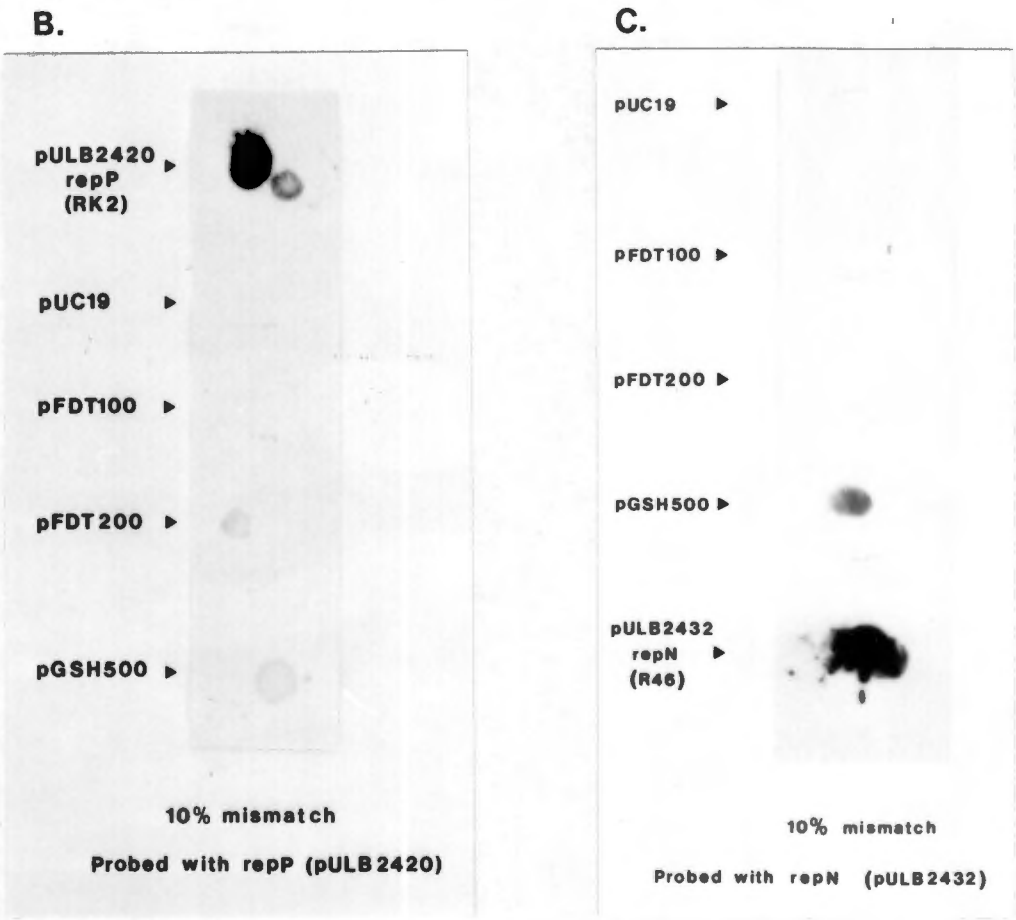


Figure 2.2. Autoradiographs of colony hybridisation blots of strains tested during replicon typing (cont.).

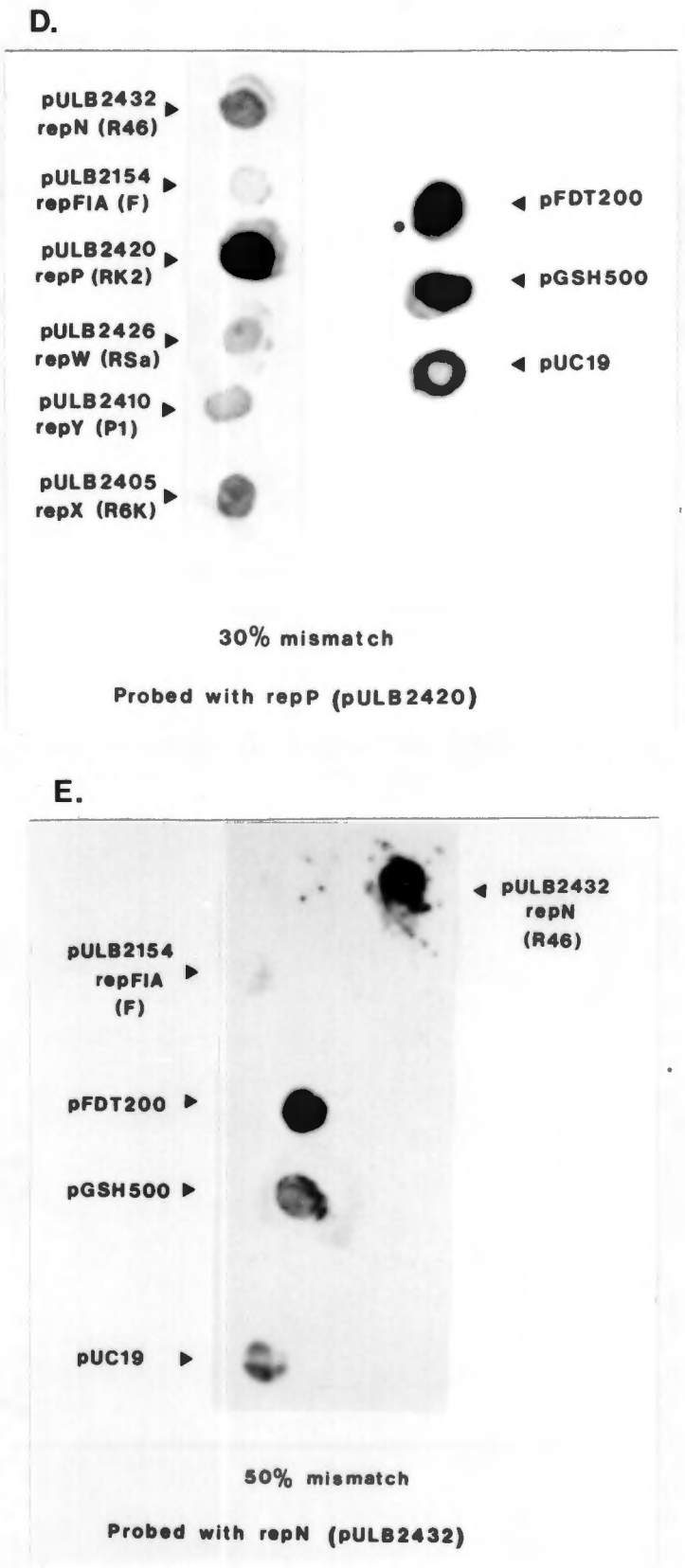


Figure 2.2. Autoradiographs of colony hybridisation blots of strains tested during replicon typing (cont.).

Identification of 2 autonomously replicating regions in pGSH500

Plasmids containing pMB1/ColE1 derived origins of replication, such as pUC19, are dependent on host encoded functions for replication and are unable to replicate in *polA1* deficient bacterial strains. In contrast, the replication of pGSH500 is *polA1* independent, enabling it to be stably maintained in *E. coli* GW125a (*polA1*). DNA libraries of pGSH500 were constructed in pUC19 using the restriction enzymes *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Pst*I, *Sac*I, *Sau*3AI, *Sma*I and *Sph*I. Following transformation of the pGSH500 DNA libraries in GW125a(*polA1*), two distinct *polA1* independent clones were isolated, namely GW125a(pFDT100) and GW125a(pFDT200) (*Figure 2.3*). These contained a 3.8 Kb *Hin*dIII and 3.3 Kb *Pst*I DNA fragment, respectively, and were shown to be unrelated by DNA hybridisation and restriction enzyme mapping (*Figures 2.4. & 2.5.*). Therefore, pGSH500 contains at least two basic replicons, designated repα (contained within the 3.8 Kb *Hin*dIII insert of pFDT100) and repß (contained within the 3.3 Kb *Pst*I insert of pFDT200). repα is present in all four spontaneous deletion mutants of pGSH500 (pGSH510-pGSH540, *Appendix 2*), whereas repß is present only in pGSH510 (*Figure 2.4.*).

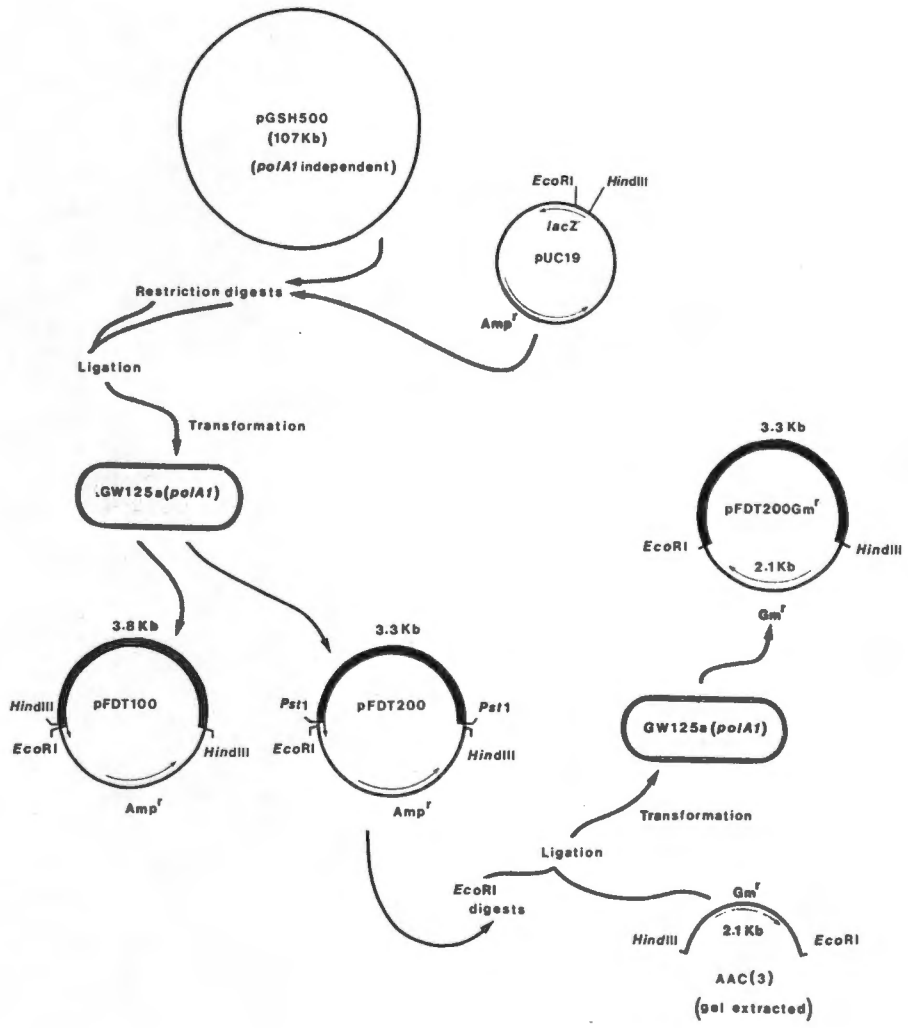


Figure 2.3. Outline of the cloning strategy used to isolate the autonomously replicating regions of pGSH500 in GW125a, GW125a(pFDT100) and GW125a(pFDT200). pFDT100 includes a 3.8 Kb *Hin*dIII fragment (repα), whereas pFDT200 contains a 3.3 Kb *Pst*I fragment (repß). The 3.3 Kb fragment in pFDT200 can form a mini-replicon when ligated to a resistance marker (Gm^r) to form pFDT200Gm^r. The latter does not contain any vector sequences.

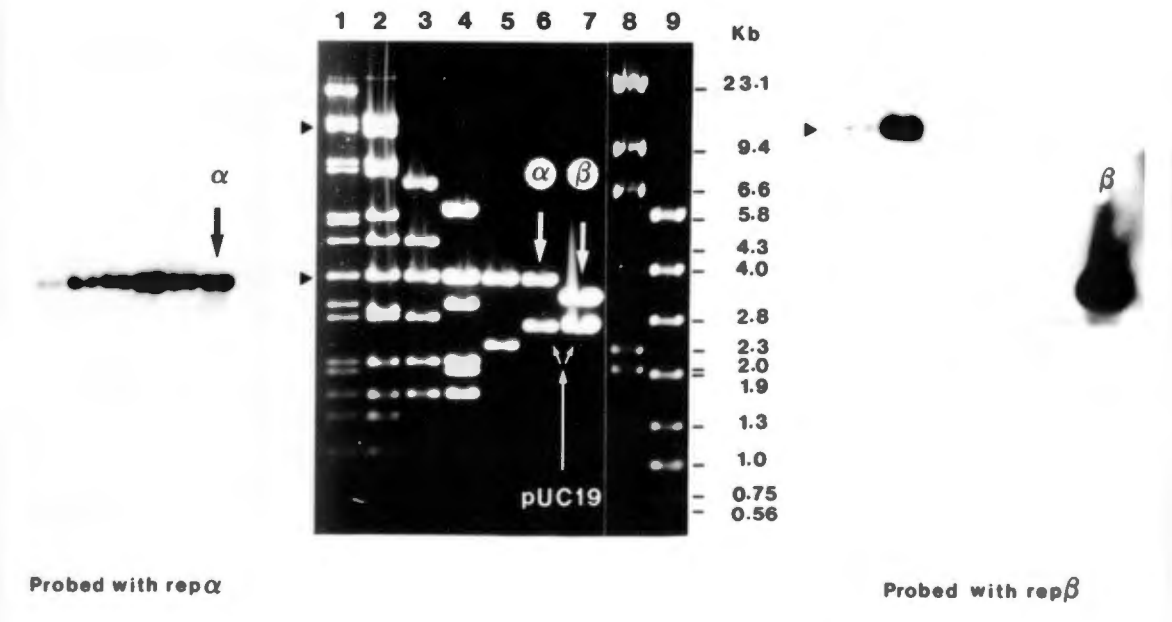


Figure 2.4. Southern hybridisation of pGSH500 and its spontaneous deletion derivatives (pGSH510-pGSH540) digested with *Hin*dIII and probed with either repα (pFDT100, left panel) or repß (pFDT200, right panel). Center panel: Lane 1, pGSH500; Lane 2, pGSH510; Lane 3, pGSH520; Lane 4, pGSH530; Lane 5, pGSH540; Lane 6, *Hin*dIII digest of pFDT100; Lane 7, *Pst*I digest of pFDT200; Lanes 8 and 9, include molecular weight markers as described in ***Figure 2.1.*** Left panel: the 3.8 Kb *Hin*dIII fragment containing repα is present in pGSH500 as well as in all its deletion derivatives. Right panel: the 3.3 Kb *Pst*I fragment containing repß is present only in pGSH500 and in pGSH510. No cross-hybridisation signal was detected between repα and repß.

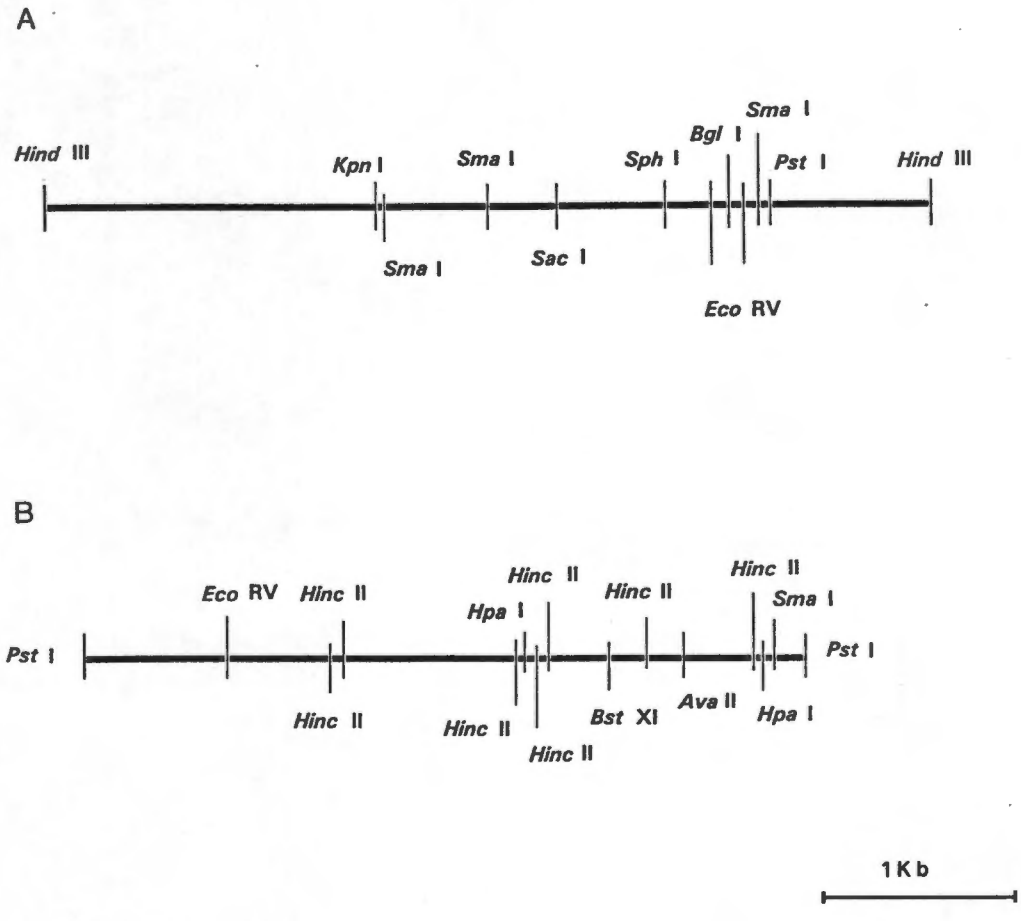


Figure 2.5. **A.** Restriction enzyme map of the 3.8 Kb *Hind*III fragment containing repα (pFDT100); **B.** restriction enzyme map of the 3.3 Kb *Pst*I fragment containing repß (pFDT200).

Identification of repß as the dominant replicon in pGSH500

The dominant replicon of pGSH500 was identified after determination of copy number of this plasmid and its replicons.

It has been proposed that the copy number of a plasmid chimera containing two or more replicons with different copy numbers, will be similar to that of a component plasmid containing the replicon which confers the higher copy number (Timmis *et al.*, 1975; Pritchard and Grover, 1981). Thus in a plasmid with multiple replicons, the replicon with the highest copy number is considered to be dominant. The dominant replicon of pGSH500 was therefore identified by estimation of plasmid copy number of pGSH500 as well as that of each of the component replicons repα (pFDT100) and repß (pFDT200) in GW125a(*polA*1).

Two methods of copy number determination were used and both gave consistent results (*Table 2.1*). The copy number of pGSH500 ranged from 3-5 copies per chromosome, while that of pFDT100 ranged from 1-2 copies per chromosome and that of pFDT200 from 3-5 copies per chromosome. In addition, the copy number of pNZ116 (*Eco*RI-f5 fragment, IncFI), pCU1 (IncN) and pUC19 was estimated as controls to assess the reliability of the methods. The copy numbers obtained correspond to those reported in the literature (*Table 2.1.*; Lane, 1981; Kozlowski, *et al.*, 1987). From these data it was concluded that repß (pFDT200) is the dominant replicon in pGSH500.

Table 2.1A. Plasmid copy number determined by dye-buoyant density gradients

Plasmid	Size (Kb)	Copy number
pGSH500	107.00	3-5
pFDT100	6.40	1-2
pFDT200	6.10	3-5
pFDT220202	4.32	3-5
pFDT220203	4.25	4-5
pFDT220204	4.00	15-20
pFDT220205	3.80	unstable*
pNZ116	13.46	1-2
pCU1	39.00	6-8

*copy number could not be determined.

Table 2.1B. Relative plasmid copy number determination. Plasmid copy number estimation by assessment of the level of antibiotic resistance of the *E. coli* host GW125a(*polA*1) carrying each of the plasmids listed.

Plasmid	**Concentration of Ap (μgml^{-}1) necessary to inhibit growth**	**Copy number relative to that of pFDT100**
pFDT100	100	1
pFDT200	200	2
pFDT210	220	2
pFDT220	200	2
pFDT220202	200	2
pFDT220203	400	4
pFDT220204	1500	15
pUC19	>3000	>>30

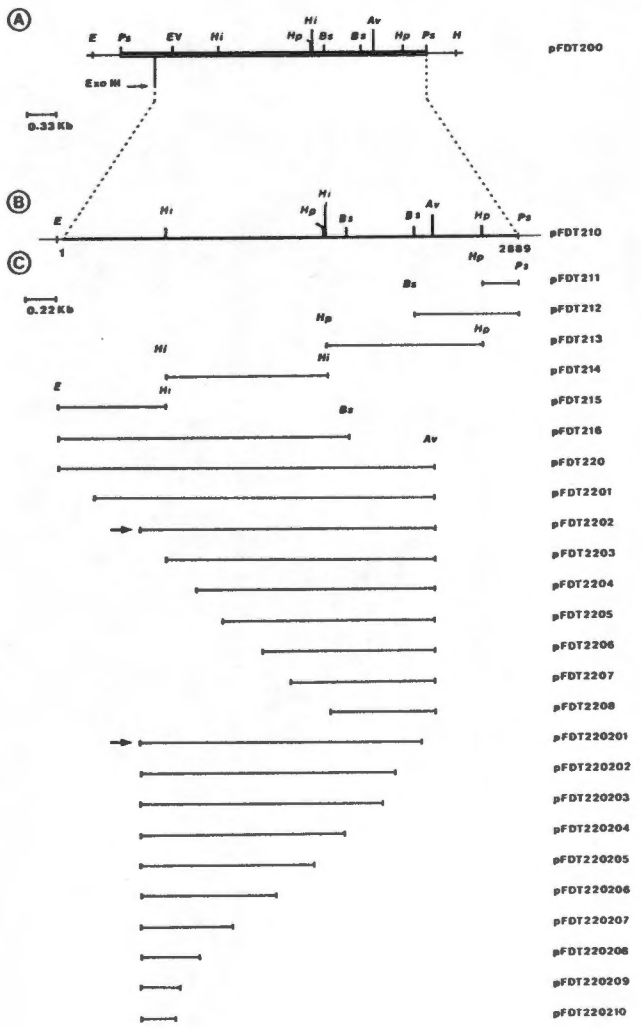
Delimitation of repß

A. Delimitation of the autonomous replicating domain

A restriction enzyme map of pFDT200 *(Figures 2.5B. & 2.6A.)* was used to plan a subcloning strategy suitable for the functional analysis of repß (*Figure 2.6B & C.*). Exonuclease III deletion from one end of the 3.3 Kb *Pst*I insert generated a 2.9 Kb derivative (pFDT210) which was capable of autonomous replication and stable plasmid maintenance in GW125a(*polA1*) (*Figure 2.6B & C.*). The plasmid derivatives of pFDT210 (pFDT211- pFDT220) that were generated following cloning of specific DNA fragments into pUC19, were tested for replication functions in GW125a(*polA1*) (*Figure 2.6C & D.*). Autonomous replication was associated with pFDT216 and pFDT220. Thus, the essential elements for replication must be contained on pFDT216. However it was noticed that the plasmid yields of pFDT216 were higher than those of pFDT220. This suggested that copy number regulatory functions present in pFDT220 have been deleted in pFDT216.

The plasmids pFDT2201 to pFDT2208, containing progressive deletions of pFDT220 (*Figure 2.6C.*), were tested in GW125a(*polA1*) for autonomous replication (*Figure 2.6C.*). Of these, only pFDT2201 and pFDT2202 were able to replicate in the *polA1* deficient host. Thus, an essential region for replication, lacking in pFDT2203, must be present in pFT2202. To determine the extent of the domain necessary for replication, pFDT2202 was deleted with exonuclease III from the opposite end. This treatment generated deletions pFDT220201 to pFDT220210. Stable, autonomous replication was demonstrated by plasmids pFDT220201 to pFDT220204. The plasmid pFDT220205 could replicate in GW125a(*polA1*), but it was highly unstable. When GW125a(pFDT220205) was grown for more than 4-5 hours in broth containing ampicillin, a large proportion of the cells lysed with resultant low plasmid yields on purification. It can be concluded that the minimal sequence required for autonomous replication of repß, *rep* domain, is contained on the

1.3 Kb insert of pFDT220204. This was confirmed by the construction of a pFDT220204Gm^r mini-replicon in the absence of any vector sequences (*Appendix 2B.*).



	REPLICATION		INCOMPATIBILITY
	GW125 a	GW125 a (pFDT200Gm^r)	
pFDT200	+		+
pFDT210	+		+
pFDT211	-	-	-
pFDT212	-	-	+
pFDT213	-	-	+
pFDT214	-	-	+
pFDT215	-	-	-
pFDT216	+		+
pFDT220	+		+
pFDT2201	+		+
pFDT2202	+		+
pFDT2203	-	-	+
pFDT2204	-	-	+
pFDT2205	-	-	+
pFDT2206	-	-	+
pFDT2207	-	-	+
pFDT2208	-	-	+
pFDT220201	+		+
pFDT220202	+		+
pFDT220203	+		+
pFDT220204	+		+
pFDT220205	+*		-
pFDT220206	-	+	-
pFDT220207	-	+	-
pFDT220208	-	+	-
pFDT220209	-	+	-
pFDT220210	-	-	-

Figure 2.6. Outline of subcloning strategy used for the functional analysis of repß (**A.-C.**) to identify regions involved in replication, incompatibilty and copy number control functions (**D.**) (full legend on following page).

Figure 2.6. Outline of subcloning strategy for functional analysis of repß. **A.** Restriction enzyme map of the pGSH500 insert in pFDT200 containing repß. The restriction sites for *Ava*II (*Av*), *Bst*XI (*Bs*), *Eco*RV (*EV*), *Hin*cII (*Hi*), *Hpa*I (*Hp*) and *Pst* (*Ps*) are shown; *Eco*RI (*E*) and *Hin*dIII (*H*) belong to the multiple cloning site of the vector. The region designated ExoIII indicates the portion of DNA which was deleted (0.4 Kb) to give rise to pFDT210 **(B). C.** Outline of the strategy used to sequence and to determine the functional regions of pFDT210. The plasmids pFDT211 to pFDT220 contain subcloned regions of pFDT210, whereas the plasmids pFDT2201 to pFDT220210 are exonuclease III deletions of pFDT220. **D.** Recombinant plasmid constructs of repß that are able to replicate autonomously in *E. coli* GW125a are indicated (+) in the first column; those that can be rescued *in trans* by the mini-replicon pFDT200Gm^r in GW125a are indicated in column two, while those that are incompatible with pFDT200Gm^r are indicated in the third column (see *Table 2.2).* * Unstable plasmid.

B. The cis-acting region of repß (*oriV*)

The plasmids pFDT220206-pFDT220210 were unable to sustain autonomous replication in GW125a(*polA1*). However, when the plasmid pFDT200Gm^r was co-resident in GW125a(*polA1*), all of these plasmids with the exception of pFDT220210, were rescued *in trans*. A *cis*-acting domain in repß (*ori*V) is, therefore, contained on the 0.25 Kb insert of the plasmid pFDT220209 (*Figure 2.6C & D.*).

C. Location of the incompatibility and copy number control determinants

The mini-replicon pFDT200Gm^r is fully compatible with pUC19, pFDT100 as well as with its parent pGSH500 (*Table 2.2*). The ability of some derivatives of pFDT210 to displace the mini-replicon pFDT200Gm^r from DK-1 after 120 generations was used to locate the region containing the plasmid incompatibility determinants. Incompatibility was associated with the plasmids pFDT212-pFDT214 and pFDT216 (*Figure 2.6C. & D.*; *Table 2.2*). Since the insert in pFDT216 overlaps those of pFDT214 and pFDT215 and as pFDT215 did not displace pFDT200Gm^r, the incompatibility determinants are located within the cloned fragments of pFDT212-pFDT214. Furthermore, as pFDT200Gm^r was not displaced by pFDT211, the incompatibility determinants do not extend over the full length of pFDT212. Thus, the incompatibility determinants of ß replicon must be contained within the region overlapped by pFDT213 and pFDT216. In addition to phenotypic characteristics, plasmid DNA was isolated to confirm the presence of 2 plasmids in the same host.

To identify the precise location of the incompatibility determinants, the exoIII derivatives of pFDT220 (pFDT2201-pFDT220210) were tested against pFDT200Gm^r. Results showed that the plasmids pFDT220 to pFDT220204 were incompatible with pFDT200Gm^r (*Table 2.2.*). In contrast, the plasmids pFDT220206-pFDT220210 were compatible with

pFDT200Gm^r. Thus the incompatibility domain must be contained within a 1 Kb fragment *Hpa*I, contained in plasmid pFDT213.

The plasmid DNA yield of pFDT220204 and of its corresponding mini-replicon pFT220204Gm^r was greater than that from GW125a(pFDT200), GW125a(pFDT210) or GW125a(pFDT220) (*Table 2.1*). This indicated that elements involved in plasmid copy number control were deleted in the smaller plasmids (pFDT220203-pFDT220204). The deleted region corresponds to that containing the incompatibility domain of repß (pFDT213) (*Figure 2.6*). The involvement of pFDT213 in copy number control was confirmed following estimation of plasmid copy number of intact and deleted plasmids by two independent assays (*Table 2.1.*). Attempts at reducing the copy number of pFDT220204 by providing pFDT213 *in trans* were unsuccessful due to the incompatibility between these two plasmids.

Table 2.2. Identification of regions of repß* containing incompatibility determinants.

	Selection for		
Incoming Plasmid	**Incoming Plasmid**	**Resident Plasmid**	***inc***
pUC19	1	1	-
pGSH500	1	1	-
pFDT100	1	1	-
pFDT200	0	0	+
pFDT210	0	0	+
pFDT211	1	0.9	-
pFDT212	0.2	0.2	+
pFDT213	0	0	+
pFDT214	0.2	0.2	+
pFDT215	1	1	-
pFDT216	0.1	0.1	+
pFDT220	0	0	+
pFDT2201	0	0	+
pFDT2202	0	0	+
pFDT2203	0	0	+
pFDT2204	0	0	+
pFDT2205	0	0	+
pFDT2206	0	0	+
pFDT2207	0	0	+
pFDT2208	0	0	+
pFDT220201	0	0	+
pFDT220202	0	0	+
pFDT220203	0	0	+

Table 2.2. (continuation)

	Selection for		
Incoming Plasmid	**Incoming Plasmid**	**Resident Plasmid**	***inc***
pFDT220204	0.1	0.1	+
pFDT220205**	1	1	-
pFDT220206	1	1	-
pFDT220207	1	1	-
pFDT220208	1	1	-
pFDT220209	1	1	-
pFDT220210	1	1	-

*In these assays, pFDT200Gm^r mini-replicon is the resident plasmid. The degree of incompatibility is indicated by a value ranging from 0 (incompatible) to 1 (compatible) (see Methods). In these assays only every alternate plasmid from pFDT220-pFDT2210 were tested, unless the results were ambiguous. The incompatibility of the plasmids not tested was inferred by examining the behaviour of plasmids including flanking regions. These results are also summarised in *Figure 2.6D*.

**Replicon unstable.

In summary, the replication, incompatibility and copy number functions of repß must be contained on the 2.2 Kb sequence in the plasmids pFDT220204 and pFDT213 (overlapping pFDT210) (*Figures 2.6 & 2.7*).

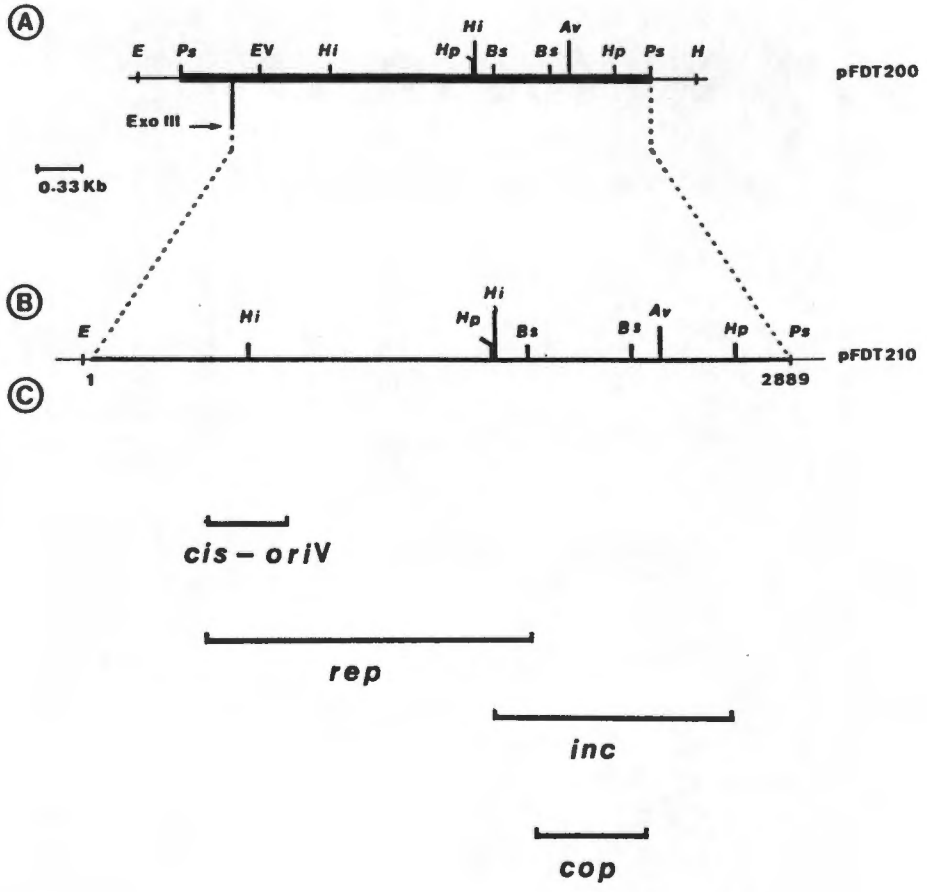


Figure 2.7. Location of the replication, incompatibility and copy number functions of repß within a 2.2 Kb sequence in pFDT210. *ori*V, indicates the region containing the *cis*-acting region of plasmid replication. *rep*, indicates the minimal region of repß capable of autonomous replication in GW125a. *inc/cop* indicate the presence of incompatibility determinants and copy number control function.

Plasmid stability properties

The plasmids pGSH500, its mini-replicons and some plasmids with deletions of repß (pFDT210-pFDT220205, *Table 2.3.*) were tested in GW125a(*polA1*) and in DK-1 to assess their stability in the absence of selective antibiotic pressure. There was no difference between the results obtained with the two strains. The results obtained with GW125a(*polA1*) are shown in *Table 2.3.* pGSH500 was stably maintained in GW125a(*polA1*) in the absence of selection over 180 generations. The stability of the mini-replicons pFDT100 (repα) and pFDT200Gm^r (repß) was the same as that of pGSH500 over 120 generations. Beyond 180 generations, these plasmids were not as stable as the parent plasmid. The stability of each of the pUC19 recombinants pFDT210, pFDT216, pFDT220, pFDT2202, pFDT220204 and pFDT220205 was also tested in GW125a(*polA1*). After 180 generations ≈80% of the cells tested still contained the plasmid (*Table 2.3.*). Thus, these constructs are slightly less stable than the parent plasmids pGSH500 and pFDT200. The plasmid pFDT220205 was unstable in GW125a(*polA1*); after 60 generations 66% of the cells tested were plasmid free. In addition to the loss of plasmid, the surviving colonies of GW125a(pFDT220205) exhibited a crenated morphology, as if sections of the colony had lost viability. This correlated with the previous observation that after 4-5 hours growth in broth containing ampicillin, a large proportion of GW125a(pFDT220205) lysed. Consequently, a low plasmid yield was obtained from this strain (see section, **A. Delimitation of the autonomous replicating domain**). The loss of viability could not be prevented by providing the plasmid pFDT200Gm^r *in trans*. A region deleted in pFDT220205, but present in pFDT220204, is essential for the stable maintenance of repß.

Table 2.3. Determination of plasmid stability in GW125a(*polA1*).

Plasmid	% cells containing plasmid after 60, 120 and 180 generations in the absence of antibiotic selection		
	60	**120**	**180**
pGSH500	100	100	100
pFDT100	100	97	82
pFDT200Gm^r	100	95	84
pFDT210	100	95	81
pFDT216	100	86	84
pFDT220	100	96	84
pFDT220202	97	88	80
pFDT220204Gm^r	98	82	84
pFDT220205	66	41	13

DISCUSSION

The conjugative, multi-drug resistance, plasmid pGSH500 exhibits a moderately promiscuous host-range mostly within the *Enterobacteriaceae*. The inability to determine the incompatibility group of pGSH500 by classical methods suggested that the incompatibility group is novel. Yet, during replicon typing a hybridisation signal was obtained between the IncN probe and pGSH500. This suggested that pGSH500 may contain a replicating region with sequence similarity to the IncN rep probe. Two unique *polA1* independent clones corresponding to the replicons, repα (pFDT100) and repß (pFDT200) were identified in pGSH500. Of these, only pFDT200 was shown to hybridise to the IncN probe. This signal was considered to be non-specific as it was obtained under low stringency (50%); at this level of stringency a signal was also obtained with pUC19. In consequence, a third replicating region distinct from repα and repß may be present in pGSH500. This additional region, however, may include only vestiges of an intact replicon or may contain restriction sites for the enzymes used during the cloning of repα and repß. Alternatively, the third replicating region may not be *polA1* independent and would require a different strategy of isolation.

Large plasmids such as those belonging to the IncFI group, often contain several replicons, some of which may be inactive (Bergquist, 1987). In F, repFIC is inactivated due to the insertion of Tn*1000* (Guyer, 1978; Bergquist *et al.*, 1986). In contrast, this replicon is intact and functional in EntP307 (Picken *et al.*, 1984; Saadi *et al.*, 1987). The structure of the replicons of pGSH500 could be similar to that described for F (Bergquist *et al.*, 1986).

Since pGSH500 was shown to be compatible with the IncN plasmid N3 (classical incompatibility assays), the significance of the signal with the IncN probe is unclear. These results illustrate that the usefulness of the *inc/rep* probes is restricted to plasmids whose replicons are similar to those from which the probes were constructed. Furthermore, the

bank of replicon probes needs to be continuously updated. In addition, the short repetitive sequences present in most replicons, may facilitate recombination events to generate new basic replicons.

The mini-replicons repα and repß are fully compatible with each other and with the parent plasmid. Furthermore, it appears that if one of the replicons of pGSH500 is challenged, the unchallenged replicon becomes active and no incompatibility occurs. This conclusion is supported by the observation that pGSH540, which contains only repα, was incompatible with pFDT100, but compatible with pFDT200. It was concluded that the unique incompatibility properties of pGSH500 were due to the presence of at least two fully compatible, unique replicons.

The presence of several basic replicons in the same plasmid is a frequent occurrence among members of the IncF group (F, P307, R386, R6-5, R124) (Bergquist *et al.*, 1986; Bergquist, 1987) and has been shown to lead to ambiguities in incompatibility classification of plasmids, as observed in pGSH500. In multi-replicons, some incompatibility determinants are suppressed and the interaction of the replicons may result in the generation of entirely different incompatibility behaviour (Couturier *et al.*, 1980) This is illustrated with the bireplicon plasmid pCG86 (Picken *et al.*, 1984; Maas *et al.*, 1989). pCG86 was shown to be compatible with IncFI plasmids and to displace IncFII plasmids. Thus, this plasmid was assigned to the latter group. Subsequently, it was observed that a spontaneous deletion of pCG86, which lacked repFIIA/FIC, was still able to replicate autonomously. This suggested that an additional replicon was present in pCG86. This additional replicon was shown to be compatible with IncFII plasmids but not with those of the IncFI group. Therefore, the second replicon was assigned to IncFI. Further characterisation of pCG86 confirmed the presence of 2 basic replicons (repFIIA/FIC and repFI) (Picken *et al.*, 1984).

More recently it was shown that a derivative of pCG86 (pRM133), containing repFIIA/FIC and repFIB, replicated under repFIIA/FIC control (Maas *et al.*, 1989). Thus, it was

expected that challenge of pRM133 by an incoming plasmid containing repFIIA/FIC, would lead to a repFIB take over of replication. Rescue of pRM133 replication by repFIB was, however, only partial. This led Maas *et al.* (1989) to postulate that repFIIA/FIC and repFIB interact with each other *in cis*, possibly by steric hindrance at the replication fork. This was most unexpected as these replicons were shown to be fully compatible *in trans* (Maas *et al.*, 1989).

Ambiguities in incompatibility classification of plasmids are also known to occur with plasmids whose basic replicons use anti-sense RNA as the replication control mechanism (Couturier *et al.*, 1988). These plasmids are common among the IncFII, IncI, com9, IncB/O, IncK and IncZ groups. The replicons of plasmids in these groups differ only by a few base changes in the *inc* domain, yet, these changes are sufficient to generate new incompatibility groups.

The presence of a dual replicon in pGSH500 may be advantageous to this multi-drug resistance plasmid by enabling it to use alternative origins. This could avoid incompatibility with related plasmids. The presence of more than one type of plasmid per host cell is quite probable within the hospital flora, as many clinically relevant plasmids transfer efficiently among the *Enterobacteriaceae* and are maintained within these bacterial populations owing to antibiotic selective pressure.

The regulation of replicon usage in these multi-replicon plasmids is poorly understood (Bergquist, 1987; Kline, 1988). In the mini-F plasmids, the primary origin may be silenced when the others become operative (Eichenlaub *et al.*, 1977; Lane, 1981). The mechanism determining this switch is unknown (Kline, 1988). To resolve which replicon is dominant in pGSH500, we assumed that the generalisation of Pritchard & Grover (1981) was valid for this plasmid. These authors proposed that the copy number of a plasmid with more than one origin would be similar to that of a component plasmid with the higher copy number. Evidence in support of this argument has recently been reported by Maas *et al.* (1989) on

the bi-replicon pCG86. Since the copy number of pFDT200 (3-5 per chromosome) corresponded to that of pGSH500, we concluded that in *E. coli* GW125a(*polA1*) repß was the dominant replicon. The criterion chosen to determine replicon dominance in pGSH500 is not infallible considering that the copy number differences between repα and repß are slight, 1-2 versus 3-5 plasmid copies per chromosome, respectivelly. But more complex techniques such as electron microscopy or Fangman 2-D gel electrophoresis systems (see Schvartzman *et al.*, 1990), may still give ambiguous results, which are difficult to interpret (see Martín-Parras *et al.*, 1991; Professor Rawlings, personal communication). Therefore, while assuming that repß was dominant in this instance, the possibility that repα (pFDT100) or another replicon of pGSH500 may be dominant in a different host or under different growth conditions is not excluded (see Nordström, 1990).

Although repα is not the dominant replicon in pGSH500, it is retained in the deletion mutants pGSH510-pGSH540. Retention of repα in these plasmids may simply be a consequence of the mechanism by which the deletions occurred.

A series of deletions of the 3.3 Kb *PstI* fragment of pFDT200 was used to locate the functional regions on the repß. A 2.2 Kb region appears to contain all the elements necessary for replication in *E. coli*. The behaviour of repß closely resembles that of pGSH500. The minimal replication region spans a mere 1.3 Kb, pFDT220204. This plasmid is not as stable as pFDT200 or pGSH500 and has a higher copy number than either of the latter. Further reduction of the size of this replication region leads to plasmid instability.

The domains of repß identified within a 2.2 Kb region of pFDT210 are shown in *Figure 2.7.* These include a 0.25 Kb fragment which is required *in cis* (*oriV*). Downstream from the *oriV*, there is a region which is involved in the control of replication, incompatibility and plasmid copy number. The replication region, *rep* domain, is *trans*-acting. The region containing the incompatibility determinants is also responsible for the

control of copy number and spans approximately 1.0 Kb and is designated the *inc/cop* domain. Involvement of incompatibility determinants with copy number control has been reported for the narrow-host-range plasmids F, P1 and R6K (Tsutsui *et al.*, 1983; Kline & Trawick, 1983; Abeles *et al.*, 1984; Filutowicz *et al.*, 1985) as well as for the broad-host-range plasmids RK2, R1162 and pCU1 (Meyer & Hinds, 1982; Meyer *et al.*, 1985; Kim & Meyer, 1986; Persson & Nordström, 1986; Krishnan & Iyer, 1990).

To further identify and characterise repß, a 2.7 Kb fragment within pFDT210, containing the 2.2 Kb with the essential replication and incompatibility determinants was sequenced. This is described in the next chapter.

CHAPTER 3

MOLECULAR CHARACTERISATION OF repß

CHAPTER 3

MOLECULAR CHARACTERISATION OF repß

SUMMARY

The nucleotide sequence of the 2.7 Kb DNA fragment containing the basic replicon of pGSH500 (repß) was determined from pFDT210. A series of sequential exonuclease III deletions from both ends of pFDT210 ensured that the sequence was obtained from overlapping templates on both strands of the plasmid. The ß replicon was shown to span a 1.8 Kb and includes 3 domains: *ori*V, *repB* and *inc/cop*. The *cis*-acting *ori*V or minimal origin region is located within a 245 bp region (nucleotides 478-723). This region is predominantly AT-rich (90%) and contains two DnaA boxes and two 16 bp direct repeats. The DnaA boxes are essential for replication. Several other direct and inverted repeats capable of forming energetically favourable stem and loop structures are also present in this region. A second domain contains ORF295, which encodes a 30.8 KDa protein (*repB*) that is required for autonomous replication. The third domain consists of eighteen 30-36 bp iterons and was shown to be involved in incompatibility and plasmid copy number control (*inc/cop*). Four of these iterons (1-4) are also essential for stable plasmid replication.

INTRODUCTION

The functional organisation of a replicon can be deduced from analysis of deleted forms of the replicon. The molecular organisation is apparent from sequence analysis of the replicon. Correlation of the functional regions with the nucleotide sequence may lead to the identification of specific sequences involved in replication, incompatibility, control of plasmid copy number and host-range. This approach was followed for F (see Lane, 1981; Murotsu *et al.*, 1981; Tsutsui *et al.*, 1983; Murotsu *et al.*, 1984), P1 (see Sternberg & Austin, 1981; Abeles *et al,* 1984), R6K (Kolter, 1981; McEachern *et al.*, 1986), pSC101 (Vocke & Bastia, 1983), RK2 (Figurski & Helinski, 1979; Stalker *et al.*, 1981; Thomas, 1981; Thomas *et al.*, 1984) and pCU1 (Kozlowski *et al.*, 1987; Krishnan & Iyer, 1990).

The analysis of plasmid primary DNA structure has also revealed binding sites for host-encoded proteins such as DnaA (Rosen *et al.*, 1980; Ryder *et al.*, 1981; Veltkamp & Stuitje, 1981; Abeles *et al.*, 1984; Murotsu *et al.*, 1981; Fuller *et al.*, 1984) and integration host factor (IHF) (Stenzel *et al.*, 1987). In addition, binding sites for plasmid-encoded proteins such as those involved in replication control (Abeles, 1986; Kline, 1988), primosomes or single-strand initiation sites (*ssi*) (Imber *et al.*, 1983; Masai *et al.*, 1990; Nomura *et al.*, 1991) as well as conserved repeated sequences (see Filutowicz *et al.*, 1985; Seelke & Kline, 1984) have also be identified.

The basic replicons of F, P1, RK2 and pSC101 contain ORFs that encode proteins (Rep proteins) which are involved in the initiation and control of plasmid replication (Murotsu *et al.*, 1981; Masson & Ray, 1988; Sϕgaard-Andersen *et al.*, 1984; Chattoraj *et al.*, 1984 and 1985b; Smith & Thomas, 1984; Kornacki *et al.*, 1984; Shingler & Thomas, 1984; Vocke & Bastia, 1983 and 1985; Yamaguchi & Masamune, 1985). Promoters and other regulatory elements of gene expression, as described for other genes (Hawley & McClure, 1983) have also been identified in the basic replicons of these plasmids. Plasmid promoter

sequences often have the potential for secondary structure that may be of significance *in vivo* (McClure, 1985; Collado-Vides *et al.*, 1991; Inouye, 1988; Simons, 1988).

Comparison of amino-acid sequences and codon frequencies of some *rep* genes was carried out by Kües & Stahl (1989). They identified codons which are rare in *E. coli* or in *Ps. aeruginosa*, yet occur commonly in *rep* genes. Codon usage may influence gene expression in different hosts. Codon usage may also disclose evolutionary relationships not immediately apparent at the DNA level (Grantham *et al.*, 1981; Maruyama *et al.*, 1986; Lewin, 1990). The G+C content of the first, second and third codon positions has been used to assess differences in variability or potential for neutral mutations in prokaryotic, viral and vertebrate genes (Muto & Ozawa, 1987; Bernardi & Bernardi, 1986; Sueoka, 1988; Filipski, 1991). A systematic analysis of the relationship between the genomic G+C content and codon usage is, however, lacking in many organisms. In addition, it appears that the genetic code is still evolving (Osawa *et al.*, 1992).

This chapter details the sequence analysis of a 2.7 Kb *Eco*RI-*Hpa*I fragment containing repß. This sequence analysis was compared to the functional analysis described in Chapter 2.

MATERIALS AND METHODS

Bacterial strains and plasmids

The bacterial strains and plasmids used are listed in *Appendix 1 & 2*. Plasmids to be sequenced were maintained in the *E. coli* LKIII.

Enzymes and Chemicals

The sources of restriction enzymes are described in the Materials and Methods of Chapter 2. Sequenase 2.0 kits were acquired from United States Biochemical Corporation (Cleveland, Ohio). [^{35}S]dATP (> 1000 Ci/mmol), [^{35}S]methionine and an *in vitro* prokaryotic DNA directed translation kit were obtained from Amersham International. All reagents were of analytical grade (See also *Appendix 4*).

DNA sequencing strategy and analysis

Plasmid pFDT210 and its exonuclease III derivatives were described in Chapter 2. They were used as templates for DNA sequencing to ensure that both strands of the ß replicon were sequenced (*Figure 3.1.C*). Plasmid DNA was purified using a modified QIAGEN plasmid minipreparation protocol (*Appendix 5*). The sequence of repß was determined by the enzymatic dideoxy chain termination method (Sanger *et al.*, 1977) as described by Biggin *et al.* (1983). This sequencing method was adapted for double stranded DNA templates (Korneluk *et al*, 1985) using Sequenase 2.0 kits and [α-^{35}S]dATP (s.a. > 1000 Ci/mmol). No single strand DNA sequencing was carried as neither the insert of pFDT200 nor that of pFDT210 could be stably maintained in M13 phages. A similar phenomenon was observed during sequencing of repFIB (Saul *et al.*, 1989). The sequence was analysed with the GCG sequence analysis software package (version 7) (Devereux *et al.*, 1984). Genpro (version 5) (Riverside Scientific) was used to identify open reading frames (ORFs).

in vitro protein expression

An *in vitro* prokaryotic DNA directed translation kit (Amersham International) was used to express the open reading frames identified in pFDT210. ^{35}S-methionine labelled products were separated electrophoretically on 12 and 15%, homogeneous, SDS-polyacrylamide,

gels (Laemmli, 1970). The gels were stained with Coomassie blue, dried and autoradiographed at room temperature. The size of the expression products was determined by comparison with protein molecular weight markers from Pharmacia (94 KDa, 67 KDa, 43 KDa, 30 KDa, 20.1 KDa and 14.4 KDa).

RESULTS

Nucleotide sequence of repß

The 2.7 Kb *Eco*RI-*Hpa*I fragment which contained repß (*Figure 3.1 A & B*) was sequenced according to the strategy outlined in *Figure 3.1C*. The 2686 bp sequence of the insert in pFDT210 is shown in *Figure 3.2*. This sequence has a G+C content of 47% and contains 11 sites for each of the restriction enzymes *Hinc*II (GT[T/C][A/C]AC) and *Bst*XI (CCANNNNNNTGG) within a region rich in repeated sequences (nucleotides 1605-2281). Analysis of the repß sequence was compared with the functional analysis described in Chapter 2.

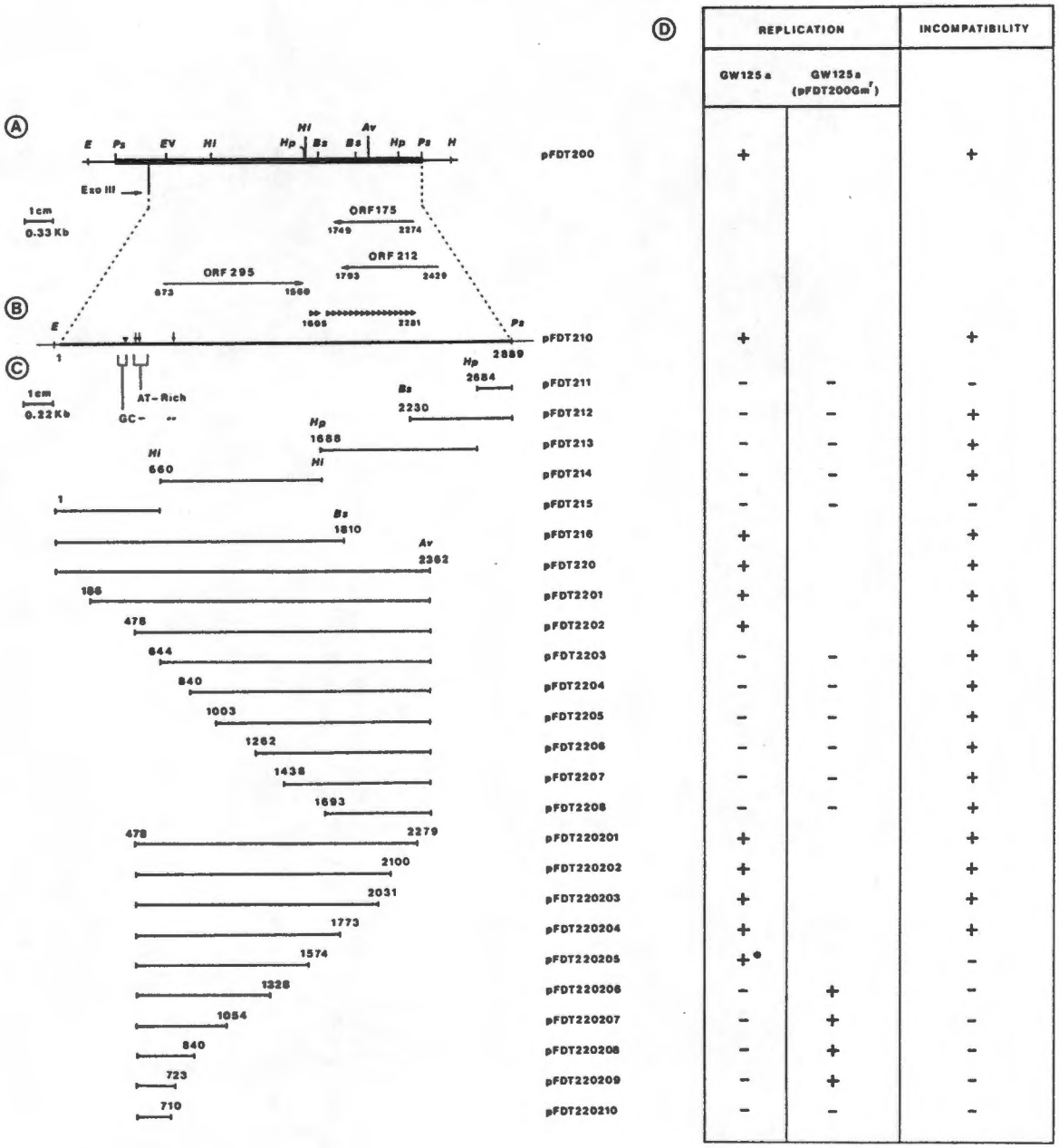


Figure 3.1. Outline of the strategy used during molecular characterisation of repß, to determine the precise location of the *ori*V, *rep* and *inc/cop* domains (see following page for full legend).

Figure 3.1. Outline of the strategy used during molecular characterisation of repß, to determine the precise location of the *ori*V, *rep* and *inc/cop* domains. **A.** Restriction enzyme map of the autonomously replicating 3.3 Kb *Pst*I DNA fragment of pGSH500 (pFDT200). The restriction sites for *Ava*II (*Av*), *Bst*XI (*Bs*), *Eco*RV (*EV*), *Hin*cII (*Hi*), *Hpa*I (*Hp*) and *Pst* (*Ps*) are shown; *Eco*RI (*E*) and *Hin*dIII (*H*) belong to the multiple cloning site of the vector. The region designated ExoIII indicates the portion of DNA which was deleted from pFDT200 (0.4 Kb) to give rise to pFDT210. **B.** Diagrammatic representation of the main features of repß contained within pFDT210. The long arrows (—►) indicate the 3 ORFs predicted from the DNA sequence. Each ORF has been qualified with a suffix, representing the number of codons per ORF. The direction of the arrows indicates the predicted direction of transcription. The numbers below each ORF, represent the first and the last base pair of each ORF. Horizontal filled arrows (►) indicate the position of the 18 iterons of repß. The numbers indicate the position in the sequence of the iteron region within pFDT210. The vertical arrows (↓) indicate the position of the DnaA boxes; the vertical filled arrow (▼) indicates the position of the primosome site. The AT- and GC-rich areas are also indicated. **C.** Outline of the strategy used to sequence and to determine the functional regions of pFDT210. The plasmids pFDT211 to pFDT220 contain subcloned regions of pFDT210, whereas the plasmids pFDT2201 to pFDT220210 are exonuclease III deletions of pFDT220. **D.** Recombinant plasmid constructs of repß that are able to replicate autonomously in *E. coli* GW125a are indicated (+) in the first column; those that can be rescued *in trans* by the mini-replicon pFDT200Gm^r in GW125a are indicated in column two, while those that are incompatible with pFDT200Gm^r are indicated in the third column. * Unstable plasmid.

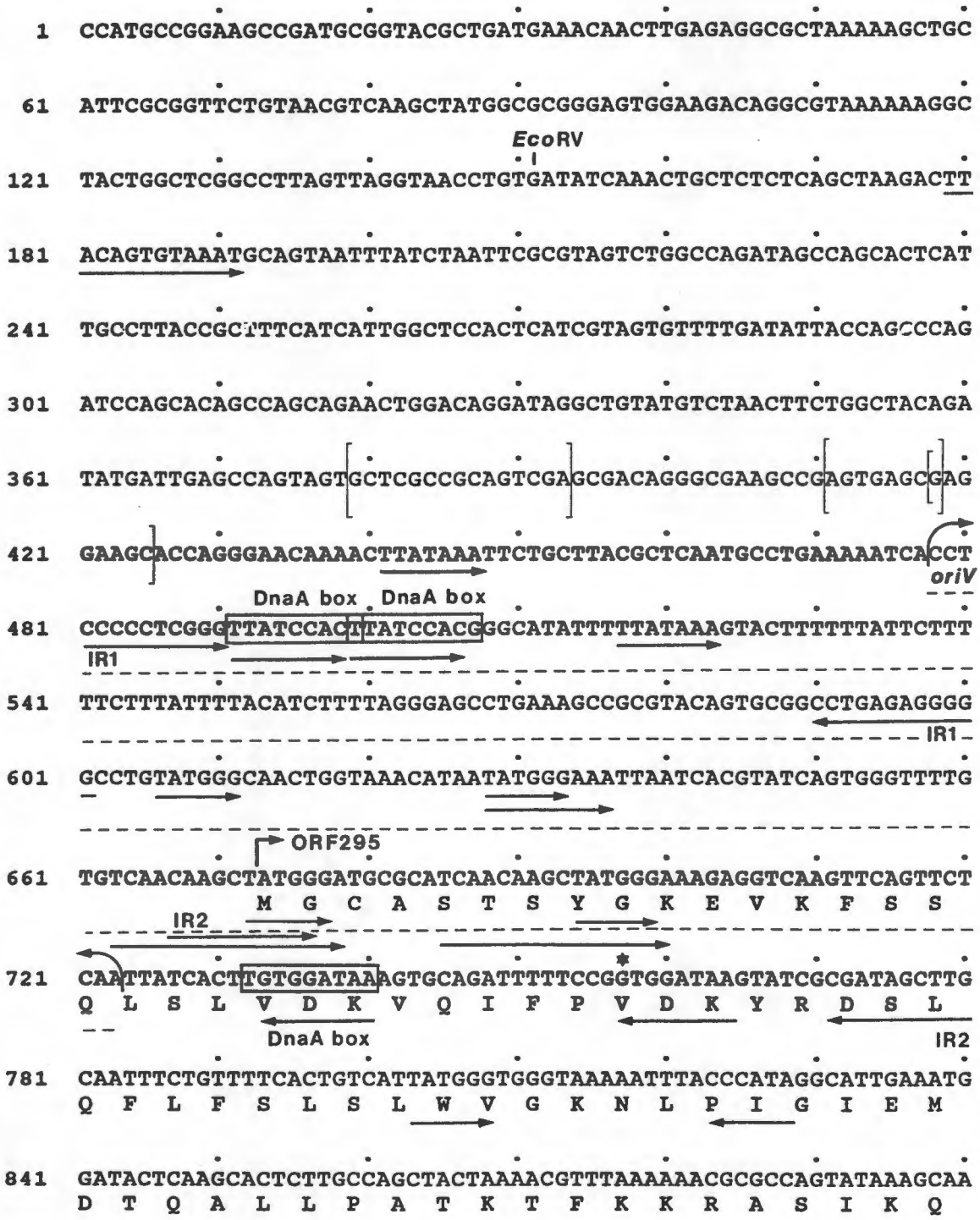
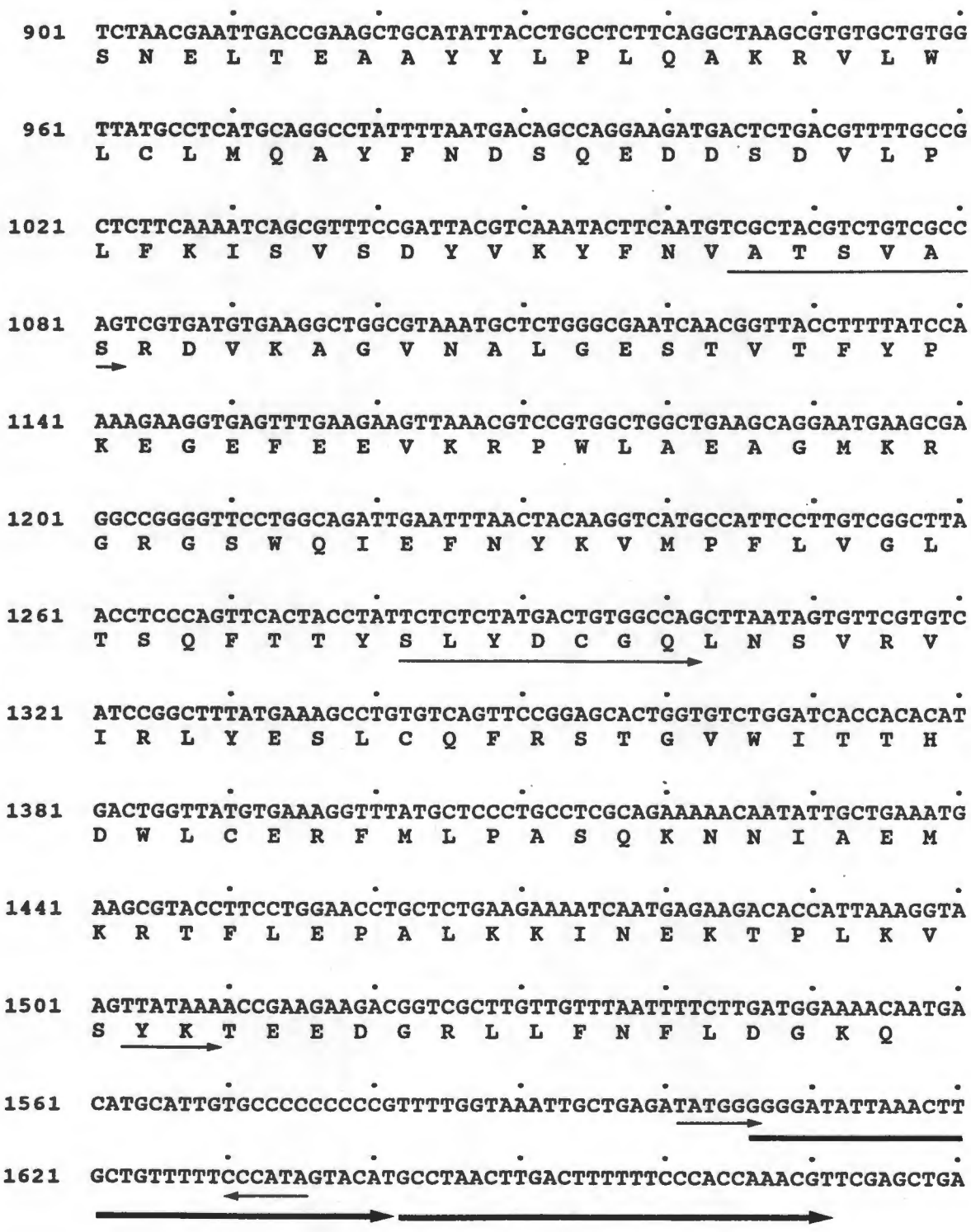


Figure 3.2. Nucleotide sequence of the 2.7 Kb region of pFDT200 containing repß (full legend in following pages).


901 TCTAACGAATTGACCGAAGCTGCATATTACCTGCCTCTTCAGGCTAAGCGTGTGCTGTGG
S N E L T E A A Y Y L P L Q A K R V L W
961 TTATGCCTCATGCAGGCCTATTTTAATGACAGCCAGGAAGATGACTCTGACGTTTTGCCG
L C L M Q A Y F N D S Q E D D S D V L P
1021 CTCTTCAAAATCAGCGTTTCCGATTACGTCAAATACTTCAATGTCGCTACGTCTGTCGCC
L F K I S V S D Y V K Y F N V A T S V A
1081 AGTCGTGATGTGAAGGCTGGCGTAAATGCTCTGGGCGAATCAACGGTTACCTTTTATCCA
S R D V K A G V N A L G E S T V T F Y P
1141 AAAGAAGGTGAGTTTGAAGAAGTTAAACGTCCGTGGCTGGCTGAAGCAGGAATGAAGCGA
K E G E F E E V K R P W L A E A G M K R
1201 GGCCGGGGTTCCTGGCAGATTGAATTTAACTACAAGGTCATGCCATTCCTTGTCGGCTTA
G R G S W Q I E F N Y K V M P F L V G L
1261 ACCTCCCAGTTCACTACCTATTCTCTCTATGACTGTGGCCAGCTTAATAGTGTTCGTGTC
T S Q F T T Y S L Y D C G Q L N S V R V
1321 ATCCGGCTTTATGAAAGCCTGTGTCAGTTCCGGAGCACTGGTGTCTGGATCACCACACAT
I R L Y E S L C Q F R S T G V W I T T H
1381 GACTGGTTATGTGAAAGGTTTATGCTCCCTGCCTCGCAGAAAAACAATATTGCTGAAATG
D W L C E R F M L P A S Q K N N I A E M
1441 AAGCGTACCTTCCTGGAACCTGCTCTGAAGAAAATCAATGAGAAGACACCATTAAAGGTA
K R T F L E P A L K K I N E K T P L K V
1501 AGTTATAAAACCGAAGAAGACGGTCGCTTGTTGTTTAATTTTCTTGATGGAAAACAATGA
S Y K T E E D G R L L F N F L D G K Q
1561 CATGCATTGTGCCCCCCCCCGTTTTGGTAAATTGCTGAGATATGGGGGGATATTAAACTT
1621 GCTGTTTTTCCCATAGTACATGCCTAACTTGACTTTTTTCCCACCAAACGTTCGAGCTGA

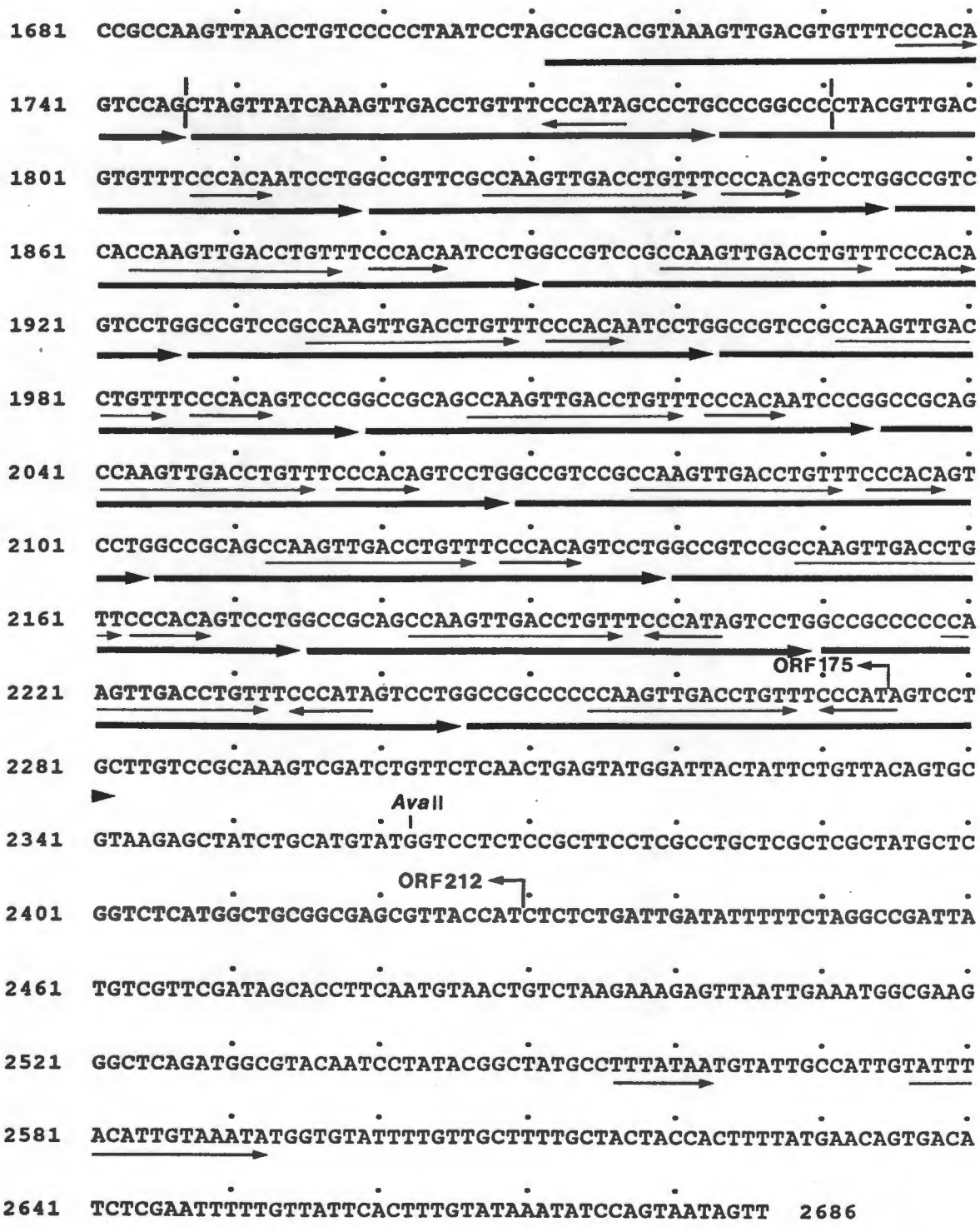

1681 CCGCCAAGTTAACCTGTCCCCCTAATCCTAGCCGCACGTAAAGTTGACGTGTTTCCCACA
1741 GTCCAGCTAGTTATCAAAGTTGACCTGTTTCCCATAGCCCTGCCCGGCCCCTACGTTGAC
1801 GTGTTTCCCACAATCCTGGCCGTTCGCCAAGTTGACCTGTTTCCCACAGTCCTGGCCGTC
1861 CACCAAGTTGACCTGTTTCCCACAATCCTGGCCGTCCGCCAAGTTGACCTGTTTCCCACA
1921 GTCCTGGCCGTCCGCCAAGTTGACCTGTTTCCCACAATCCTGGCCGTCCGCCAAGTTGAC
1981 CTGTTTCCCACAGTCCCGGCCGCAGCCAAGTTGACCTGTTTCCCACAATCCCGGCCGCAG
2041 CCAAGTTGACCTGTTTCCCACAGTCCTGGCCGTCCGCCAAGTTGACCTGTTTCCCACAGT
2101 CCTGGCCGCAGCCAAGTTGACCTGTTTCCCACAGTCCTGGCCGTCCGCCAAGTTGACCTG
2161 TTCCCACAGTCCTGGCCGCAGCCAAGTTGACCTGTTTCCCATAGTCCTGGCCGCCCCCCA
ORF175
2221 AGTTGACCTGTTTCCCATAGTCCTGGCCGCCCCCCAAGTTGACCTGTTTCCCATAGTCCT
2281 GCTTGTCCGCAAAGTCGATCTGTTCTCAACTGAGTATGGATTACTATTCTGTTACAGTGC
AvaII
2341 GTAAGAGCTATCTGCATGTATGGTCCTCTCCGCTTCCTCGCCTGCTCGCTCGCTATGCTC
ORF212
2401 GGTCTCATGGCTGCGGCGAGCGTTACCATCTCTCTGATTGATATTTTTCTAGGCCGATTA
2461 TGTCGTTCGATAGCACCTTCAATGTAACTGTCTAAGAAAGAGTTAATTGAAATGGCGAAG
2521 GGCTCAGATGGCGTACAATCCTATACGGCTATGCCTTTATAATGTATTGCCATTGTATTT
2581 ACATTGTAAATATGGTGTATTTTGTTGCTTTTGCTACTACCACTTTTATGAACAGTGACA
2641 TCTCGAATTTTTGTTATTCACTTTGTATAAATATCCAGTAATAGTT 2686

Figure 3.2. Nucleotide sequence of the 2.7 Kb region of pFDT210 containing repß of pGSH500. The *cis-ori*V (nucleotides 478-723) is indicated within curved arrows and is marked by a dashed line. Several direct and inverted repeated sequences ranging from 6 to 20 bp are underlined (thin arrows underneath sequence). The *rriA*/primosome sites are marked by horizontal square brackets. The DnaA boxes are delineated by rectangles. The beginning of an ORF is indicated by a bent arrow and the end by a vertical bar. The amino acid sequence of the putative polypeptide encoded within ORF295 is shown below the corresponding codons. The 30.8 KDa expressed polypeptide (RepB) corresponds to a product translated from codon 29 indicated by *. The direction of transcription/ translation is indicated by the bent arrow (↱). Thick arrows mark the position of the 18 iterons. These iterons overlap ORF175 and ORF212, from which no polypeptide(s) were expressed *in vitro*. This sequence has been submitted to the EMBL Data Base, Heidelberg, accession no Z11775.

A. Structural features of the oriV

In Chapter 2, the *ori*V was identified to the sequence contained in pFDT220209 (*Figure 2.6C & D*). The features of this 245 bp sequence (nucleotides 478-723; *Figure 3.1C., 3.1D, 3.2. & 3.3*) are described below.

An AT-rich sequence (90%) occurs between nucleotides 491 and 560. This unique sequence includes 2 overlapping DnaA boxes corresponding to the 9 bp consensus sequence of TTATCCACT and TTATCCACG (nucleotides 491-507).

There are several repeated sequences within the ß *ori*V. Degenerate inverted repeats (IR1/1', *Figure 3.2.*) flank the 90% AT-rich sequence, are present between nucleotides 481-490 (CCCCCTCGGG) and 590-601 (CCTGAGAGGGGG); an additional set of IRs (IR2/2') is found downstream from IR1/1' and is partially outside the *ori*V (IR2 at nucleotides 667-676, CAAGCTATGG; IR2' at nucleotides 771-780, CGATAGCTTG). IR1/1' and IR2/2' along with their intervening sequences are capable of forming stem and loop structures (ΔG=-28.9 Kcal/mol and ΔG=-28.5 Kcal/mol, respectively) (*Figure 3.4.A & B*). The degree of secondary structure is further illustrated in *Figure 3.4.C.*

Two identical 16 bp direct repeats (TCAACAAGCTATGGGA) occur at nucleotides 663 to 700. In addition, several 7 bp (TTATAAA) and 6 bp (TTATCC; TATGGG) direct repeats are also present in this region.

Upstream from the *ori*V (nucleotides 379-425; *Figure 3.2*) there are sequences corresponding to the consensus sequence for the primosome sites of ColE1 and F, which are designated *rriA* sites (GCTCGCCGCAGTCGA, AGTGAGCG, GAGGAAGC) (Nomura *et al.*, 1982; Marians *et al.*, 1982; Imber *et al.*, 1983). These are also referred to

as single-strand initiation sequences (*ssi*) (Nomura *et al.*, 1991) and suggest that as with the *ori-2* of F, replication in repß proceeds unidirectionally to the left. For the significance of *ssi/rriA* sites near the *ori*V regions see Chapter 4.

Downstream from the *ori*V, there is a third DnaA box (nucleotides 732-740) in the opposite orientation to those described earlier. The DnaA boxes between nucleotides 491-507 and associated sequences were shown to be essential, as their deletion led to loss of autonomous replication (pFDT2203). The role of the third DnaA box is unknown.

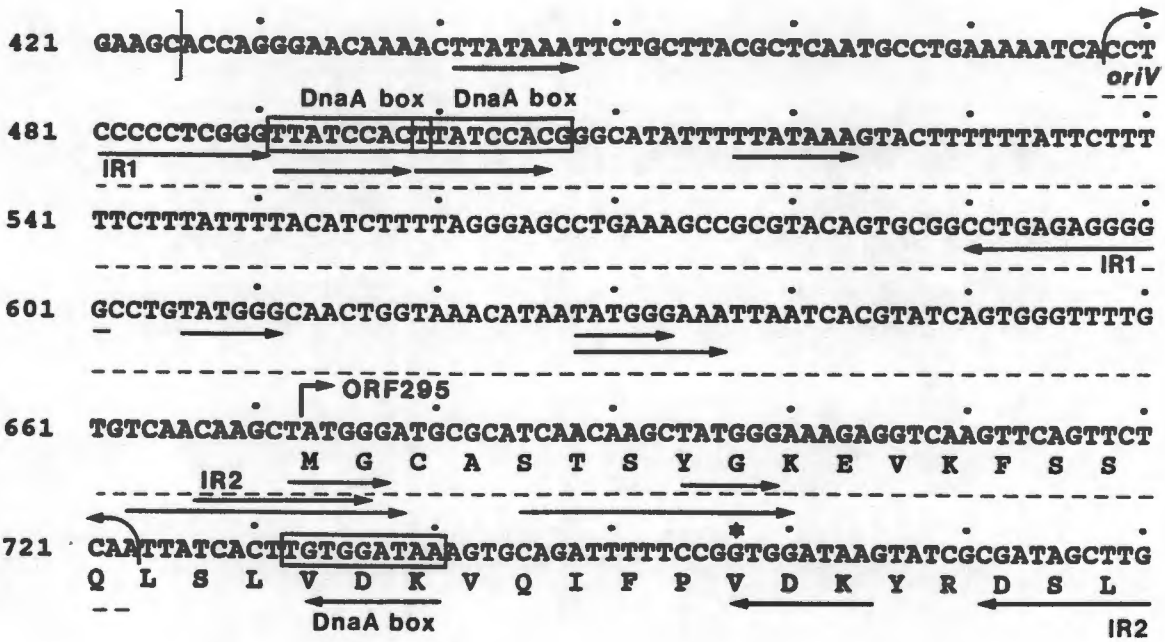


Figure 3.3. Nucleotide sequence of the *cis-ori*V of repß (between curved brackets, dashed line). The position of the DnaA boxes, repeated sequences and inverted repeats (IR1/IR1'; IR2/IR2') are indicated (see ***Figure 3.2***). The sequence between nucleotides 491 and 560 is 90% AT-rich.

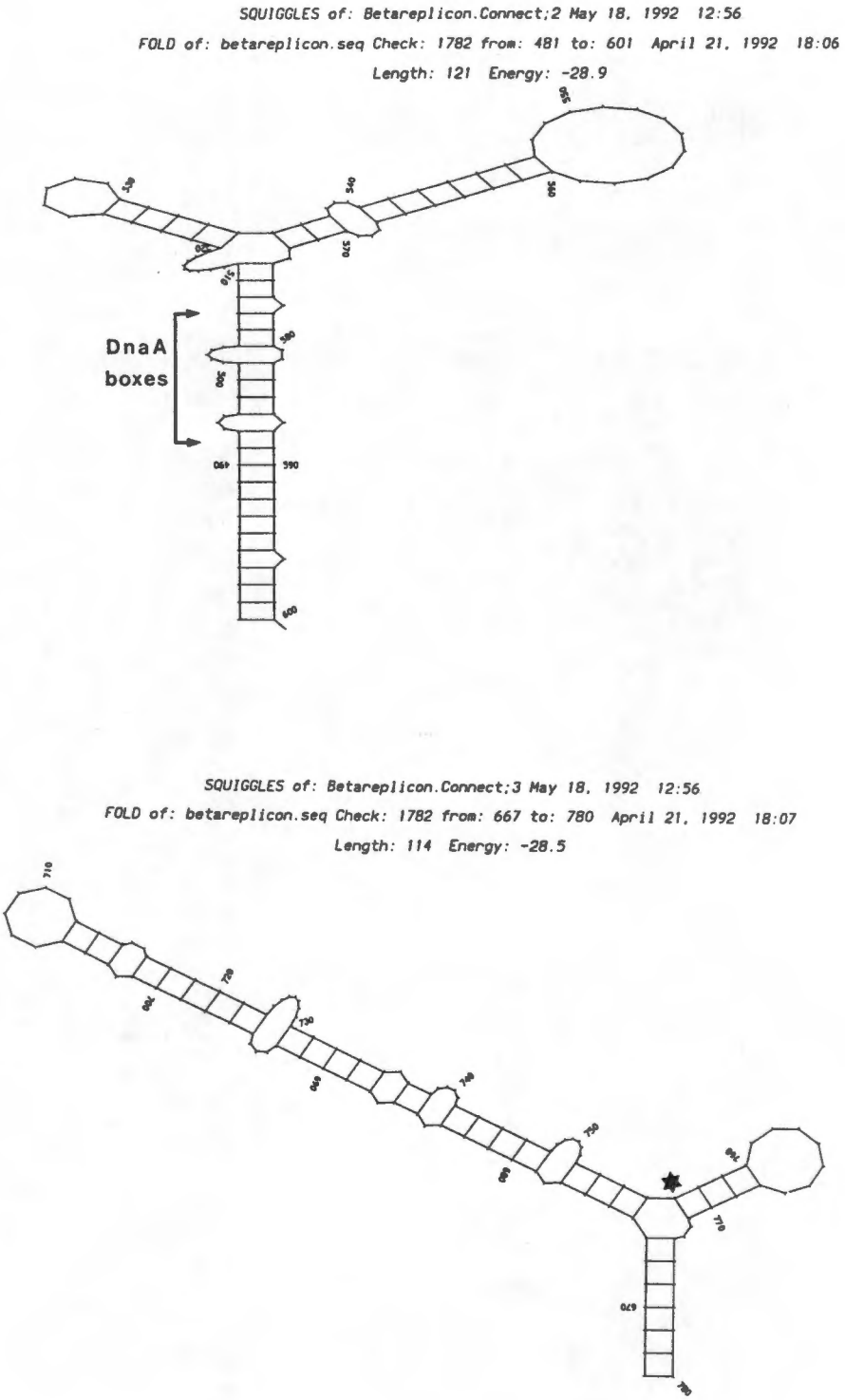


Figure 3.4. Secondary structure of the region (nucleotides 450-780) containing the *cis-ori*V (nucleotides 478-723). Stem and loop structures capable of forming between IR1/IR1'(nucleotides 481-490/590-601) (**A**); between IR2/IR2' (nucleotides 667-676/771-780) (**B**); and between nucleotides 450 and 780 (**C**). The position of the DnaA boxes in (**A**) is marked. The position of the putative start of *repB* (*) is indicated (**B**).

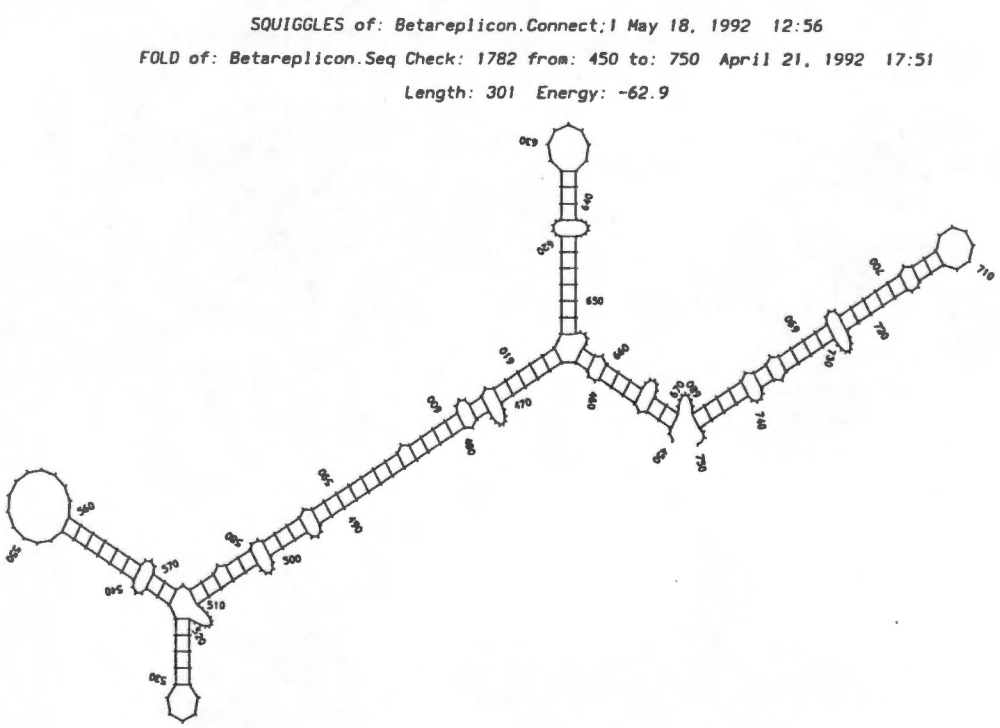


Figure 3.4. Secondary structure within the *cis-ori*V of repß (legend in previous page).

B. The *rep* domain

The functional analysis described in Chapter 2. demonstrated that the plasmid pFDT220204 contains the minimal autonomously replicating region of repß. This insert corresponds to the sequence between nucleotides 478-1773 and includes the *ori*V, an open reading frame (ORF295) as well as four 30-36 bp iterons (see later below) (*Figure 3.2)*. These iterons and ORF295 are essential for origin function, as their deletion (pFDT220205-pFDT220209, *Figure 3.1D*) led to either unstable plasmids or complete loss of autonomous replication function (Chapter 2.). The properties of ORF295 are described later (this Chapter).

C. The iteron containing domain

The complementation analysis described in Chapter 2 and is summarised in *Figure 3.1C & D*, suggests that the incompatibility and copy number determinants of repß are associated with the overlapping regions of pFDT213, pFDT216 and pFDT220. The common fragment to these plasmids corresponds to nucleotides 660 to 2364 (*Figure 3.1C & 3.2*).

The most striking feature of the sequence in this region is the presence of 16 tandem direct repeats (iterons) separated from 2 similar iterons by a 40 bp gap (nucleotides 1605 to 2281). These 18 iterons are located downstream to ORF295 and their nucleotide sequence and organisation is shown in *Figure 3.5*. Iteron 1 is closest to ORF295, whereas iteron 18 is the furthest from this ORF. Except for iteron 2, which is 30 bp long, all of the other 17 iterons are 34 to 36 bp. Iterons 17 and 18 are identical to each other (*Figure 3.5.*). In contrast, iterons 1 and 2 are the most degenerate. The degree of sequence degeneracy decreases from iteron 1 towards iteron 17 (*Figure 3.2*). Within each repeat there are discrete AT-rich nucleotide sequences interspersed with GC-rich regions. Iterons 6 to 18 contain a 15 bp conserved sequence, CCAAGTTGACCTGTT. The pentanucleotide CCTGT, which is a subset of this 15 bp conserved sequence, occurs 14 times within the

iteron containing region (nucleotides 1760 to 2270) and 7 times between nucleotides 100 and 1700. A family of 6 bp conserved sequence is present in iterons 3 and 5 to 14 (CCCACA). Two ORFs, ORF175 and ORF212 overlap this iteron containing region (*Figure 3.1 & 3.2*) (see later, this Chapter).

```
1)    GGGGATATTAAACTTGCTGTTTTTCCCATAGTACAT
2)          GCCTAACTTGACTTTTTTCCCACCAAACGT
3)    GCCGCACGTAAAGTTGACGTGTTTCCCACAGTCCAG
4)    CTAGTTATCAAAGTTGACCTGTTTCCCATAGCCCTG
5)    CCCGGCCCCTACGTTGACGTGTTTCCCACAATCCTG
6)    GCCGTTCGCCAAGTTGACCTGTTTCCCACAGTCCTG
7)    GCCGTCCACCAAGTTGACCTGTTTCCCACAATCCTG
8)    GCCGTCCGCCAAGTTGACCTGTTTCCCACAGTCCTG
9)    GCCGTCCGCCAAGTTGACCTGTTTCCCACAATCCTG
10)   GCCGTCCGCCAAGTTGACCTGTTTCCCACAGTCCCG
11)   GCCGCA_GCCAAGTTGACCTGTTTCCCACAATCCCG
12)   GCCGCA_GCCAAGTTGACCTGTTTCCCACAGTCCTG
13)   GCCGTCCGCCAAGTTGACCTGTTTCCCACAGTCCTG
14)   GCCGCA_GCCAAGTTGACCTGTTTCCCACAGTCCTG
15)   GCCGTCCGCCAAGTTGACCTGTT_CCCACAGTCCTG
16)   GCCGCA_GCCAAGTTGACCTGTTTCCCATAGTCCTG
17)   GCCGCCCCCCAAGTTGACCTGTTTCCCATAGTCCTG
18)   GCCGCCCCCCAAGTTGACCTGTTTCCCATAGTCCTG

CONCENSUS:    GCCGNNCGCCAAGTTGACCTGTTTCCCACAGTCCTG
```

Figure 3.5. Sequence alignment of the 18 iterons in repß. Iterons 1 and 2 are the most degenerate and proximal to the end of ORF295. In contrast, iterons 17 and 18 are identical to each other. The sequence degeneracy decreases from iteron 1 to iteron 18. AT-rich regions are interspersed with GC-rich regions. The consensus sequence is 36 bp long.

C.1. Association of the iteron containing region with the incompatibility locus (inc). Sequence analysis showed that the region associated with the incompatibility properties of repß contained the 18 iterons described above. To assess if incompatibility was related to the number of iterons present, plasmids containing different sectors of the iteron region (pFDT212-pFDT216, pFDT2208-pFDT220204), were tested in DK-1(pFDT200Gm^r) (*Figure 3.1C & D*). None of these plasmids was compatible with pFDT200Gm^r, in fact, the incompatibility of pFDT214 which contains only iterons 1 and 2 was similar to that of other plasmids containing all or most of the iterons. These results show that there is no association between the number of iterons present and plasmid incompatibility.

C.2. The role of iterons in the plasmid copy number control locus (cop). In the previous Chapter it was demonstrated that deletions on the right side of pFDT220 (*Figure 2.6)* led to a family of plasmids with raised copy number (*Table 2.1*), suggesting that copy number control elements were deleted in these smaller plasmids (pFDT220203-pFDT220204, *Figure 3.1*).

When these data were related to the sequence of pFDT210 it was found that deletion of iterons 14 to 18 (nucleotides 2100-2281; *Figure 3.1 & 3.5*) was not associated with an alteration in the copy number (pFDT220202; *Table 2.1*). This implied that the 13 remaining iterons in pFDT220202 are sufficient to regulate the plasmid copy number at the same level as that of wild type. In fact, the progressive deletion of iterons 5-13 (1774-2104) was associated with a sequential increase in copy number of pFDT220203 and pFDT220204 (up to 7 fold in the latter). Deletion of the remaining 4 iterons in pFDT220204 led to the "unstable" plasmid pFDT220205. When this plasmid was transformed into GW125a(*polA1*), the colonies exhibited a crenated morphology, indicating that sectors of the colony were dying. The reason for this is not clear.

D. Other features of the repß sequence

Other sequences, whose role is unknown at present, occur within the 2.7 Kb DNA fragment containing repß. These include: three degenerate, iteron-like direct repeats (19-23 bp long) with internal palindromic sequences can be found at nucleotide positions 378-420 (TGCTGCCCGCAGTCGAGCGACAG), nucleotides 1064-1082 (TCGCTACGTCTGTCGCC) and nucleotides 1282-1302 (TCTCTCTATGACTGTGGCCAG). Two other degenerate direct repeats, are present between nucleotides 179-191 (TTACAGTGTAAA) and nucleotides 2578-2590 (TTTACATTGTAAA). The direct repeat TTACAGTGTAAA contains an internal palindrome. The 7 bp and 6 bp repeats mentioned above also occur in other regions of the replicon.

Recognition sites for IHF (YAA----TTGATW; Leong *et al.*, 1985; Gamas *et al.*, 1986) or for Dam (GATC; Hattman *et al,* 1978) were not found.

ORFs encoded by repß

There are 3 major open reading frames contained within this 2.7 Kb sequence: ORF295 (nucleotides 673 to 1560), ORF212 (nucleotides 1793 to 2429) and ORF175 (nucleotides 1749 to 2274). ORF295 has several in frame transcription initiation sites (ATG, GTG, TTG) between nucleotides 658 and 840. ORF175 and ORF212 overlap each other as well as the 18 iterons and are in the opposite orientation to ORF295 (*Figure 3.1, 3.2 & 3.6*).

The sequences immediately prior the putative transcription initiation sites of ORF175, ORF212 and ORF295 show poor homology with the consensus recognition sequences of the *E. coli* δ^{70}, δ^{54} and δ^{32} factors (Hawley and McClure, 1983; McClure, 1985). In fact

the putative promoter of ORF295 corresponding to the transcription initiation site at nucleotide 838 (ATG), does not have sequence consensus to the -35 hexamer of *E. coli* (*Figure 3.7*). In addition, the spacing between the -35 and -10 hexamers is not optimal (mostly less than 17). Some of the best matches of these putative promoters to the consensus sequences described by Rosenberg & Court (1979), Hawley and McClure (1983) and McClure (1985), are shown in *Figure 3.7.*

Apart from poor putative promoters, the ORFs identified were preceded by sequences showing poor homology to the consensus ribosome binding site (RBS) of *E. coli* (Shine & Dalgarno, 1974 and 1975) (*Table 3.1*). Computer analysis of all three ORFs predicted that if translation occurred, polypeptides ranging in size from 18.8 to 33.8 KDa would be expressed. In addition, since several translation initiation sites were recognised for ORF295, expression products from this ORF were expected to range from 27.4 to 33.8 KDa (*Table 3.1*).

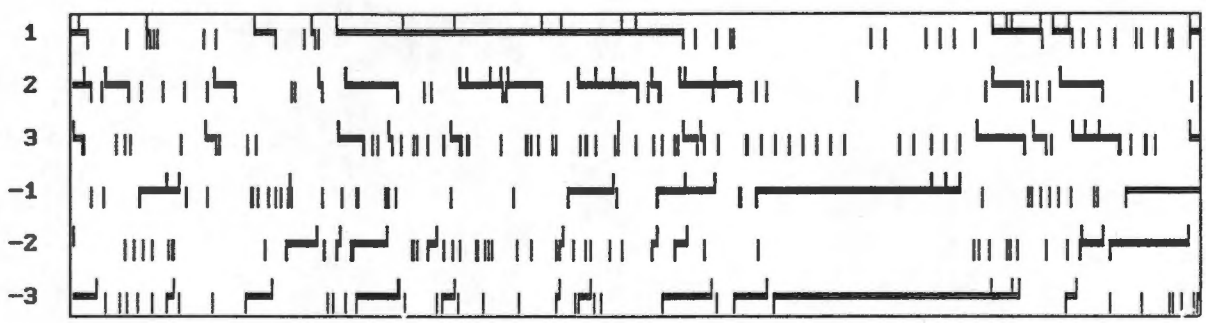


Figure 3.6. ORFs identified within the 2.7 Kb sequence (pFDT210) containing repß. ORF295 is the major ORF within the first reading frame of the sequence (1), whereas ORF175 and ORF212 correspond to the major ORFs within the first (-1) and third (-3) reading frames in the opposite direction, respectively.

	-35		-10		*E. coli* consensus
	tcTTGACat	≈N17	**t tg TAtAaT**		
ORF175:					
	ACTCAG**TTGAGA**	N5	**TCGACT**	N17	ATG
	****** ***		*** * ***		2274
			2291		
ORF212:					
	TTTCAA**TTAACT**	N14	**TAGATT**	N54	ATG
	**** ****		**** * ***		2429
			2483		
ORF295:					
	GGGGCC**TGTATG**	N15	**TAATAT**	N43	ATG
	*** ***		**** ****		673
			630		
	GTCAAG**TTCAGT**	N20	**TAAAGT**	N13	GTG
	**** ***		**** * ***		757
			744		
	AGCTTGCAATTTCTGTTTTAC**TGTCAT**			N36	ATG
			*** * ****		838
			802		
RepE:					
	TTGACT	N16	**GACAAT**	N34	ATG
	*********		*** *****		304
			270		

Figure 3.7. Comparison between the putative promoters of the ORFs identified in repß and that of RepE of mini-F (Murotsu *et al.*, 1984; Sϕgaard-Andersen *et al.*, 1984). The consensus sequences for the *E. coli* -35 and -10 regions are also shown (Hawley & McClure, 1983). Numbers under sequence correspond to the nucleotide position of beginning of ORF's (ATG/GTG) and putative promoter sequences.

Table 3.1. Putative polypeptides that may be expressed from the open reading frames identified within repß.

ORF	Position	Translational initiation site	Stop codon	No aas	M_r (KDa)
175	1749-2274	UUUGCGGACAAGCAGGACUA**UG**	UGA	175	18.8
212	1793-2429	AAAAAUAUCAAUCAGAGAGA**UG**	UAG	212	23.4
295	673-1560	GGUUUUGUGUCAACAAGCUA**UG**	UGA	295	33.8
	757-1560	UAAAGUGCAGAUUUUUCCGG**UG**	UGA	267	30.6
	840-1560	UUACCCAUAAGGCAUUGAAA**UG**	UGA	240	27.4

Partial ribosomal binding sites are underlined. Consensus ribosomal binding sites as proposed for *E. coli* by Shine & Dalgarno (1974, 1975) and Gold *et al.*, (1981): AGGA, UGGA, GGAG. Initiation codons (**AUG, GUG**) are indicated in bold.

Analysis of proteins encoded by repß

A. *in vitro* expression of polypeptides

Expression of the putative products of ORF175, ORF212 and ORF295 (*Figure 3.1*, *Table 3.1*) was investigated using a DNA directed translation system (Amersham). The expressed products were analysed by SDS-polyacrylamide gel electrophoresis (*Figure 3.8A & B*). The expression products from pUC19 were compared to those of pFDT213 (ORF175 and ORF212) (*Figure 3.8A*, Lanes 2 & 4). No additional expression products corresponding to those predicted for ORF175 and ORF212 (18.8 and 23.4 KDa, respectively) were obtained from pFDT213. But, the possibility that products of ORF175 and ORF212 were obscured by those expressed from pUC19 had to be excluded. Thus, the plasmids pFDT213, pFDT210 and pUC19 were restricted with enzyme *BglI*. This enzyme does not restrict either the *lacI* region of the vector or the repß sequence. In addition, pFDT213 was restricted with *Hinc*II (negative control). Since there are 10 *Hinc*II sites in this region of repß, no expression products were expected from this plasmid.

No expression products were obtained from *Bgl*I restricted pUC19 and pFDT213 (*Figure 3.8A*, Lanes 3 and 5). From the *Hinc*II restricted pFDT213, the only polypeptides expressed were those corresponding to vector ß-lactamase (*Figure 3.8A*, Lane 6). Also no expression products corresponding to those predicted from ORF175 and ORF212 were expressed from *Bgl*I restricted pFDT210 (*Figure 3.8A*, Lane 7). Yet, two polypeptides of approximately 30 KDa were poorly expressed from *Bgl*I restricted pFDT210. These polypeptides correspond to those of ORF295 (see below; *Figure 3.8B*).

A 30.8 KDa protein was the major expression product of ORF295 (pFDT220204), although a 32 KDa polypeptide was also observed, albeit at a lower concentration (*Figure 3.8B,* Lane 2). These polypeptides become visible only after disruption of the ß-lactamase gene of pUC19. This was achieved by restriction of the pUC19 derived plasmids

with *Ava*II. Expression of the 30.8 KDa protein is consistent with translation from one of the alternative initiation sites of ORF295, corresponding to a protein of 267 amino acids (starting at nucleotide 757, codon 29, *Table 3.1*). It is possible that the minor expression product (32 KDa polypeptide) results also from sporadic initiation at one of the other in frame initiation sites (GTG or TTG, between nucleotides 673 to 838, see nucleotide sequence of repß, *Figure 3.2*).

To confirm that the expression of these products from ORF295 was not due to premature termination, two plasmids with progressive deletions of the 3'end of ORF295 were also tested in this system. Smaller polypeptides, 28 KDa (pFDT220206) and 16 KDa (pFDT220207) were obtained (*Figure 3.8B*, *lanes 5* and *7*). The size of the polypeptides expressed from these plasmids was consistent with those predicted for the shortened ORF295 in each plasmid. In addition, the 28 KDa and 16 KDa polypeptides were accompanied by two other polypeptides (29 KDa and 17 KDa, respectively). It was concluded that the 30.8 KDa protein is the major product of ORF295. Since this protein appears to be required for autonomous replication, it was designated RepB.

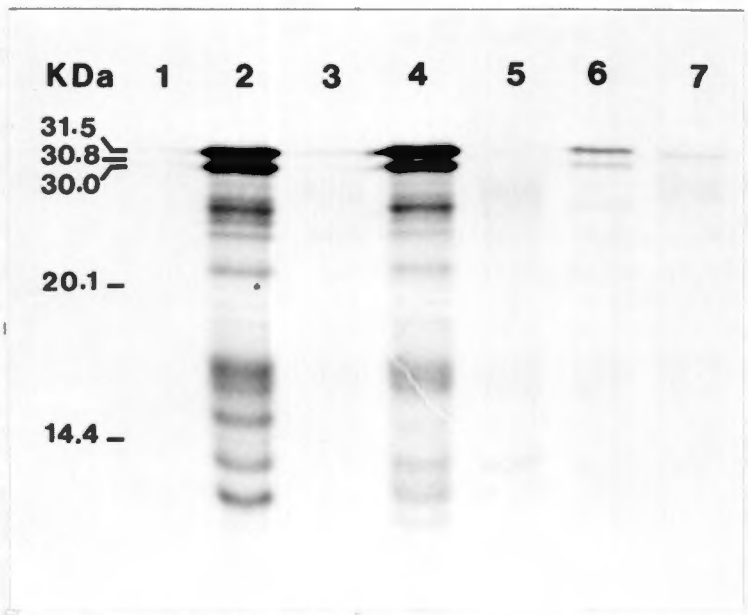


Figure 3.8A. Autoradiograph of ^{35}S-methionine labelled repß polypeptides translated *in vitro* and separated by SDS-PAGE. **A.** Polypeptides expressed from pFDT213 and pFDT210. Lanes 1: non radio-active protein molecular weight markers. Lane 2: pUC19, unrestricted; Lane 3: pUC19, *Bgl*I digested; Lane 4: pFDT213, unrestricted; Lane 5: pFDT213, *Bgl*I digested; Lane 6: pFDT213, *Hinc*II digested; Lane 7: pFDT210, *Bgl*I digested. Restriction of pUC19 derived plasmids with *Ava*II or *Bgl*I disrupts the ß-lactamase gene. The polypeptides were separated on a homogeneous 15% SDS-polyacrylamide gel.

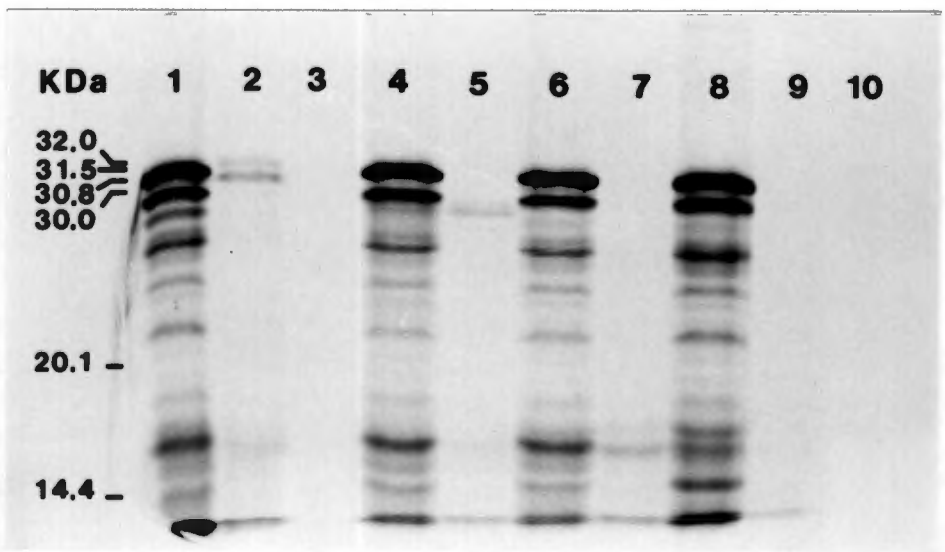


Figure 3.8B. Autoradiograph of ^{35}S-methionine labelled repß polypeptides translated *in vitro* and separated by SDS-PAGE. **B.** Polypeptides expressed from intact and deleted forms of pFDT220204. Lane 1: pFDT220204, unrestricted; Lane 2: pFDT220204, *Ava*II digested; Lane 4: pFDT220206, unrestricted; Lane 5: pFDT220206, *Ava*II digested; Lane 6: pFDT220207, unrestricted; Lane 7: pFDT220207, *Ava*II digested; Lane 8: pUC19, unrestricted; Lane 9: pUC19, *Ava*II digested; Lanes 3 and 10: non radio-active protein molecular weight markers. Restriction of the pUC19 derived plasmids with *Ava*II disrupts the ß-lactamase gene. The polypeptides were separated on a homogeneous 12% SDS-polyacrylamide gel.

B. Codon usage of the ORFs in repß

The codon usage (frequencies per thousand) of the three ORFs identified in the sequence of repß is shown in *Table 3.2A*. Included in this table is also the codon usage for *repE* of mini-F (Murotsu *et al*, 1981; Murotsu *et al.*, 1984) and that of ORF239 of the plasmid pCU1 (Krishnan & Iyer, 1990; Krishnan *et al,* 1990). *Table 3.2B* shows an analysis of the G+C content of each of the three nucleotide positions in the codons of the ORFs described in *Table 3.2A*. Codon usage tables for genes from *E. coli*, *K. pneumoniae*, *S. marcescens* and *Salmonella* (species not given) are shown in *Table 3.3*. The average codon frequency of 13 highly expressed (H13) and 16 weakly expressed *E. coli* genes (W16) (Grantham *et al.*, 1981) as well as that of the *nifh* gene of *K. pneumoniae* (Maruyama *et al.*, 1986) are also included in *Table 3.3*.

ORF175 and ORF212 show a similar bias for codon selection. Both ORFs show a bias for CUG and UUG of Leu, for all 4 codons of Pro, for UCC and AGU for Ser and for the rarely used codons. In contrast, the codon bias of ORF295 differs from that of ORF175 and ORF212, being similar to that of *repE* and ORF239. Of the ORFs of repß, only ORF295 shows similar bias for codon selection to that of genes in several bacterial species, especially to the weakly expressed genes in *E. coli* (*Eco* W16, *Table 3.3*).

The average G+C content for ORF175 and ORF212 is similar (59.81 and 57.03%, respectively) and greater than that of ORF295 (44.14%). In fact the G+C content of ORF295 is similar to that of *repE* and ORF239 (*Table 3.2B*). In ORF175, the highest G+C ratio occurs at the first codon position (74.12%). In ORF212, it is the second codon position that exhibits G+C preference. In contrast, ORF295 has a lower G+C ratio in the second position. This is true also for *repE* and ORF239.

Table 3.2A. Codon usage(*) of ORF175, ORF212 and ORF295 encoded by the ß mini-replicon as well as *repE* (mini-F) and ORF239 (pCU1).

Amino Acid	Codon	ORF175	ORF212	ORF295	*repE*	ORF239
Ala	GCG	0	0	0	24	9
	GCA	6	15	14	8	9
	GCU	0	10	31	8	50
	GCC	46	24	14	12	0
Arg	AGG**	0	0	4	0	0
	AGA**	0	0	0	12	0
	CGG**	6	5	10	16	9
	CGA**	0	0	4	4	13
	CGU	6	42	17	28	17
	CGC	46	14	10	28	17
Asn	AAU	0	14	27	16	17
	AAC	0	0	10	16	25
Asp	GAU	0	5	27	28	42
	GAC	51	5	20	20	17
Cys	UGU	6	28	10	8	9
	UGC	0	10	7	4	4
Gln	CAG	17	14	30	44	34
	CAA	46	10	17	8	17
Glu	GAG	0	5	10	24	25
	GAA	0	0	54	44	34
Gly	GGG**	0	0	4	12	4
	GGA	0	0	10	8	4
	GGU	0	0	17	4	9
	GGC	29	42	20	24	21
His	CAU	12	10	4	16	9
	CAC	17	33	0	8	0
Ile	AUA**	6	0	7	12	13
	AUU	0	0	14	8	17
	AUC	23	10	14	69	21
Leu	CUG	85	24	31	24	17
	CUA**	6	5	0	4	9
	CUU	0	5	20	8	34
	CUC	0	5	17	16	25
	UUG	6	28	20	8	17
	UUA	0	0	24	8	21
Lys	AAG	6	33	37	24	42
	AAA	6	0	48	44	46
Met	AUG	6	5	24	24	13
Phe	UUU	46	5	37	28	25
	UUC	12	24	24	28	29
Pro	CCG	23	47	10	12	9
	CCA	40	61	14	12	4
	CCU	51	66	10	16	9
	CCC**	68	34	4	4	9
Ser	UCG**	0	10	10	0	9
	UCA	0	0	17	12	21
	UCU	0	19	17	16	21
	UCC	40	71	10	29	4
	AGU	52	57	17	28	17
	AGC	6	10	20	8	17
Thr	ACG	0	0	10	12	13
	ACA	40	14	10	12	13
	ACU	0	0	13	16	13
	ACC	6	24	24	12	13
Trp	UGG	6	24	20	8	21
Tyr	UAU	0	14	31	32	17
	UAC	6	10	14	20	25
Val	GUG	6	5	20	4	9
	GUA	0	5	7	8	21
	GUU	91	47	17	20	17
	GUC	51	5	27	12	9
Ter	UGA	0	0	4	0	4
	UAG	6	5	0	4	0
	UAA	0	0	0	0	0

* Codon usage defined as frequency of codon occurance per thousand codons.

** indicates rarely used codons in *E. coli*.

Table 3.2B. A comparinson of % G+C content of the nucleotides at three positions in the codons used in ORF175, ORF212, ORF295, *repE* (mini-F) and ORF239 (pCU1).

	ORFs/Genes				
Positions	ORF175	ORF212	ORF295	*repE*	ORF239
1st	74.12	54.02	48.64	50.40	50.00
2nd	50.58	63.98	37.16	38.49	36.25
3rd	54.71	53.08	46.62	52.78	45.83
Average G+C content of ORF	59.81	57.03	44.14	47.22	44.03

Table 3.3. Frequency per thousand for each of the 61 codons in some *E. coli*, *S. marcescens* and *Salmonella* (species not given) genes (adapted from Grantham *et al.*, 1981).; *K. pneumoniae* gene (adapted from Maruyama *et al*, 1986).

		Species/gene								
	Codon	*Eco s12*	*Eco recA*	*Eco ompA*	*Eco fol*	*Sma lpp*	*Sal hin*	*Eco* H13	*Eco* W16	*Kle nif*
Ala	GCG	33	54	9	44	23	32	21	26	13
	GCA	16	31	32	6	39	21	43	21	1
	GCU	24	9	64	6	105	21	65	16	1
	GCC	0	11	3	25	0	21	13	27	20
Arg	AGG**	0	0	0	0	0	0	0	4	0
	AGA**	0	0	0	0	0	16	1	9	0
	CGG**	0	0	0	13	0	26	0	8	0
	CGA**	0	0	0	0	0	26	0	0	0
	CGU	106	34	29	6	39	21	44	18	4
	CGC	16	6	9	38	13	16	17	24	9
Asn	AAU	0	3	3	13	13	32	3	18	1
	AAC	41	40	52	25	79	11	39	15	11
Asp	GAU	0	26	17	51	26	21	16	33	5
	GAC	24	31	46	32	79	21	36	22	11
Cys	UGU	8	6	3	6	0	5	1	8	2
	UGC	24	3	3	6	13	0	5	7	7
Gln	CAG	24	34	43	13	39	21	28	29	8
	CAA	8	3	6	13	39	16	8	18	2
Glu	GAG	8	26	12	19	0	21	10	23	11
	GAA	24	60	26	57	0	32	39	33	18
Gly	GGG**	0	3	0	0	0	16	1	11	1
	GGA	0	3	0	6	0	21	2	7	1
	GGU	0	48	70	25	13	5	45	19	5
	GGC	33	45	41	32	26	32	32	27	20
His	CAU	0	0	6	19	0	26	10	20	0
	CAC	24	6	9	13	26	5	8	11	2
Ile	AUA**	0	0	0	0	0	11	2	8	0
	AUU	0	6	3	32	0	53	17	30	7
	AUC	24	71	43	44	26	21	42	22	17
Leu	CUG	49	68	61	32	79	32	58	39	11
	CUA	0	3	0	0	0	5	0	5	0
	CUU	8	9	0	0	13	16	3	15	1
	CUC	8	6	0	13	0	0	2	12	7
	UUG	0	6	3	13	0	0	3	10	0
	UUA	0	0	3	13	0	53	3	14	0
Lys	AAG	33	17	12	6	13	11	19	15	5
	AAA	73	60	43	32	66	42	62	31	11
Met	AUG	0	26	14	25	13	21	19	25	16
Phe	UUU	0	11	6	13	0	26	5	29	2
	UUC	8	17	20	25	0	11	15	19	4
Pro	CCG	24	26	43	32	0	5	22	15	4
	CCA	8	3	9	19	0	0	5	9	1
	CCU	24	3	3	6	0	11	5	6	0
	CCC**	0	0	0	6	0	5	1	8	3
Ser	UCG**	0	3	0	13	0	0	3	14	2
	UCA	0	6	0	6	0	16	1	13	1
	UCU	0	17	12	6	66	5	28	9	0
	UCC	41	17	23	0	13	5	16	10	6
	AGU	0	0	0	13	0	32	1	13	0
	AGC	8	14	12	19	26	16	10	9	1
Thr	ACG	0	14	3	19	0	0	1	16	4
	ACA	8	0	3	0	0	5	4	6	0
	ACU	41	9	29	0	26	16	32	11	0
	ACC	16	26	32	19	0	21	22	20	12
Trp	UGG	0	6	14	32	0	5	5	16	0
Tyr	UAU	16	0	6	19	0	11	5	19	1
	UAC	16	20	43	6	13	11	14	12	8
Val	GUG	33	28	6	32	13	16	14	19	12
	GUA	41	11	20	25	26	16	32	11	0
	GUU	49	11	49	13	39	0	34	20	2
	GUC	0	9	3	0	0	16	7	12	8

** Indicates rarely used codons in *E. coli*.

DISCUSSION

The 2.7 Kb fragment of pFDT210 containing repß was sequenced. Association of the sequence features with the functional analysis described in the previous chapter, delineated the basic ß replicon (nucleotides 478-2281). In addition, 3 structural domains, *ori*V, *rep* and *inc/cop* were identified.

The structural organisation of the *cis*-acting ß *ori*V (nucleotides 478-723) shows a few common features with the *ori*-2 of F (Tsutsui *et al.*, 1983; Murotsu *et al.*, 1984) and with the *ori*Vs of P1 (Abeles *et al.*, 1984), pSC101 (Vocke & Bastia, 1983; Linder *et al.*, 1985) and R6K (McEachern *et al.*, 1986). These features include an AT-rich region containing DnaA boxes and repeated sequences. In these plasmids, the repeats are the binding sites for the Rep proteins and are also involved in plasmid copy number control and incompatibility (Rokeach *et al.*, 1985; Kline, 1988; Abeles *et al.*, 1989; Armstrong *et al.*, 1986; Filutowicz *et al.*, 1985)

The repß *cis-ori*-V and the sequence immediately downstream (724-780) contain several unique direct and inverted repeats (IR1/1' and IR2/2'). Both of these IR's have the potential to form stem-loop structures (*Figure 3.4)*. This may be functionally significant since the AT-rich region containing the DnaA boxes was demonstrated to be essential for replication (*Figure 3.1D.*). Thus if this region become part of the stem and loop structures, it would interfere with replication.

DnaA-dependence of F plasmid was demonstrated by Kline *et al.* (1986) and Hansen & Yarmolinsky (1986). This has also been observed for RK2 (Gaylo *et al.*, 1987), pSC101 (Hasunuma & Sekiguchi, 1977), P1 (Hansen & Yarmolinsky, 1986; Abeles *et al.*, 1990) and pCU1 (Krishnan & Iyer, 1990). The position of the DnaA boxes at the origin of the ß replicon is similar to that of the DnaA boxes in F (Kline *et al.*, 1986) and in P1 (Abeles *et al.*, 1990) (*Figure 3.1B.*). This suggests that similar roles may be played by the DnaA

boxes in all these plasmids. These boxes precede a 90% AT-rich sequence which may be the site for DNA melting. An additional DnaA box is present within the first 70 bp of ORF295. Since this region is also characterised by a high degree of secondary structure, these features may contribute to the poor expression of ORF295. In the *E. coli* chromosome, DnaA boxes have been associated with the autoregulation of the *dnaA* gene, which is equivalent to the *rep* gene of plasmid replicons (Braun *et al.*, 1985; Wang & Kaguni, 1987). In addition to autoregulation, the binding of DnaA to the 4 boxes within the *ori*C causes strand opening within the origin by distorting the DNA helix (Bramhill & Kornberg, 1988).

The most striking feature of repß is its family of iterons. These are involved in copy number control and incompatibility (*inc/cop*). The structure of these repeats is similar to the Group I (13, thirty-seven base-pairs) iterons of the 2 Kb mini-replicon of pCU1, pCU714 (Kozlowski *et al.*, 1987; Krishnan & Iyer, 1990). Unlike pCU714, two degenerate iterons of unequal length (36 and 30 bp) are also present in this region of repß. The degenerate iterons (1 and 2) of repß are not similar to the Group II iterons of pCU714. Group II iterons confer *polA1* independence on pCU1, whereas iterons 1 and 2 of repß appear to be essential for plasmid replication regardless of the *polA1* status of the host.

Deletion of iterons 14 to 18 did not have any effect on the replication and copy number of repß. In fact, the plasmid pFDT220202 which contains only 13 iterons, has the same copy number as pFDT210 and pFDT220 (3-5 copies/chromosome). This suggests that the first 13 iterons (*Figure 3.5*) regulate copy number at the same level of the parent plasmid. Sequential deletion of iterons 13 to 4 led to a seven fold increase in copy number of the plasmid pFDT220204 (*Table 2.1B*). This increase in copy number is analogous to that described for similar deletions of mini-F, mini-P and mini-pCU1 (Tsutsui *et al.*, 1983; Pal *et al.*, 1986; Krishnan & Iyer, 1990). The mechanism that controls copy number and replication in repß is unknown and is thus inferred by analogy to that described for similarly organised plasmids. Data available are consistent with iteron mediation as

described for the F and P1 (Tsutsui *et al.*, 1983; Trawick & Kline, 1985; Pal *et al.*, 1986; Abeles & Austin, 1991).

Further deletion of iterons 1 to 4 led to an unstable replicon (pFDT220205), which may have led to run-away replication with consequent deleterious effects on the host. This phenomenon requires further investigation. These mutations are known to reduce the autorepressor activity of RepE, but never to abolish it, otherwise run-away replication occurs (Kline, 1988). Proposed models of iteron mediated control of copy number and corresponding adjustment of replication frequency are discussed by Norsdtröm (1990) and Abeles & Austin (1991). The latter researchers propose that the role of iterons in P1 is to align the newly replicated plasmids and to inhibit further rounds of replication until partition occurs. This ensures that the correct number of plasmids is allocated to each daughter cell.

In addition to the copy number control functions of iterons 1 to 13, all 18 iterons of repß (nucleotides 1605-2281), were associated with incompatibility. This incompatibility was independent of the number of iterons present (*Figure 3.1D.*) and is unlike the findings on the number of iterons required to express incompatibility of RK2 and mini-F (Stalker *et al.*, 1981; Tsutsui *et al.* 1983). Thus, the mechanism regulating incompatibility in repß may differ from that of these plasmids.

Besides the 18 iterons, other families of direct and indirect repeats also occur in repß. All of these may serve as potential binding sites for proteins, including that expressed from ORF295 (RepB). Abeles *et al.* (1989) carried out footprinting and gel retardation assays to show that the RepA of P1 binds uniformly to individual repeats both at the origin and *incA* locus. Similarly, RepE of F was shown to bind to the *incB* and *incC* repeats as well as to the *ori-2* region (Masson & Ray, 1986 and 1988; Tokino *et al.*, 1986). Since the *incB* repeats are separated from each other by integral numbers of helical turns, Kline (1988) proposed that these RepE binding sites lie exclusively on one side of the helix.

Other families of direct and inverted repeats, ranging from 5 to 15 bp are conserved within repß. The hexanucleotide ACACCC and the 15 bp sequence CCAAGTTGACCTGTT are conserved in the majority of the iterons. The sequence ACAGGG present in repß has also been reported in the mini-replicon of pCU1 (Iyer, 1989). The pentanucleotide sequence ACAGG which occurs within the iterons of the ß replicon has also been reported in the *ori*V region of the plasmids R6K (Stalker *et al.*, 1979), F (Murotsu *et al*, 1981), RK2 (Stalker *et al*, 1981), pSa (Tait *et al*, 1983), Rts1 (Kamio *et al* 1984 and 1988) and pCU1 (Iyer, 1989). Similar hexanucleotides have also been observed in the origins of replication of bacteriophage lambda and the *E. coli* chromosome (Tolun & Helinski, 1981). A crucial role for these small sequences has been proposed by Seelke & Kline (1984) and Filutowicz *et al.* (1985). Furthermore, it has been suggested that the similarities among the repeats of these replicons may reflect the interaction with a common cellular component that is essential for the initiation of replication (Kline, 1988).

Of the 3 ORF's identified in repß, only ORF295 was shown to be essential for autonomous replication. Two proteins, postulated to be translated from different start sites, were expressed from this ORF (30.8 and 32 KDa). Since the 30.8 KDa protein was the major product, it was designated RepB. Recently, different forms of at least one Rep protein (RepE) have been described in the same host. Kline *et al.* (1992) demonstrated that RepE (mini-F) undergoes site-specific proteolysis *in E. coli* CSH50 to give rise to Δ1- and Δ17-RepE (the prefixes Δ1- and Δ17- refer to the number of amino-acids deleted in the proteolytic processing of RepE). These RepE forms exhibit different binding affinities for target DNA (*ori-2, incC*) in the presence of *E. coli* chromosomal DNA. It is possible that the 30.8 and 32 KDa RepB forms have an equivalent function to the Δ1- and Δ17-RepE forms.

The size of the RepB protein (30.8/32 KDa (*Figure 3.8B)* is similar to that of initiator proteins expressed from other basic replicons such as the RepE of F (29 KDa, Watson *et al.*, 1982), Rep of pCU1 (Krishnan *et al.*, 1990), RepA of P1 (32 KDa, Abeles, 1986), TrfA of RK2 (32/43 KDa, Thomas and Helinski, 1989), RepA of pSC101 (37 KDa, Armstrong *et al.*, 1984) and RepA1 of R100 (33 KDa, Rosen *et al.*, 1980; Armstrong *et al.*, 1986).

It is unclear why only ORF295 is expressed *in vitro* since none of the promoter sequences of the 3 ORFs in repß show good homology to those of *E. coli* (*Figure 3.7*). In addition, all putative RBS identified are weak and are not ideally positioned from the start codons (*Table 3.1*). This distance is critical, and if it is shorter than 5 or longer than 9 nucleotides it impairs translation (Kozak, 1983).

There is also the possibility that ORF175 and ORF212 are non-coding open reading frames. Though, differences in the G+C content and codon usage of ORF295 and those of the other ORFs, may explain the lack of expression from ORF175 and ORF212. The codon usage of ORF295 is similar to that of other genes involved in replication functions. At present, however, there is no consensus as to whether or not codon usage can influence regulation of gene expression (Bonekamp & Jensen, 1988; Holm, 1986; Robinson *et al.*, 1984; Sharp & Li, 1986).

In summary, the functional organisation of repß includes three major domains (*Figure 3.9*). The AT-rich *cis*-acting *ori*V domain, followed by the *repB* gene which encodes the essential replication protein (RepB) and the eighteen iterons, which are involved in incompatibility function and maintenance of a low copy number (*inc/cop* locus). These domains are contained within nucleotides 478-2281. The minimal sequence required for autonomous replication, however, spans only nucleotides 478-1773, and includes the *ori*V, *repB* as well as the 4 proximal iterons to the latter.

This compact molecular organisation is similar to that of the mini-replicon of plasmid pCU1 (Krishnan & Iyer, 1990), to the repFIA of F plasmid (Murotsu *et al.*, 1981) and to the mini-replicon of P1 (Abeles *et al.*, 1984). The organisation of these replicons may indicate that plasmids use the same host specified enzymes for replication and that an initiation complex has to conform with a specific configuration and stoichiometry (Kline, 1988).

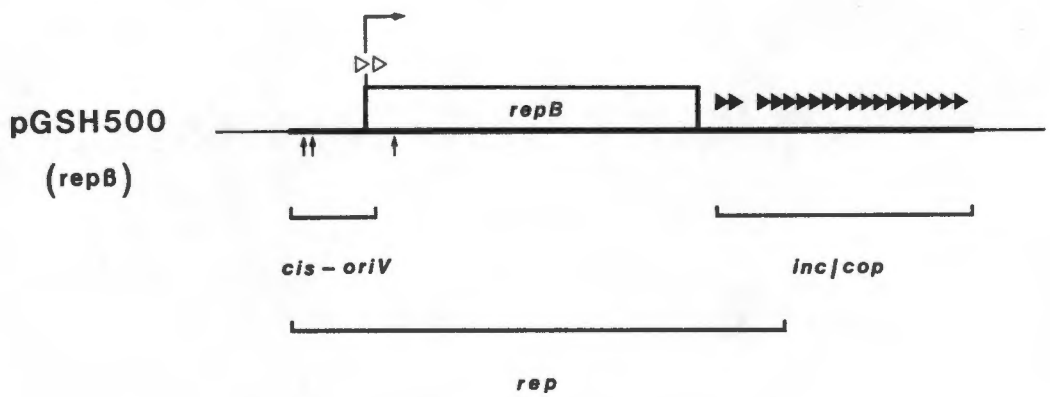


Figure 3.9. The functional organisation of repß in pGSH500. The *cis-ori*V domain delineates the minimum region required *in cis*. *rep* indicates the minimal essential region for autonomous replication; this domain consists of the *cis-ori*V, the gene *repB* and 4 iterons. The *inc/cop* domain include 18 iterons, of which 16 are tandem; this domain is associated with the incompatibility and copy number control properties of pFDT200. Vertical arrows indicate the position of DnaA boxes, whereas horizontal arrows indicate the location of iterons (open arrows, two 15 bp direct repeats overlapping the start or *repB*; filled arrows, eighteen 30-36 bp iterons).

CHAPTER 4

COMPARISON OF repß WITH OTHER MINI-REPLICONS

CHAPTER 4

COMPARISON OF repß WITH OTHER MINI-REPLICONS

SUMMARY

The 2.7 Kb sequence of repß was compared with that of mini-pCU1. There is an overall 53% identity between these two replicons. DNA and amino-acid sequence identities (both 58%) were also found between the respective Rep proteins. In addition, RepB showed a 31% amino-acid sequence identity with RepE of mini-F. Both RepE and the Rep of pCU1 were shown to complement the function of deleted forms of repß. The *cis-ori*V region of repß shows an 80% DNA sequence identity (over 200 bp) with the corresponding region in *ori-2* of mini-F. Analysis of the sequence of repß suggests that this replicon is either a "hybrid" of two unrelated replicons (mini-pCU1-like and mini-F-like) or an ancestral form from which these other origins have evolved.

INTRODUCTION

DNA and amino-acid sequences of newly identified loci can be compared with those available in sequence data banks. Nucleotide sequence comparison can disclose regions of similarity, which may not have been suspected. Such similarities may reflect evolution from an ancestral locus as has been postulated for some of the basic replicons of the IncF plasmids. Furthermore, these similarities may contribute to incompatibility of related plasmids.

The F plasmid contains three discrete replicons, repFIA, repFIB and repFIC (Bergquist *et al.*, 1986). Although the sequence organisation of repFIA (Murotsu *et al.*, 1981) is distinct from that of repFIB (Lane *et al.*, 1984) and repFIC (Rosen *et al.*, 1980; Picken *et al.*, 1984), nucleotide sequence similarities occur between these and other basic replicons of IncF plasmids. Examples include similarities between repFIC (pCG86 and P307) and repFIIA (R100 and R6) (Picken *et al.*, 1984; Saadi *et al.*, 1987). In addition, repFIII, repFVIB and repFVII of the pSU plasmid family also show similarities with repFIIA (López *et al.*, 1989a, b; López *et al.*, 1991). Similarities have also been reported between the repFIIA replicon of R100 and that of the more distantly related R1 (Ryder *et al.*, 1982).

Analysis of nucleotide sequence similarities within these replicons supports the proposal that evolutionary divergence in the IncF group has occurred from an ancestral plasmid through genetic rearrangements (Bergquist *et al.*, 1986). Sequence analysis of the repFIIA replicons has led to the proposal that these may have evolved by module-exchange (gene shuffling) (Saadi *et al.*, 1987; López *et al.*, 1991). This process allows for rapid gene divergence, while recruiting control genes or determinants from the chromosome and other plasmids. It may also be similar to the exchange of exons in eukaryotic chromosomes (López *et al.*, 1991).

In this chapter the DNA sequence of the basic replicon of pGSH500 (repß) is compared with that of mini-pCU1 (Krishnan & Iyer, 1990; Krishnan *et al.*, 1990) and that of mini-F (repFIA) (Murotsu *et al.*, 1981; Murotsu *et al.*, 1984). The amino-acid sequence of some of the proteins encoded by these replicons is also compared.

MATERIALS AND METHODS

Computer programs used for DNA and amino acid sequence analysis

The DNA sequences were analysed using the GCG software package (version 7.1) (Devereux *et al.*, 194) running on a Vax 6000-330 mainframe and using VMS version 5.41. Genpro (Riverside Scientific, version 5.0) was used to generate dot plots and to identify open reading frames on an IBM compatible personal computer for illustrative purposes only. The Clustal V (Higgin *et al.*, 1991) program was used for multiple amino-acid sequence alignments. FASTA (Pearson & Lipman, 1988) was used for DNA alignments. FASTA was also used to screen sequences at GENBANK by email. This was done to ensure that no sequences published since the GENBANK release 70.0 and EMBL release 29.0, were missed. The 2.7 Kb sequence of repß has been submitted to the EMBL data bank, accession no Z11775.

RESULTS

The database searches revealed a meaningful similarity between the DNA sequences of repß and mini-pCU1 (Kozlowski *et al.*, 1987; Krishnan & Iyer, 1990). In addition, similarities between the amino-acid sequences of RepB, Rep (pCU1) (Krishnan & Iyer, 1990) and RepE (Murotsu *et al.*, 1981) were disclosed. Similarity between the *ori*V of repß and the *ori-2* of mini-F was also revealed. These similarities are discussed below.

Comparison of repß with the mini-replicon of pCU1

The 2684 bp sequence of the ß replicon (pFDT210) was aligned with the 2056 bp sequence of the replicon of pCU1 (pCU714) (Krishnan & Iyer, 1990; Krishnan *et al*, 1990) (*Figure 4.1.*). There is an overall DNA sequence identity of 53% between these two

replicons. Furthermore, the DNA regions encoding ORF295 (repß) and ORF239 (pCU1) show 58% identity. In both replicons the incompatibility and copy number control functions coincide with a region of tandem direct repeats (iterons). The 16 distal conserved iterons of repß (3-18), which are 34-36 bp in length, appear to be equivalent to the 13, 37 bp tandem iterons of pCU714 (group I iterons). A number of short sequences which are repeated in repß, are also present in pCU714. These include ACAGG (21 times in repß) and TGTGGG (12 times in repß). Despite the lack of similarity between the *cis*-acting *ori*V's of the two replicons, the molecular organisation of the ß replicon resembles that of the replicon of pCU1 (*Figure 4.2.*).

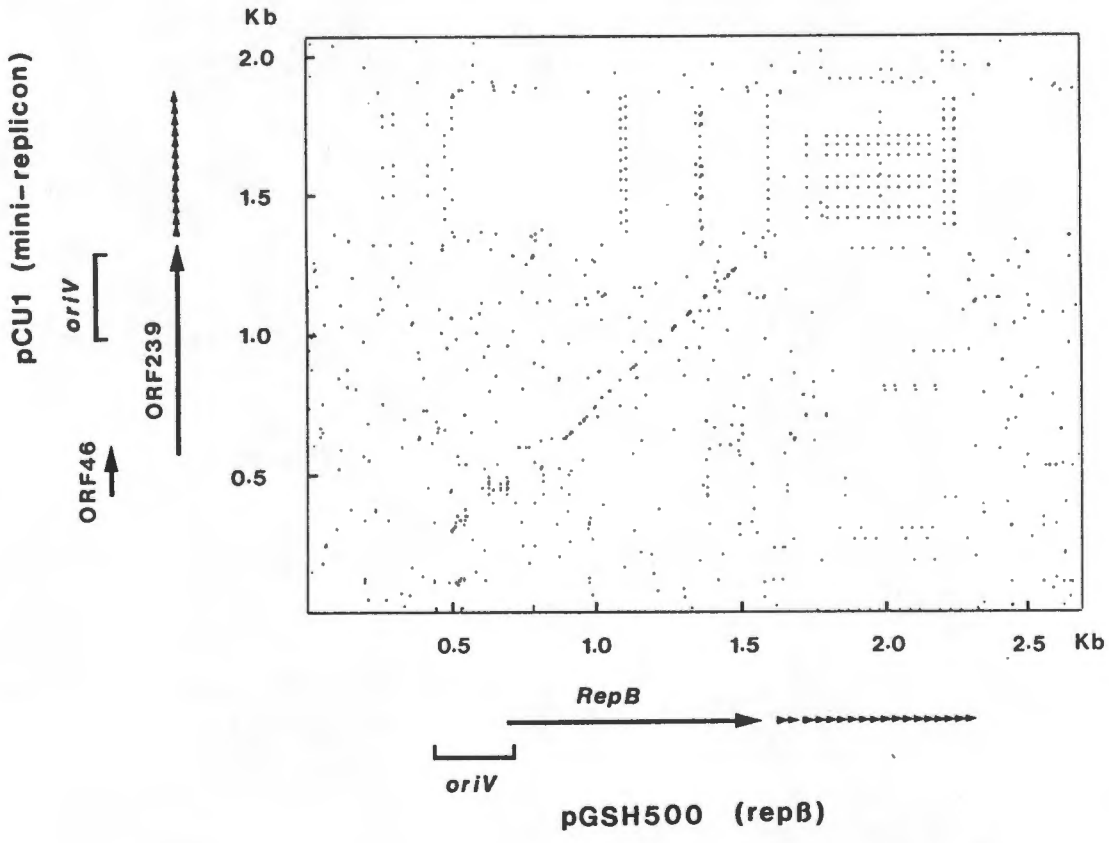


Figure 4.1. Dot plot of the nucleotide sequences of the 2.7 Kb DNA fragment (pFDT210) containing the basic ß replicon (nucleotides 478-2281) compared to the 2 Kb *Pvu*II mini-replicon of pCU1.

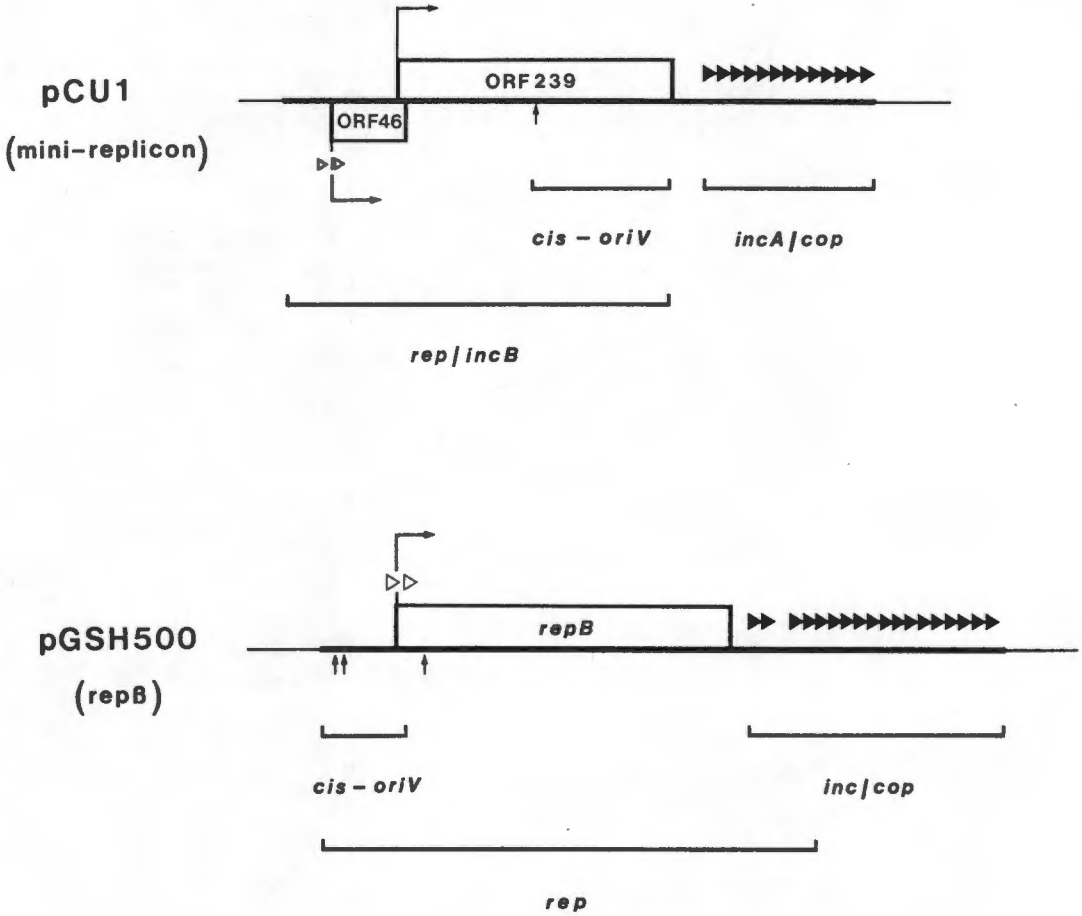


Figure 4.2. Comparison of the functional organisation of repß with that of mini-pCU1 (Krishnan & Iyer, 1990). The *cis-ori*V indicates the *cis*-acting origin of replication. In pGSH500, the *rep* domain demarcates the minimal region of repß capable of autonomous replication in a *polA1* deficient *E. coli*, although with a 7 fold elevated copy number. In pCU1, *rep* indicates the minimal region capable of autonomous replication in a wild-type (PolA+) *E. coli*. *repB*, indicates the RepB locus. In pCU1, the products of ORF239 and ORF46 (corresponding to a 27 and 5.5 KDa proteins) are essential for replication (Krishnan *et al,* 1990). The *inc/cop* domain encodes plasmid incompatibility and copy number control functions (▶). In repß, this region includes 18 iterons, whereas only 13 iterons are present in the equivalent region of mini-pCU1 (*incA/cop*). DnaA boxes in both basic replicons are indicated (↟). The open arrows (▷) indicate the presence of direct repeats at the beginning of ORF46 and *repB*. The arrow (↱) shows direction of transcription.

Comparison of the amino acid sequence of RepB with other Rep proteins

The putative amino-acid sequence of the RepB was compared with those in the GENBANK and EMBL databases. Two clear similarities emerged. The first was a 58% identity with the product of ORF239 of pCU1 (Krishnan & Iyer, 1990; Krishnan *et al.*, 1990) (*Figure 4.3.*). The second was a 31% identity with the RepE of the mini-F plasmid (Murotsu *et al.*, 1981) (*Figure 4.4.*). This similarity is not reflected at the DNA level. The alignment of the predicted amino-acid sequence of RepB with that of the ORF239 polypeptide (pCU1, Krishnan & Iyer, 1990; Krishnan *et al.*, 1990) and the RepE protein (mini-F plasmid; Murotsu *et al.*, 1981; Watson *et al.*, 1982) is shown in *Figure 4.5*. This multiple sequence alignment identifies identical amino-acid and neutral substitutions in all 3 polypeptides.

Two regions of the three polypeptides showing more than 80% identity were identified (region **I** and **II**). In region **I**, 10/12 amino-acids are identical, whereas in region **II**, 7/8 amino-acids were conserved in all 3 sequences. The sequence conservation in these regions suggests that regions **I** and **II** may have similar functions in all three proteins (see later). Another region of similarity has also been identified between a putative α helix- turn- α helix consensus sequence in repE (after Pabo & Sauer, 1984) and the equivalent region in RepB and in the product of ORF239 (*Figures 4.5. & 4.6.*).

```
RepB        MGCASTSYGKEVKFSSQLSLVDKVQIFPVDKYRDSLQFLFSLSLWVGKNLPIGIEMDTQA
                    10        20        30        40        50        60

                       10        20        30        40        50
ORF239         MDKLLNKKIKVKQSNELTEAAYYLSLKAKRVLWLCLMQTYFTASVSEDDDEMAVLGD
                 : ..:. ..:::::::::::::::.:.:::::::::::::.::..:    ..:. .::
RepB        LLPATKTFKKRASIKQSNELTEAAYYLPLQAKRVLWLCLMQAYFNDS---QEDDSDVL--
                    70        80        90       100       110

             60        70        80        90       100       110
ORF239      STFKVKVADYQQIFQVSRNQAIKDVKEGVFELSRSAVIFYPKEGSFDCVARPWLTEAGSR
            . X:..:.:: . :.:. . : .:::.:: .:. :.:.:::::::.:. : ::::.::: .
RepB        PLFKISVSDYVKYFNVATSVASRDVKAGVNALGESTVTFYPKEGEFEEVKRPWLAEAGMK
              120       130       140       150       160       170

            120       130       140       150       160       170
ORF239      SARGIWEIEFNHKLLRYIYGLTNQFTTYSLRDCGSLRNPRTIRLYESLAQFKSSGLWVTT
            ..:: :.::::.:.... :::.:::::::: ::: :.. :.:::::::: ::.:.:.:.::
RepB        RGRGSWQIEFNYKVMPFLVGLTSQFTTYSLYDCGQLNSVRVIRLYESLCQFRSTGVWITT
              180       190       200       210       220       230

            180       190       200       210       220       230
ORF239      HAWLNDRFLLPESQQKNLAELKRSFLDPALKQINEKTPLLAKYSIDDSGKFLFSIIDKQN
            :.:: .::.::.::..:.::.::.::.::::.:::::::: ..:.....:..:X...:
RepB        HDWLCERFMLPASQKNNIAEMKRTFLEPALKKINEKTPLKVSYKTEEDGRLLFNFLDGKQ
              240       250       260       270       280       290


ORF239      PV
```

Figure 4.3. Alignment of the amino acid sequences of the product of ORF239 of pCU1 with RepB of repß. Conserved amino-acids (:) and neutral changes (.) are indicated (see text for further details). There is a 58% identity over a 232 amino-acid overlap.

```
RepE                                                                  MAET
                                                                      ...
RepB        MGCASTSYGKEVKFSSQLSLVDKVQIFPVDKYRDSLQFLFSLSLWVGKNLPIGIEMDTQA
                    10        20        30        40        50        60

                10        20        30        40        50        60
RepE        AVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRMLYLFVDQIRKSDGTLQEHDGIC--EI
             . . :. :....: :::.:::::: :. ..::.:.: . :   .:.  .. : .   .:
RepB        LLPATKTFKKRASIKQSNELTEAAYYLPLQAKRVLWLCLMQAYFNDSQEDDSDVLPLFKI
                    70        80        90       100       110       120

              70        80        90       100       110       120
RepE        HVAKYAEIFGLTSAEASKDIRQALKSFAGKEVVFYRPEEDAGDEKGYESFPWFIKRAHSP
             :..:.. :..... ::.:............:.:: :.:.. .:    . X:. . . ..
RepB        SVSDYVKYFNVATSVASRDVKAGVNALGESTVTFY-PKEGEFEE---VKRPWLAEAGMKR
                130       140       150       160           170

            130       140       150       160       170       180
RepE        SRGLYSVHINPYLIPFFIGLQNRFTQFRLSETKEITNPYAMRLYESLCQYRKPDGSGIVS
            .:: . ...:  ..::..:: ..:: ..: .  ....  ..:::::::X.:   ..:.
RepB        GRGSWQIEFNYKVMPFLVGLTSQFTTYSLYDCGQLNSVRVIRLYESLCQFR---STGVWI
             180       190       200       210       220           230

            190       200       210       220       230       240
RepE        LKIDWIIERYQLPQS-YQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFS
             . ::. ::. ::.:  .......: ::.  ...::..::...:: ....::
RepB        TTHDWLCERFMLPASQKNNIAEMKRTFLEPALKKINEKTPLKVSYKTEEDGRLLFNFLDG
                240       250       260       270       280       290

              250
RepE        FRDITSMTTG

RepB        KQ
```

Figure 4.4. Alignment of the amino acid sequences of RepE of mini-F with RepB of repß. Conserved amino acids (:) and neutral changes (.) are indicated (see text for further details). There is a 31% identity over a 235 amino-acid overlap.

FIGURE 6

Figure 4.5. Alignment of the amino acid sequences of the product of ORF239 of pCU1 and RepE of mini-F with RepB of repß. Conserved amino acids in all 3 sequences (*) and neutral changes (·) are indicated. Alignment of the amino acid sequences was done using the ClustalV program (Higgin *et al.*, 1991). The arrow indicates the position from where a 30.8 KDa polypeptide could be expressed. **I** and **II** indicate conserved motifs in all 3 sequences, whereas (----) shows the position of a putative α-helix-turn-α motif.

Protein	Position	Sequence
		1 5 10 15 20
		α Helix Turn α Helix
		. . . * **A** . . * **G** * . . . . **I/V/L** . . * * .
ORF239	66	V A D Y Q Q I F Q V S R N Q **A** I K D V K
RepB	102	V S D Y V K Y F N V A T S V **A** S R D V K
RepE	64	V A K Y **A** E I F **G** L T S A E **A** S K D I R
		* . * . * . . * . * . * . .

Figure 4.6. α-helix-turn-α-helix region in RepE (mini-F) and putative corresponding motifs in RepB (repß) as well as in the product of ORF239 (mini-pCU1). The highly conserved amino acids in positions 5, 9 and 15 are shown in bold type face. Asterisk and dots represent non-polar and hydrophobic amino acids, respectively. ( ) represent amino acids conserved in all 3 sequences, whereas ( ) represent neutral substitutions.

Comparison of the sequence of repß with repFIA

Because of the amino-acid sequence similarity of the RepB protein to the RepE protein of mini-F, the two replicons were compared for other sequence similarities (*Figure 4.7.*). An 80% identity was found over a 200 bp segment immediately upstream of the *ori*-2 region (mini-F plasmid) (Murotsu *et al.*, 1984) and an equivalent region of the repß *cis*-acting *ori*V. No other similarity was observed between pFDT210 and the 2.3 Kb sequence of repFIA (Murotsu *et al.*, 1981 and 1984; Bergquist *et al.*, 1986). In the 200 bp region of the repß (nucleotide 370-580) sequences identical to those of the primosome sites of ColE1 and F, two DnaA boxes and an 90% AT-rich sequence (70 bp) were identified. Several small repeated sequences were also found in this region (see Sequence analysis, Chapter 3). In this respect, the overall functional organisation of the *cis-ori*V of repß resembles that of the *ori-2* of mini-F, to the left of the 4 *incB* direct repeats (Murotsu *et al.*, 1981; Kline, 1988) (*Figure 4.8.*).

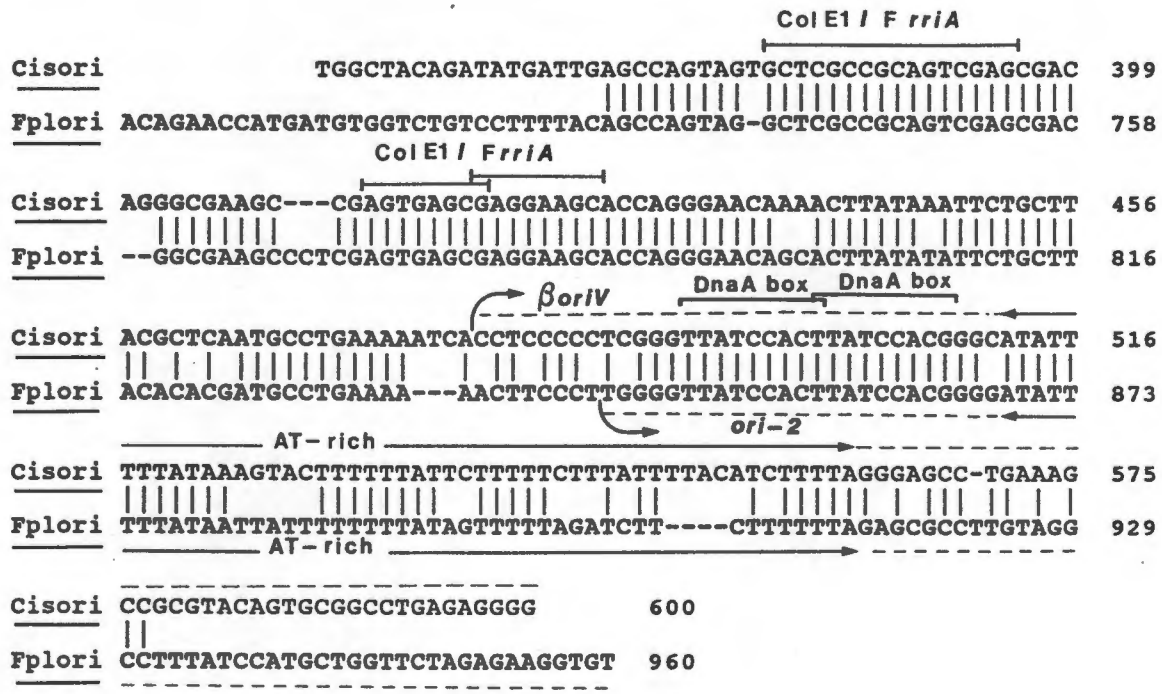


Figure 4.7. Alignment of the *cis-ori*V of repß (Cisori) with the *ori-2* region of the mini-F plasmid (Fplori). Sequences homologous to either 5'-GTGAGCG-3' or 5'-GNGGAAGC-3' are indicated (⟼). These sequences are present within the primosome sites of the ColE1 and F plasmids (Imber *et al.*, 1983). The start of *ori*V and *ori-2* are marked with a curved arrow ( ). The extent of the origins is indicated (- -). The AT-rich sequences and position of the DnaA boxes is also shown.

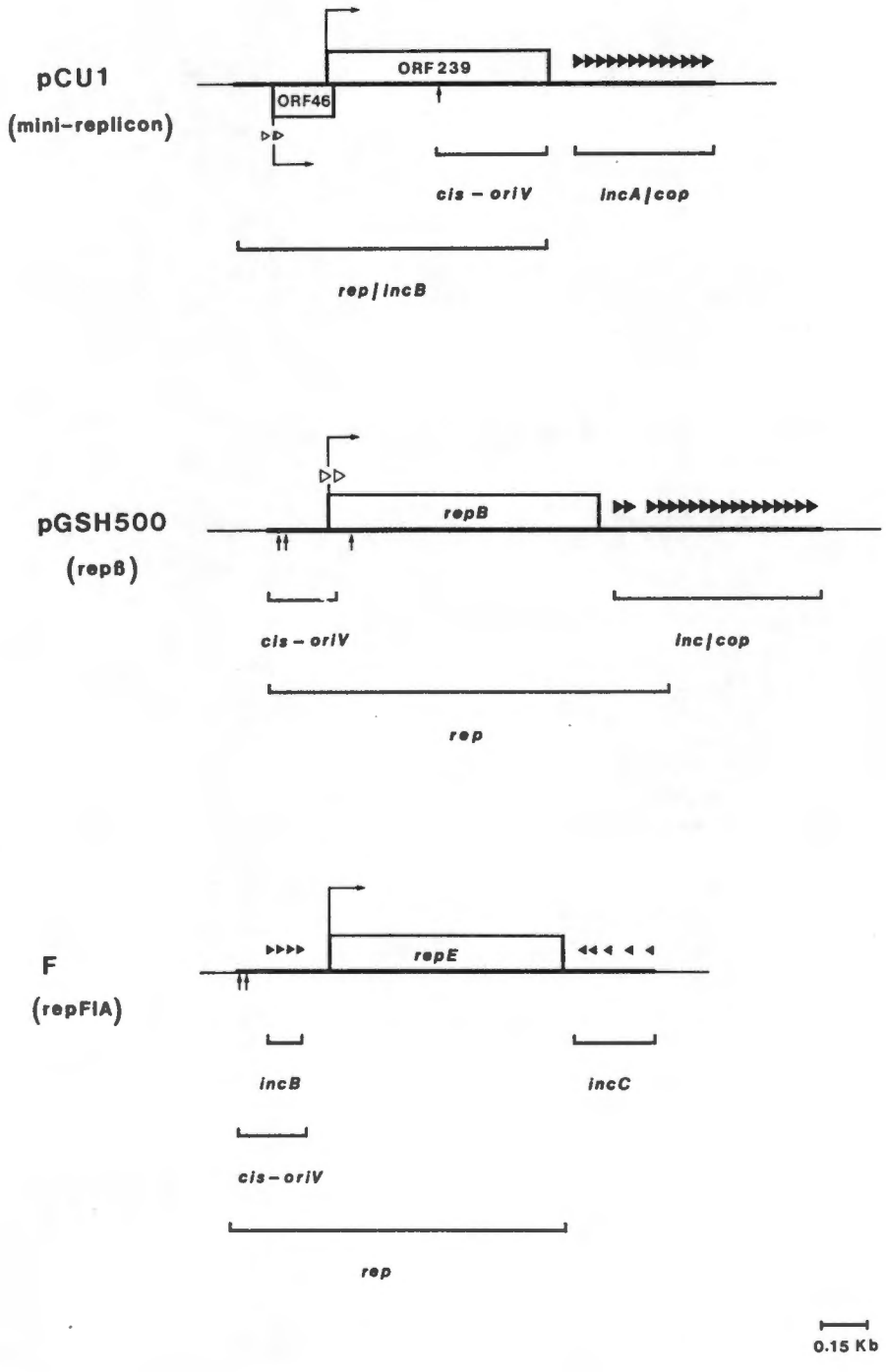


Figure 4.8. The functional organisation of the basic replicon, repß, in pGSH500 is compared to that of the basic replicons of pCU1 (Krishnan & Iyer, 1990) and mini-F (repFIA) (Murotsu *et al.*, 1981; Tsutsui *et al.*, 1983) (full legend in the following page).

Figure 4.8. The functional organisation of the basic replicon, repß, in pGSH500 is compared to that of the basic replicons of pCU1 (Krishnan & Iyer, 1990) and mini-F (repFIA) (Murotsu *et al.*, 1981; Tsutsui *et al.*, 1983). The *cis-ori*V indicates the minimal *cis*-acting region required for replication in all replicons. In F, the *cis-ori*V is also designated *incB*. *rep*, designates the essential domain for autonomous replication. ORF239, *repB* and *repE* indicate the genes for essential proteins. ORF46, is necessary for pCU1 replication, but its role is unclear. The *inc/cop* domains of repß and mini-pCU1 encode plasmid incompatibility and copy number control determinants. These domains are equivalent to *incC* in repFIA. The iterons in these replicons are marked (▶). In mini-pCU1 there are thirteen 37 bp iterons, whereas in repß (pGSH500) there are eighteen 30-36 bp. In repFIA, the *incC* domain contains five 19 bp iterons, whereas only four are present in *incB*. DnaA boxes (▮) and direct repeats at the start of ORF46 and *repB* are indicated (▷). Direction of transcription is also shown (⌐).

DISCUSSION

As shown in *Figures 4.1 & 4.2*, the molecular and functional organisation of repß is similar to that of the mini-replicon of the broad-host-range plasmid pCU1 (Krishnan & Iyer, 1990; Krishnan *et al.*, 1990). Both repß and mini-pCU1 are compact replicons with all the essential features contained in less than 2 Kb. This compact organisation also resembles that of the replicons of the narrow-host-range plasmids F (Murotsu *et al.*, 1981), P1 (Abeles *et al.*, 1984) and pSC101 (Yamaguchi & Yamaguchi, 1984a and 1984b). In contrast, the compact organisation of repß differs from that of the large basic replicons of RSF1010 (Frey & Bagdasarian, 1989; Haring & Scherzinger, 1989) and of RK2 (Thomas & Helinski, 1989).

The protein RepB (repß) shows a 58% amino-acid sequence identity with the product of ORF239 of mini-pCU1 (Krishnan & Iyer, 1990; Krishnan *et al.*, 1990) (*Figure 4.3*). RepB also shows a 31% identity to RepE (mini-F, Murotsu *et al*, 1981 and 1984) (*Figures 4.4*). The similarity between RepB and RepE was most unexpected as no similarity was observed at the DNA level. The similarity in amino-acid composition of these 3 proteins may reflect functional similarity *in vivo* (see Chapter 5). Furthermore, this similarity may indicate that the roles of these proteins in replication control are similar.

Sequence comparison of these polypeptides (*Figure 4.5*) revealed the presence of 2 highly conserved regions (**I** and **II**). The amino-acid identity in regions **I** (10/12) and **II** (7/8) suggests that these sequences may represent protein motifs. These conserved regions may be important in protein-DNA or protein-protein interactions within the replisome. Regions of homolgy (HR) alternating with non-homologous (NHR) regions have been described within repFIC of P307, F, R100 and R1 (Saadi *et al.*, 1987) as well as repFVIB of pSU212 (López *et al.*, 1991). These HR and NHR regions are thought to function as modules in the evolution of F plasmids in a similar manner to "gene shuffling" proposed for the exchange of exons in eukaryotic chromosomes (López *et al.*, 1989a).

An additional region (*Figure 4.5*) shows similarities to the α-helix-turn-α-helix consensus sequence in RepE (Pabo & Sauer, 1984). Since some amino-acid substitutions have occurred in RepB and the polypeptide of ORF239 (*Figure 4.6*), the necessary polar/hydrophilic amino-acid arrangement may not exist and this region may no longer be α helix-turn-α helix motif in these polypeptides.

The region of repß associated with incompatibility and copy number control (*inc/cop*) is similar to that of pCU1. This locus in repß contains eighteen 30-36 bp iterons, whereas in pCU1 this region comprises of thirteen 37 bp iterons (*Figure 4.2.*). These similarities may indicate that these replicons utilise similar incompatibility and copy number control mechanisms. An additional *inc* locus with no corresponding region in repß, is present in pCU1 (*Figure 4.2 & 4.8*).

Despite the overall similarities between these two replicons, differences exist in the organisation of the minimal region required *in cis*. The *cis-ori*V (368 bp) of the pCU1 mini-replicon overlaps part of ORF239 and includes 3 GATC sequences and 1 DnaA box (Krishnan & Iyer, 1990). In contrast, the *cis-ori*V of repß is located upstream of ORF295 (*repB*). This is an AT-rich region, which contains two DnaA boxes that are essential for replication. No GATC sequences were found in repß (Chapter 3.). It is interesting to note that the essential ORF46 in pCU1 corresponds to the *cis-ori*V of repß (*Figure 4.2 & 4.8*). Nevertheless, there is no significant sequence homology between the two replicons in this region. Interestingly, during replicon typing (Chapter 2) a signal was obtained between the IncN probe (pULB2432) and pGSH500, but not with either pFDT100 (repα) or pFDT200 (repß). This suggests that pGSH500 may contain a third replicon with sequence similarity to the IncN fragment used as a probe.

The DNA sequence identity between a 200 bp region stretching over part of the ß *cis-ori*V region and the corresponding region in *ori-2* of mini-F (Murotsu *et al*, 1981 and 1984)

(*Figure 4.7*) was both unexpected and remarkable. No hybridisation signal was obtained between the repFIA *inc/rep* probe (Couturier *et al.*, 1988) and repß during replicon typing (Chapter 2). The similarity between repß and repFIA stops abruptly before the beginning of the *incB* direct repeats, at nucleotide 931 (*ori-2*) (*Figures 4.7 & 4.8*; see also *Figure 1.3*). This could reflect a functional module incorporated into these replicons. It is not surprising that repß and mini-F are compatible (Chapter 2), since they share no similarity over one of the regions in mini-F that is involved in incompatibility (*incB*). Furthermore, the reason why this region of identity was not recognised during replicon typing of repß is now evident (Chapter 2). The replicon probe repFIA (Couturier *et al.*, 1988) includes the sequence corresponding to nucleotides 930-1846 of the basic replicon on mini-F (Murotsu *et al.*, 1981; Couturier *et al.*, 1988). Since this sequence overlaps the corresponding sequence in repß by only 2 bp (*Figure 4.7*), no hybridisation signal should be obtained.

The common regions adjacent, but outside of the *ori*Vs of repß and repFIA (*Figure 4.7*), include *rri* sites (rifampicin resistance initiation sites), also designated *ssi* or single-strand initiation sequences (Nomura *et al.*, 1991). Within the *ori*Vs, the sequence identity spans a 90% AT-rich region containing two DnaA boxes. Similarities among 11 *ssi* sites in the *ori*V regions of F, R6K, R100 and ColE2 plasmids have been reported (Nomura *et al.*, 1991). In F, it is likely that the *ssiA* near the *ori-2* functions as a primosome assembly site for the initiation of multiple RNA priming on the lagging strand. Nomura *et al* (1991) have suggested that although the *ssiA* sites of the plasmids F (f5), R100, R6K, ColE1 and ColE2 are mapped adjacent to the "minimal-*ori*s" and are not essential for autonomous replication, they may enhance the replication activity of these plasmids. Enhancement of replication of pBR322 containing an n' site (equivalent to *ssiA*) was demonstrated both *in vivo* and *in vitro* (Masai & Arai, 1988 and 1989). In repß, the *rriA* sites are not essential for replication, but the possibility that these sites enhance the replication ability of the replicon has not been excluded.

In conclusion, the molecular organisation of repß appears to be either a "hybrid" between that of two unrelated replicons, pCU1 (broad-host) and F (narrow-host), or an ancestral form from which the other, more specialised, replicons have evolved.

CHAPTER 5

COMPLEMENTATION AND COMPATIBILITY OF repß WITH MINI-pCU1 AND MINI-F

CHAPTER 5

COMPLEMENTATION AND COMPATIBILITY OF repß WITH MINI-pCU1 AND MINI-F

SUMMARY

Deleted forms of RepB were shown to be complemented by the Rep proteins of pCU1 and mini-F. In addition, repß was demonstrated to be fully compatible with both of these plasmids.

INTRODUCTION

Data reported in Chapter 4 suggested that RepB is functionally similar to the Rep proteins of mini-pCU1 (Krishnan & Iyer, 1990) and mini-F (Murotsu *et al,* 1981 and 1984). To test this possibility, deleted forms of repß were tested for functional complementation by these plasmids. Furthermore, the incompatibility properties of pFDT200Gm^r towards pCU1 and mini-F were examined.

MATERIALS AND METHODS

Bacterial strains and plasmids

E. coli strain LKIII was used for general plasmid maintenance and incompatibility studies, whereas GW125a(*polA1*) was used to test *polA1*-independent replication (Chapter 2, *Appendix 1*). Plasmid pNZ116 (IncFI) containing the 9.1 Kb f5 fragment of F (mini-F) was obtained from Professor P. Bergquist and Dr. A. Spiers (University of Auckland, N. Zealand). This plasmid was transformed into GW125a(*polA1*), using the method of Chung & Miller (1988) (Chapter 2). The plasmid pCU1 (IncN) was isolated from the strain J53-2(RM98) as described (Konarska-Kozlowska & Iyer, 1981) and was transferred by conjugation (As detailed in Chapter 2) to the GW125a(*polA1*) strain. The strain J53-2(RM98) was obtained from Dr K. Towner (PHLS Laboratory, Nottingham, U.K.). Specific constructs for complementation studies of pFDT200 deletions (see *Figure 3.1.*) with pCU1 and pNZ116 are shown in *Figure 5.1.* Other plasmids are summarised in *Appendix 1.*

Media and Antibiotics

These were prepared and used as described in Chapter 2.

Construction of deletion derivatives of pFT200

The plasmid pFDT200, its deletions, pFDT220206-pFDT220210, as well as the mini-replicon pFDT200Gm^r (Chapter 2, *Figure 2.6, Appendix 1*) were used. The ability of pNZ116 and pCU1 to complement the pFDT200 deletions derivatives was tested. Since all

the plasmids used in this assay encoded ampicillin resistance, an additional resistance marker (Gm^r) was introduced in pFDT220206-pFDT220210 as follows: the ampicillin gene of each of the plasmids pFDT220206-pFDT220210 was inactivated by insertion of a gel purified, blunt-ended 2.1 Kb *Eco*RI-*Hin*dIII DNA fragment encoding gentamycin resistance (aminoglycoside acetyl transferase, AAC(3), Elisha & Steyn, 1991) at the *Ava*II sites (after blunt-ending the latter) (*Figure 5.1*). These Gm^rAp^s pFDT220206-pFDT220210 plasmids were designated pFDT220206Gm^r-pFDT220210Gm^r, respectively.

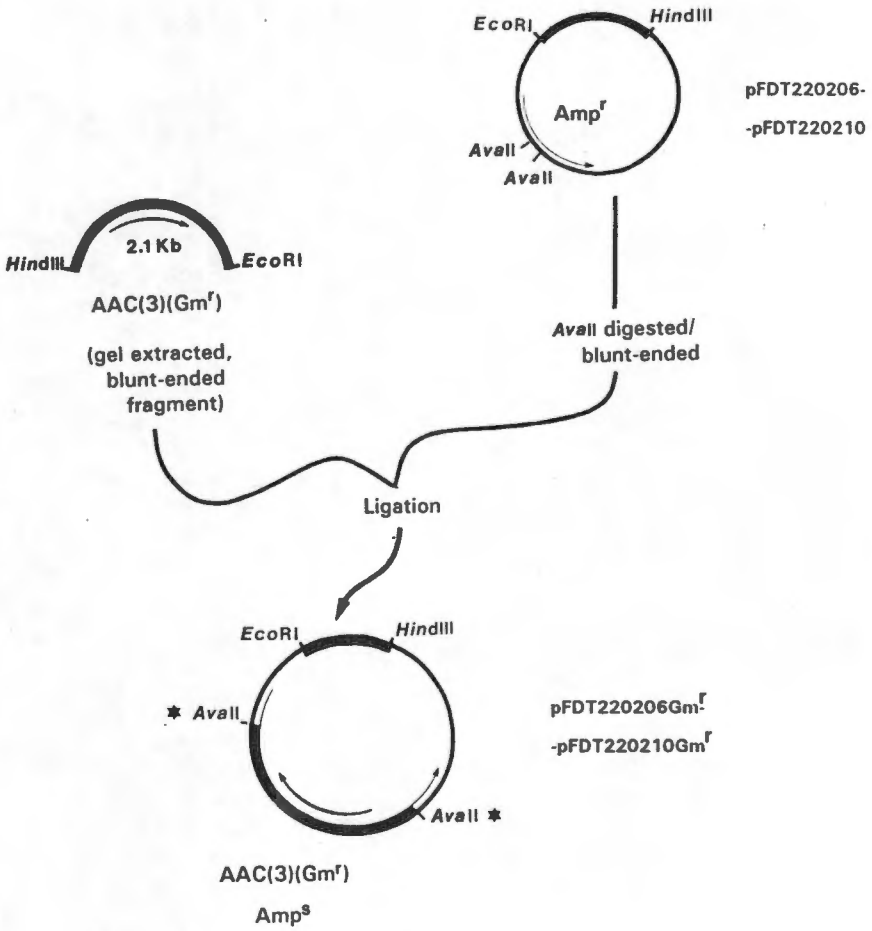


Figure 5.1. Inactivation of the ampicillin gene of pUC19 in pFDT220206Gm^r-pFDT220210Gm^r. The Ampicillin gene was inactivated by restriction with *Ava*II, followed by blunt-ending and ligation to a 2.1 Kb gentamycin resistance gene (AAC(3)Gm^r; Elisha & Steyn, 1991). The AAC(3) gene was also blunt-ended. These ligations resulted in the construction of pFDT220206Gm^r-pFDT220210Gm^r, whose sizes corresponded to 4.03 Kb, 3.75 Kb, 3.54 Kb, 3.43 Kb and 3.41 Kb, respectively (see *Figure 3.1*). The (*) designates the destroyed *Ava*II restriction site.

Complementation assays

The plasmids pFDT220206-pFDT220210 were unable to replicate autonomously in GW125a(*polA1*), but could be rescued *in trans* by pFDT200Gm^r (see Chapter 3, *Figure 3.1D*). The plasmids pFDT220206Gm^r-pFDT220209Gm^r were tested for complementation *in trans* by pCU1 or pNZ116 after transformation into competent GW125a(pCU1) and GW125a(pNZ116). The strains GW125a(pCU1) and GW125a(pNZ116) were made competent by the modified method of Chung & Miller (1988) (Chapter 2). Transformants containing both incoming and resident plasmids were selected and maintained on L-agar containing ampicillin and gentamycin.

Plasmid stability and incompatibility tests

The stability properties of the plasmids pNZ116 (IncFI) and pCU1 (IncN) were assessed in GW125a(*polA1*), as described previously (Chapter 2).

The incompatibility properties of the plasmid pFDT200Gm^r towards pNZ116 and pCU1 were investigated also as described in Chapter 2. Briefly, bacterial cells containing a resident plasmid were transformed with a second plasmid (incoming plasmid). Large incoming plasmids were introduced in the bacterial cells by conjugation. Hosts carrying both plasmids were maintained on double selective media. These hosts were subcultured as previously described in the absence of antibiotic selection or with selection for either the resident or the incoming plasmid. Aliquots of these cultures were screened for the presence of both plasmids at the generations indicated above. As before, the criteria of Meacock & Cohen (1980) were used to determine the degree of incompatibility.

RESULTS

Complementation of repß deletions by pCU1 and mini-F plasmids

Because of the similarity of RepB with the product of ORF239 of pCU1 and with the RepE of the mini-F plasmid (pNZ116), the deleted forms of RepB were tested for functional complementation with these plasmids. These results are shown in *Table 5.1.* Both pCU1 and pNZ116 were able to complement the plasmids containing the deleted forms of RepB (pFDT220206Gm^r-pFDT220208Gm^r) *in trans*. In addition, pNZ116 alone (not pCU1) was able to rescue pFDT220209Gm^r. The extent of functional complementation of deleted forms of the RepB protein by the other Rep proteins was unexpected, especially in the case of the RepE of mini-F, with which RepB shares only a 31% identity at the amino-acid level.

Table 5.1. Complementation of deleted forms of repß* by pCU1 and mini-F.

	Replication in GW125a (*polA1*)	
Plasmid	**with pCU1**	**with pNZ116 (mini-F)**
pFDT220206Gm^r	+	+
pFDT220207Gm^r	+	+
pFDT220208Gm^r	+	+
pFDT220209Gm^r	-	+
pFDT220210Gm^r	-	-

* These deleted forms of repß were unable to replicate autonomously in GW125(*polA1*).

Compatibility of repß with pCU1 and mini-F

Previously (Chapters 2 & 3) it had been show that the incompatibility determinants of repß were located between nucleotides 1605-2281 (pFDT210, *Figure 3.1*). This incompatibility was associated with the presence of the 18 iterons (Chapter 3, *Figures 3.2 & 3.2*). Classical incompatibility studies showed that the parent plasmid pGSH500 was compatible with both IncFI and IncN plasmids (Chapter 2). Furthermore, no significant hybridisation signal was obtained between either repα or repß with any of the replicon probes tested, though a signal was obtained between the IncN probe (pULB2432) and pGSH500 (Chapter 2). The compatibility of these plasmids could have been due to the presence of multiple replicons in pGSH500. Since deletions of pFDT200 (repß) could be complemented by pNZ116 and pCU1 (*Table 5.1*), this suggested that repß is compatible with these plasmids; the formal demonstration of the compatibility of these plasmids is described below.

The plasmid pFDT220204 that contains only 4 iterons was shown to elicit an incompatibility reaction towards pFDT200 (see Chapter 2, *Table 2.2*). Since both pFDT200 and pFDT220204 can form mini-replicons (*Appendix 2B*, pFDT200Gm^r and pFDT220204Gm^r, respectively), these were used in the incompatibility assays with pCU1 and pNZ116. The use of the mini-replicons, which lack Ap^r, was convenient as both pCU1 and pNZ116 contain the latter resistance marker.

The plasmids pCU1 and pNZ116 were stably maintained in GW125a(*polA1*) for 180 generations. In addition, the mini-replicons pFDT200Gm^r and pFDT220204Gm^r were stably maintained for more than 120 generations in the absence of selective pressure in a GW125a(*polA1*) containing pCU1, indicating that the ß replicon and the mini-replicon of pCU1 are fully compatible. The mini-replicons pFDT200Gm^r and pFDT220204Gm^r were also shown to be compatible with the mini-F plasmid. No loss of any of these plasmids was observed in GW125A(*polA1*) after 120 generations without selective pressure. No plasmid

displacing behaviour of any plasmid was noticed following selection favouring either the resident or the incoming plasmid.

DISCUSSION

The Rep proteins from both mini-pCU1 and mini-F (pNZ116) complemented deleted forms of repß (*Table 5.1*). The extent of complementation of repß differed with mini-F and pCU1. Complementation by mini-F required only an intact *ori*V in repß. In contrast, complementation by pCU1 required an additional 117 bp. Inspection of this region in *Figure 3.2* showed it to contain a DnaA box (inverted with respect to other present within the origin), IR2' and several 5 and 6 bp direct and inverted repeats. It is not clear which of these features is required for complementation by pCU1. Theoretically, pBR322 and pUC19 should incompatible. Therefore, the complementation of deleted forms of repß by F is entirely due to the f5 fragment in pNZ116.

Although these experiments show that complementation of repß by other plasmids is possible, this may be difficult to demonstrate in pGSH500, because of its "multi-replicon" organisation. Besides, interaction between repα and repß in pGSH500 cannot be excluded, although no evidence is available for this.

It is evident that the complementation and compatibility properties of repß are a consequence of its "hybrid organisation". Furthermore, at least in *E. coli*, repß is responsible for control of replication and incompatibility of pGSH500 (see Chapter 2 for a detailed discussion on the incompatibility properties of pFDT200 and pGSH500).

Irrespective of the evolutionary origin of repß, it is clear that its incompatibility group is novel. Analysis of the incompatibility group of repß using the *inc/rep* probes supplied by Dr Couturier (Couturier, *et al.*, 1988) showed no homology to known incompatibility

groups (Chapter 2). These results were confirmed by those obtained from classical incompatibility studies (Chapter 2) and by the demonstration of compatibility of repß with both F and pCU1 (this Chapter). The latter plasmids are frequently isolated from members of the *Enterobacteriaceae*. Since the host-range of repß includes members of hospital acquired *Enterobacteriaceae* and other nosocomial pathogens (Chapter 2), it is proposed that this replicon is especially adapted to replicate within this environment. Thus the survival of the moderately promiscuous plasmid pGSH500, which contains both α and ß replicons, does not depend on having a broad-host range, but rather on its capacity to avoid incompatibility reactions with other plasmids which co-inhabit the same hosts. The plasmid pGSH500 is therefore an example of a specialised group of plasmids which are adapted to the flora of specific environments (e.g. hospitals). In evolutionary terms, super-specialisation of an organism may lead to a risk of extinction. Yet, extinction of the multiple drug-resistant atypical promiscuous plasmid pGSH500 is unlikely to occur as long as antibiotic selective pressures and the family *Enterobacteriaceae* abound in hospital environments.

CHAPTER 6

GENERAL CONCLUSIONS

pGSH500 was obtained from a nosocomial isolate of *Klebsiella pneumoniae* (App) from the urine of a leukaemic patient, at Groote Schuur Hospital. Interest in this plasmid increased when it was associated with multiply resistant bacterial isolates (*E. coli, P. mirabilis*) from immuno-compromised patients. This suggested that pGSH500 was a conjugative plasmid encoding resistance to several antibiotics.

pGSH500 was experimentally transferred, by conjugation, from *K. pneumoniae* (App) to other members of the *Enterobacteriaceae* family. The recipients included *polA1* encoding and deficient *E. coli* strains, *C. diversus*, *M. morganii* and *P. mirabilis*. In addition, pGSH500 was also transmitted to *A. baumanii*. Conjugation of pGSH500 with either *Ps. aeruginosa* or *Ps. putida* was unsuccessful, however. This implied that the host-range of pGSH500 was intermediate between that of the narrow host-range plasmids and that of the typical promiscuous plasmids. Therefore, pGSH500 is a moderately promiscuous, *polA1* independent, plasmid.

Bacteriophage typing of strains APP(pGSH500) and J53(pGSH500) showed that the conjugative pili of pGSH500 are distinct from those of typical promiscuous plasmids.

Restriction enzyme analysis showed that pGSH500 is a large plasmid of 107 Kb. Spontaneous deletions of pGSH500, pGSH510-pGSH540, ranging in size from 76 to 10 Kb, were isolated from *E. coli* during one of the conjugation experiments.

The resistance genotype of pGSH500 was determined after antibiotic resistance assays were carried out in donor and transconjugant organisms. This showed that pGSH500 encoded resistance to amikacin, ampicillin, chloramphenicol, kanaycin, neomycin, netilmycin, penicillin, cotrimoxazole, tobramycin, tetracycline and trimethoprim.

The incompatibility group of pGSH500 is unknown and is most probably novel. pGSH500 is compatible with all plasmids tested, using classical incompatibility assays. In addition, its incompatibility group remained unclear despite the use of replicon-specific probes for "replicon typing".

Initially, the finding that two distinct fragments in pGSH500, pFDT100 & pFDT200, were capable of autonomous replication in *polA1* deficient *E. coli* hosts was unexpected. Yet, the incompatibility properties of pGSH500 may be due to interaction between these two replicons. Furthermore, the presence of two replicons raised the question of replicon dominance in pGSH500; that is, which replicon is responsible for the control of replication in pGSH500? This was resolved by establishing that the copy number of pFDT200 was the same as that of pGSH500, namely 3-5 copies per chromosome equivalent. Thus, the 3.3 Kb *Pst*I insert of pFDT200 contains the dominant or basic replicon (repß) of pGSH500. This 3.3 Kb DNA fragment was shown to contain all the necessary elements for regulation of plasmid replication and incompatibility. Thus, when ligated to an antibiotic resistance marker, in the absence of any cloning vector sequences, this fragment was able to form a true mini-replicon, pFDT200Gm^r. The possibility that the replicon contained in pFDT100 (repα) is dominant in other hosts has not been excluded.

The basic replicon of pGSH500 (repß) is contained within a 1.8 Kb DNA fragment (nucleotides 478-2281 of pFDT210) and includes 3 domains. The first domain is the *cis*-acting *ori*V. This is 245 bp in size (nucleotides 478-723) and is required *in cis* for plasmid replication. The second domain includes the gene for a 30.8 KDa protein (RepB) that is essential for replication, while the third domain consists of eighteen, 30-36 bp iterons. The

iterons are associated with incompatibility and copy number control of repß and were identified as the *inc/cop* domain. The minimal autonomously replicating fragment of repß includes the *cis-ori*V, the *repB* gene and 4 proximal iterons (nucleotides 478-1773).

The overall molecular organisation of repß is similar to that of the basic replicons of the narrow-host-range plasmids F, P1, pSC101 as well as that of the broad-host-range plasmid pCU1. The similar molecular organisation of these plasmids suggests that they may use the same host specified enzymes for replication.

Comparison of the DNA sequence of repß to that of other replicons showed a 53% identity to that of the mini-pCU1. This similarity is also reflected in the overall molecular organisation of both replicons. Analysis of the predicted amino-acid sequence of RepB showed a 58% identity to the Rep of mini-pCU1 and a 31% identity with RepE of mini-F plasmid. The latter was surprising, since no obvious similarity had been found between the *repB* and *repE* genes. Codon usage of RepB is similar to that of RepE, ORF239 (mini-pCU1) and of poorly expressed genes in *E. coli*.

The similarity between the amino acid sequence of RepB, Rep (mini-pCU1) and RepE (F) reflects a functional similarity. This was shown by the ability of both pCU1 and mini-F plasmid to rescue deletion mutants of RepB (pFDT220206Gm^r-pFDT220209Gm^r) *in trans*. The ability of Rep proteins from unrelated plasmids to complement each other may have an advantage *in vivo*, especially if the plasmids are compatible. Since repß and pGSH500 are compatible with both pCU1 and mini-F, the effects of mutations within repß may be complemented by a wide range of plasmids within the IncF and IncN groups. This may promote stability of plasmids co-infecting the same host cell.

DNA sequence analysis also showed an 80% identity between a 200 bp fragment of the *cis-ori*V region of repß with the equivalent region of *ori-2* of mini-F. This identity stopped abruptly before the start of the direct repeats within *ori-2* (*incB*). Perhaps the compatibility

between these two replicons results from the absence of any sequence similarity between the repß *cis-ori*V and the repeats of *ori-2* (*incB*).

The region of 80% sequence identity also includes an F/ColE1 primosome site (*rriA, ssiA*) outside of the *cis-ori*V of both replicons. In *ori-2*, this site is involved in replication initiation, which proceeds unidirectionally to the left. It is possible that the primosome site in repß has a similar function.

A 90% AT-rich sequence, containing 2 DnaA boxes is another feature common to both replicons. Since the DnaA boxes are essential for replication in both plasmids, their presence in equivalent sites of the replicon may imply similar involvement in the replisome complex.

Thus, based on the sequence analysis and molecular organisation, repß appears to have a "hybrid structure". It is possible that repß is a "hybrid" between the basic replicons of pCU1-like and F-like plasmids or that it may be an ancestral replicon form, from which others have been derived.

Irrespective of its evolutionary origin, repß is a most interesting replicon and may be a key factor in the successful maintenance of pGSH500 in a bacterial population. The compatibility properties of repß ensure that pGSH500 avoids incompatibility with related plasmids. Added to this, is the complementation of repß by other replicons, should mutations in the *repB* gene occur.

The contribution of repα to pGSH500 replication is unknown at the moment and this necessitates further investigation. Since repα is compatible and distinct from repß, it may become the dominant replicon in pGSH500 in the event of incompatibility of repß with other replicons. Furthermore, repα may be the dominant replicon in different hosts. In addition, the possibility that pGSH500 may contain other replicons cannot be excluded. In

consequence, the survival of pGSH500 may not be dependent on having a broad-host-range replicon, but rather on having at least two replicons with distinct host ranges. Broad-host-range replicons are large and require a complex network of operons. In fact it may be more efficient for a plasmid to possess a compact, stable replicon.

In conclusion, the moderately promiscuous, multiple drug-resistance, *polA1* independent, plasmid pGSH500 has found an ecological niche by adapting to the *Enterobacteriaceae* of hospital environments.

LITERATURE CITED

Abeles, A. L. (1986) P1 plasmid replication, purification and DNA-binding activity of the replication protein RepA. *J. Biol. Chem.* **261**, 3548-3555.

Abeles, A. L. & Austin, S. J. (1987). P1 plasmid requires methylated DNA. *EMBO J.* **6**, 3185-3189.

Abeles, A. & Austin, S. J. (1991). Antiparallel plasmid-plasmid pairing may control P1 plasmid replication. *Proc. Natl. Acad. Sci. USA.* **88**, 9011-9015.

Abeles, A. L., Friedman, S. A. & Austin, S. J. (1985). Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organisation of the P1 partition region. *J. Mol. Biol.* **85**, 261-272.

Abeles, A. L., Reaves, L. D. & Austin, S. J. (1989). Protein-DNA interactions in regulation of P1 plasmid replication. *J. Bacteriol.* **171**, 43-52.

Abeles, A. L., Reaves, L. D. and Austin, S. J. (1990). A single DnaA box is sufficient for initiation from the P1 plasmid origin. *J. Bacteriol.* **172,** 4386-4391.

Abeles, A., Snyder, K. & Chattoraj, D. (1984). P1 plasmid replication: replication structure. *J. Mol. Biol.* **173**, 307-324.

Aoki, T, Kitao, T. & Arai, T. (1977). R-plasmids in fish pathogens. In *Plasmids, medical and theoretical aspects* (Mitsuhashi, S., Rosival, L. & Krcmery, V., eds), pp 39-45, Springer-Verlag, Berlin.

Arai, T., Yao-Sgawa, K. & Aoki, T. (1980). The stability of R-plasmids in *Vibrio cholera. Keio J. Medicine*, **29**, 125-132.

Armstrong, K. A., Acosta, R., Ledner, E., Machida, Y. Pancotto, M., McCormick, M., Ohtsubo, H. & Ohtsubo, E. (1984). A 37 KDa plasmid-encoded protein is required for replication and copy number control in pSC101 and Ts derivative pHS1. *J. Mol. Biol.* **175**, 331-347.

Armstrong, K.A., Ohtsubo, H., Bauer, W. R., Yoshioka, Y., Miyazaki, C., Maeda, Y. & Ohtsubo, E. (1986). Characterisation of the gene products produced in minicells by pSM1, a derivative of R100. *Mol. Gen. Genet.* **205**, 56-65.

Austin, S. & Abeles, A. (1983) Partition of unit-copy mini-plasmids to daughter cells. I. P1- and F-mini-plasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *J. Mol. Biol.* **169**, 353-372.

Austin, S., Hart, F., Abeles, A. Sternberg, N. (1982). Genetic and physical map of a P1 miniplasmid. *J. Bacteriol.* **152**, 63-71.

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

Austin, S. Sternberg, N. & Yarmolinsky, M. (1978). Miniplasmids of bacteriophage P1. I Stringent plasmid replication does not require elements that regulate the lytic cycle. *J. Mol. Biol.* **120**, 297-309.

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

Bachi, B. & Arber, W. (1977). Physical mapping of *Bgl*II, *Bam*HI, *Eco*RI, *Hind*III and *Pst*I restriction fragments of bacteriophage P1 DNA. *Mol. Gen. Genet.* **153**, 311-324.

Bachman, B. J. (1987). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In *Escherichia coli and Salmonella typhimurium, cellular and molecular biology* (Neidhardt, F. C., Ingraham, J. L., Low, B. K., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds.), pp. 1190-1219, American Society for Microbiology, Washington.

Backman, K., Betlach, M. Boyer, H. W. & Yanofsky, S. (1978). Genetic and physical studies on the replication of ColE1-type plasmids. *Cold Spring Harbor Symp. Quant. Biol.* **43**, 69-76.

Barber, M. & Stokes, J. E. (1966). Antibiotic sensitivity tests. Broadsheet No 55. Association of Clinical Pathologists, London.

Barth, P. T. (1979). RP4 and R300B as wide host-range plasmid cloning vehicles. In *Plasmids of medical, environmental and commercial importance* (Timmis, K. N. & Pühler, A., eds), pp 399-410, Elsevier/North-Holland Biomedical Press, Amsterdam.

Barth, P. T., Ellis, K., Bechhofer, D. H. & Figurski, D. H. (1984). Involvement of *kil* and *kor* genes in the phenotype of a host mutant of RP4. *Mol. Gen. Genet.* **197**, 236-243.

Bechhofer, D. H. & Figurski, D. H. (1983). Map location and sequence of *korA*, a key regulatory gene of promiscuous plasmid RK2. *Nucleic Acids Res.* **11**, 7453-7469.

Bechhofer, D. H., Kornacki, J. A., Firsheim, W. & Figurski, D. H. (1986). Gene control in broad host range plasmid RK2: expression, polypeptide product and multiple regulatory functions of *korB*. *Proc. Natl. Acad. Sci. USA*, **83**, 394-398.

Bernardi, G. & Bernardi, G. (1986). Compositional constraints and genome evolution. *Annu. Rev. Genet.* **23**, 637-661.

Bergquist, P. L. (1987). Incompatibility. In *Plasmids, a practical approach* (Hardy, K. G., ed.), pp. 37-78, IRL Press, Oxford, Washington DC.

Bergquist, P. L., Downard, R. D., Caughey, P. A., Gardner, R. C. & Lane, H. E. D. (1981). Analysis of mini-F plasmid replication by transposition mutagenesis. *J. Bacteriol.* **147**, 888-899

Bergquist, P. L., Lane, H. E. D., Malcolm, L. & Downard, R. A. (1982). Molecular homology and incompatibility in the IncFI plasmid group. *J. Gen. Microbiol.* **128**, 223-238.

Bergquist, P. L., Saadi, S. & Maas, W. (1986). Distribution of basic replicons having homology with repFIA, repFIB and repFIC among IncF group plasmids. *Plasmid* **15**, 19-34.

Bex, F., Piérard, P., Desmyter, A., Drèze, P., Colet, M. & Couturier, M. (1986). Mini-F E protein: the carboxy-terminal end is essential for E gene repression and mini-F copy number control. *J. Mol. Biol.* **189**, 293-303.

Bex, F., Rokeach, L., Drèze, P., Garcia, L. & Couturier, M. (1983). Mini-F encoded proteins: identification of a new 10.5 KDa species. *EMBO J.* **2**, 1853-1861.

Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA*, **80**, 3963-3965.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.

Boice, L. & Luria, S. E. (1963). Behaviour of prophage P1 in bacterial matings. 1. Transfer of the defective prophage Pld1. *Virology*. **20**, 147-157.

Bonekamp, F. & Jensen, K. J. (1988). The AGG codon is translated slowly in *E. coli* even at very low expression levels. *Nucleic Acids Res.* **16**, 3013-3024.

Bradley, D. E., Coetzee, J. N., Bothma, T. & Hedges, R. W. (1981). Phage X: A plasmid-dependent broad host range, filamentous bacterial virus. *J. Gen. Microbiol.* **126**, 389-396.

Bradley, D. E. & Taylor, D. E. (1987). Transposition from RP4 to other replicons of a tellurite-resistance determinant not normally expressed by IncPα plasmids. *FEMS Microbiol. Lett.* **41**, 237-240.

Bramhill, D. & Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *Escherichia coli* chromosome. *Cell*, **54**, 915-918.

Braun, R. E., O'Day, K. and Wright, A. (1985). Autoregulation of the DNA replication gene *dnaA* in *Escherichia coli* K12. *Cell*, **40**, 159-169.

Brendler, T., Abeles, A. & Austin, S. (1991). Critical sequences in the core of the P1 plasmid replication origin. *J. Bacteriol.* **173**,3935-3942.

Burkardt, H., Riess, G. & Pühler, A. (1979). Relationship of group P1 plasmids revealed by heteroduplex experiments: RP1, RP4, R68 and RK2 are identical. *J. Gen. Microbiol.* **114**, 341-348.

Cabello, F., Timmis, K. & Cohen S. (1976). Replication control in a composite plasmid constructed by *in vitro* linkage of two distinct replicons. *Nature*, **259**, 285-290.

Capage, M. A. & Scott, J. (1983). SOS induction by P1 Km miniplasmids. *J. Bacteriol.* **155**, 473-480.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **134**, 1141-1156.

Chattoraj, D. K., Cordes, K. & Abeles, A. (1984). Plasmid P1 replication: negative control by repeated DNA sequences. *Proc. Natl. Acad. Sci. USA.* **81**, 6456-6460.

Chattoraj, D. K., Mason, R. J.. & Wickner, S. H. (1988) Mini-p1 plasmid replication: autoregulation-sequestration paradox. *Cell*, **52**, 551-557.

Chattoraj, D. K., Pal, S. K., Swack, J. A., Mason, R. J. & Abeles, A. L. (1985a). An autoregulatory protein is required for P1 plasmid replication. *UCLA Symp. Mol. Cell Biol. New Series*, **30**, 271-280.

Chattoraj, D. K., Snyder, K. M. & Abeles, A. L. (1985b). P1 plasmid replication. Multiple functions of RepA protein at the origin. *Proc. Natl. Acad. Sci. USA*, **82**, 2588-2592.

Cheah, U. E., Weigand, W. A. & Stark, B. C. (1987). Effects of recombinant plasmid size on cellular processes of *Escherichia coli*. *Plasmid*, **18**, 127-134.

Chikami, G. K., Guiney, D. G., Schmidhauser, T. J. & Helinski, D. R. (1985). Comparison of 10 IncP plasmids: homology in the regions involved in plasmid replication. *J. Bacteriol.* **162**, 656-660.

Chung, C. T. & Miller, R. H. (1988). A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res.* **16**, 3580.

Coetzee, J. N., Bradley, D. E., du Toit, L. & Hedges R. W. (1988). Bacteriophage X-2: a filamentous phage lysing IncX-plasmid-harbouring bacterial strains. *J. Gen. Microbiol.* **134**, 2535-2541.

Coetzee, J. N., Lecatsas, G. & Coetzee, W. F. (1979). Properties of R plasmid R772 and the corresponding pilus-specific phage PR772. *J. Gen. Microbiol.* **110**, 263-273.

Cohen, S. N. & Chang, A. C. Y. (1973). Recircularisation and autonomous replication of a sheared R-factor DNA segment in *Escherichia coli* transformants. *Proc. Natl. Acad. Sci. USA*, **70**, 1293-1297.

Collado-Vides, J., Magasanik, B. & Gralla, J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* **55**, 371-394.

Collins, J. & Pritchard, R. H. (1973). Relationship between chromosome replication and F'lac episome replication in *Escherichia coli*. *J. Mol. Biol.* **78**, 143-155.

Cooper, S. & Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *J. Mol. Biol.* **31**: 519-540.

Couturier, M., Bex, F., Bergquist, P. & Maas, W. (1988). Identification and classification of bacterial plasmids. *Microbiol. Rev.* **52**, 375-395.

Cross, M. A., Warne, S. R. & Thomas, C. M. (1986). Analysis of the vegetative replication origin of broad-host-range plasmid RK2 by transposon mutagenesis. *Plasmid*, **15**, 132-146.

Currier, T. C. & Morgan, M. K. (1981). Restriction endonuclease analysis of the incompatibility group P-1 plasmids RK2, RP1, RP4, R68 and R68.45. *Curr. Microbiol.* **5**, 323-327.

Datta, N. (1985). Plasmids as organisms. In *Plasmids in bacteria* (Helinski, D. R., Cohen, S. N., Clewell, D. B., Jackson, D. A. & Hollaender, A., eds), pp. 125-140, Plenum Press, New York.

Datta, N. & Hedges, R. W. (1971). Compatibility groups among fi⁻ R factors. *Nature*, **234**, 222-223.

Datta, N. & Hedges, R. W. (1972). R factors identified in Paris, some conferring gentamicin resistance, constitute a new compatibility group. *Ann. Inst. Pasteur (Paris)*, **123**, 849-852.

Datta, N., Hedges, R. W., Shaw, E. J., Sykes, R. B. & Richmond, M. H. (1971). Properties of an R-factor from *Pseudomonas aeruginosa*. *J. Bacteriol.*, **108**, 1244-1249.

Davey, R. B., Bird, P. I., Nikoletti, S. M., Praszkier, J. & Pittard J. (1984). The use of mini-gal plasmids for rapid incompatibility grouping of conjugative R plasmids. *Plasmid*, **11**, 234-242.

Davis, M. A. & Austin, S. J. (1988). Recognition of the P1 plasmid centromere analog involves binding of the *parB* protein and is modified by a specific host factor. *EMBO J.* **7**, 1881-1888.

de Haan, P. G. & Stouthamer, A. H. (1963). F-prime transfer and multiplication of sexduced cells. *Genet. Res.* **4**, 30-41.

de Lucia, P. & Cairns, J. (1969). Isolation of an *E. coli* strain with a mutation affecting DNA polymerase. *Nature*, **224**, 1164-1166.

Davidson, N., Deonier, R. C., Hu, S. & Ohtsubo, E. (1975). Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. X. Deoxyribonucleic acid sequence organisation of F and of F-primes, and the sequences involved in Hfr formation. In *Microbiology-1974* (Schlessinger, D. ed.), pp 56-65, American Society for Microbiology, Washington.

Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.

Donoghue, D. J. & Sharp, P. A. (1978). Replication of colicin E1 plasmid DNA in vivo requires no plasmid-encoded proteins. *J. Bacteriol.* **133**, 1287-1294.

Dorrington, R. A. & Rawlings, D. E. (1989). Identification and sequence of the basic replication region of a broad-host-range plasmid isolated from *Thiobacillus ferrooxidans*. *J. Bacteriol.* **171**, 2735-2739.

Durland, R. H. & Helinski, D. R. (1987). The sequence encoding the 43-kilodalton trfA protein is required for efficient replication or maintenance of minimal RK2 replicons in *Pseudomonas aeruginosa*. *Plasmid*, **18**, 164-169.

Durland, R. H. & Helinski, D. R. (1990). Replication of the broad-host-range plasmid RK2: direct measurements of intracellular concentrations of the essential TrfA replication proteins and their effect on plasmid copy number. *J. Bacteriol.* **172**, 3849-3858.

Durland, R. H., Toukdarian, A., Fang, F. & Helinski, D. R. (1990). Mutations in the *trfA* replication gene of the broad-host-range plasmid RK2 result in elevated plasmid copy numbers. *J. Bacteriol.* **172**, 3859-3867.

Ebbers, J & Eichenlaub, R. (1981). Complementation of replication-deficient deletion derivatives of plasmid mini-F. *J. Bacteriol.* **147**, 736-743.

Eichenlaub, R., Figurski, D. & Helinski, D. (1977). Bidirectional replication from a unique origin in a mini-F plasmid. *Proc. Natl. Acad. Sci. USA.* **74**, 1138-1141.

Eichenlaub. R. & Wehlmann, H. (1980). Amber-mutatants of plasmid mini-F defective in replication. *Mol. Gen. Genet.* **180**, 201-204.

Eichenlaub, R., Wehlman, H. & Ebbers, J. (1981). Plasmid mini-F encoded functions involved in replication and incompatibility. In *Molecular biology, pathogenicity, and ecology of bacterial plasmids* (Levy, S. B., Clowes, R. C. & Koenig, E. L., eds), pp 327-336, Plenum Press, New York.

Elisha, B. G. and Steyn, L. M. (1991). Cloning of AAC(3) and AAD(2") genes from *Acinetobacter*: Differential expression in the host strain. *Current Microbiol.* **22**, 259-263.

Engberg, B. & Nordström, K. (1975). Replication of R-factor R1 in *Escherichia coli* K-12 at different growth rates. *J. Bacteriol.* **123**, 179-186.

Filipski, J. (1991). Evolution of DNA sequences. Contribution of mutation bias and selection to the origin of chromosomal compartments. *Adv. Mutagenesis Res.* **2**, 1-54.

Figurski, D. & Helinski, D. R.. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc. Natl. Acad. Sci. USA*, **76**, 1648-1652.

Figurski, D. H., Meyer, R. J. & Helinski, D. R. (1979). Suppression of ColE1 replication properties by the IncP1 plasmid RK2 in hybrid plasmids constructed *in vitro*. *J. Mol. Biol.* **133**, 295-318.

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

Filutowicz, M., McEachern, M., Greener, A., Mukhopadhyay, P., Uhlenhopp, E., Durland, R. & Helinski, D. (1985). Role of the π initiation protein and direct nucleotide sequence repeats in the regulation of plasmid R6K replication. In *Plasmids in Bacteria* (Helinski, D. R., Cohen, S.N., Clewell, D. B., Jackson, D. A. & Hollaender, A., eds), pp. 125-140, Plenum Press, New York.

Filutowicz, M., McEachern, M. J., Mukhopadhyay, P., Greener, A., Yang, S. & Helinski, D. R. (1987). DNA and protein interactions in the regulation of plasmid replication. *J. Cell Sci. Suppl.* **7**, 15-31.

Foster, T. J. (1983). Plasmid-determined resistance to antimicrobial drugs and toxic metal ions in bacteria. *Microbiol. Rev.* **47**, 361-409.

Frame, R. & Bishop, J. O. (1971). The number of sex-factors per chromosome in *Escherichia coli*. *Biochem. J.* **121**, 93-103.

Frey, J. & Bagdasarian, M. (1989). The molecular biology of IncQ plasmids. In *Promiscuous plasmids of Gram-negative bacteria*. (Thomas, C. M. ed.), pp. 79-94, Academic Press. London, San Diego C. A.

Fuller, R. S., Funnell, B. E. & Kornberg, A. (1984). The DnaA protein complex with the *E. coli* chromosomal replication origin (*ori*C) and other DNA sites. *Cell*, **38**, 889-900.

Gamas, P., Burger, A. C., Churchward, G., Caro, L., Galas, D. & Chandler, M. (1986). Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. *Mol. Gen. Genet.* **204**, 85-89.

Gaylo, J. P., Turjman, N. & Bastia, D. (1987). DnaA protein is required for replication of the minimal replicon of the broad-host-range plasmid RK2 in *Escherichia coli*. *J. Bacteriol.* **169**, 4703-4709.

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

Goebel, W. & Schrempf, N. (1972). Replication of plasmid DNA in temperature sensitive DNA replication mutants of *Escherichia coli*. *Biochim. Biophys. Acta*, **262**, 32-41.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981). Translational initiation in prokaryotes. *Ann. Rev. Microbiol.* **35**, 365-403.

Grant, A. J., Bird, P. I. & Pittard, J. (1980). Naturally-occuring plasmids exhibiting incompatibility with members of incompatibility groups I and P. *J. Bacteriol.* **144**, 758-765.

Grantham, R., Gautier, C., Gouy, M. Jacobzone, M. & Mercier, R. (1981). Codon usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Res.* **9**, r43-r74.

Griffith, J., Hochschild, A. & Ptashne, M. (1986). DNA loops induced by cooperative binding of lambda repressor. *Nature*, **322**, 750-752.

Grindley, N. D. F. & Kelley, W. S. (1976). Effects of different alleles of *Escherichia coli* K-12 *polA* gene on replication of non-transferring plasmids. *Mol. Gen. Genet.* **143**, 311-318.

Grinter, N. J. (1984). Replication control of IncP plasmids. *Plasmid*, **11**, 74-81.

Grosschedl, R. & Hobom, G. (1979). DNA sequences and structural homologies of the replication origins of lambdoid bacteriophages. *Nature*, **277**, 621-627.

Guyer, M. S. (1978). The γ-δ sequence of F is an insertion sequence. *J. Mol. Biol.* **126**, 347-365.

Guyer, M. S., Figurski, D. & Davidson, N. (1976). Electron microscope study of a plasmid chimera containing the replication region of *Escherichia coli* F plasmid. *J. Bacteriol.* **127**, 988-997.

Guiney, D. G.. (1982). Host range of conjugation and replication functions of *Escherichia coli* sex plasmid F *lac*: Comparison with the broad host range plasmid RK2. *J. Mol. Biol.* **162**, 699-703.

Hansen, E. B. & Yarmolinsky, M. B. (1986). Host participation in the plasmid maintenance: dependence upon *dnaA* of replicons derived from P1 and F. *Proc. Natl. Acad. Sci. USA*. **83**, 4423-4427.

Haring, V. & Scherzinger, E. (1989). Replication proteins of the IncQ plasmid RSF1010. In *Promiscuous plasmids of Gram-negative bacteria* (Thomas, C. M., ed.), pp 95-124, Academic Press, London.

Hashimoto-Gotoh, T. & Sekiguchi, M. (1977). Mutations to temperature sensitivity in the R plasmid pSC101. *J. Bacteriol.* **131**, 405-412.

Hasunuma, K. & Sekiguchi, M. (1977). Replication of plasmid pSC101 in *Escherichia coli* K12: requirement for *dnaA* function. *Mol. Gen. Genet.* **154**, 225-230.

Hattman, S., Brooks, J. E. & Masurekar, M. (1978). Sequence specificity of the P1 modification methylase (M. *Eco* P1) and the DNA methylase (M. *Eco dam*) controlled by the *Escherichia coli dam* gene. *J. Mol. Biol.*, **126**, 367-380.

Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**, 2237-2255.

Hedges, R. W., Jacob, A. E., Barth, P. T. & Grinter, N. J. (1975). Compatibility properties of P1 and ϕAMP prophages. *Mol. Gen. Genet.* **141**, 263-267.

Helsberg, M. & 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.

Helsberg, M., Ebbers, J. & Eichenlaub, R. (1985). Mutations affecting replication and copy number control in plasmid mini-F both reside in the gene for the 29 KDa protein. *Plasmid*, **14**, 53-63.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351-369.

Higara, S., Jaffe, A., Ogura, T., Mori, H. & Takahashi, H. (1986) F plasmid ccd mechanism in Escherichia coli. *J. Bacteriol.* **166**, 100-104.

Higgin, D. G., Bleasby, A. H. & Fuchs, R. (1991). Improved software for multiple sequence alignment. *Comp. Appl. Biol. Sci.* **5**, 151-153.

Hinchliffe, E. & Vivian, A. (1980). Naturally occurring plasmids in *Acinetobacter calcoaceticus*: a P class R factor of restricted host range. *J. Gen. Microbiol.* **116**, 75-80.

Hogrefe, C. & Friedrich, B. (1984). Isolation and characterisation of megaplasmid DNA from lithoautotrophic bacteria. *Plasmid*, **12**, 161-169.

Holm, L. (1986). Codon usage and gene expression. *Nucleic Acids Res.* **14**, 3075-3087.

Imber, R., Low, R. L. & Ray, D. S. (1983). Identification of a primosome assembly site in the region of the *ori 2* replication origin of the *Escherichia coli* mini-F plasmid. *Proc. Natl. Acad. Sci. USA*, **80**, 7132-7136.

Inouye, M. (1988). Antisense RNA: its functions and applications in gene regulation - a review. *Gene*, **72**, 25-34.

Inselburg, J. (1974). Replication of colicin E1 plasmid DNA in minicells from a unique replication initiation site. *Proc. Natl. Acad. Sci. USA,* **71**, 2256-2259.

Ish-Horowicz, D. & Burke, J. F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9**, 2989-2998.

Ishii, K., Hashimoto-Gotoh, T. & Matsubara, K. (1978). Random replication and random assortment model for plasmid incompatibility. *Plasmid*, **1**, 435-446.

Itoh, Y., Watson, J. M., Haas, D. & Leisinger, T. (1984). Genetic and molecular characterisation of the *Pseudomonas* plasmid pVS1. *Plasmid*, **11**, 206-220.

Iyer, V.N. (1989). IncN group plasmids and their genetic systems. In *Promiscuous plasmids of Gram-negative bacteria*. (Thomas, C. M. ed.), pp. 165-183, Academic Press. London, San Diego C.A.

Jaffe, A., Ogura, T. & Higara, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *J. Bacteriol.* **163**, 841-849.

Jacob, A. E., Shapiro, J. A., Yamamoto, L., Smith, D. I., Cohen, S. N. & Berg, D. (1977). Plasmids studied in *Escherichia coli* and other enteric bacteria. In *DNA insertion*

elements, plasmids and episomes (Bukhari A. I., Shapiro, J. A. & Adhya, SL), pp 607-656, Cold Spring Harbor Laboratory, New York.

Jacob, F., Brenner, S. & Cumin, F. (1963). On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.* **28**, 329-348.

Kamio, Y., Itoh, Y. & Terawaki, Y. (1988). Purification of Rts1 RepA protein and binding of the protein to mini-Rts1 DNA. *J. Bacteriol.* ***170***, 4411-4414.

Kamio, Y., Tabuchi, A., Itoh, Y., Katagiri, H. & Terawaki, Y. (1984). Complete nucleotide sequence of the mini-Rtsl and its copy mutant. *J. Bacteriol.* **158**, 307-312.

Kelley, W. & Bastia, D. (1985). Replication intitator protein of R6K autoregulates its own synthesis at the transcriptional step. *Proc. Natl. Acad. Sci. USA*, **82**, 2574-2578.

Keasling, J. D., Palsson, B. O. & Cooper, S. (1991). Cell-cycle-specific F plasmid replication: Regulation by cell size control of initiation. *J. Bacteriol.* **173**, 2673-2680.

Khatoon, H. & Iyer, V. N. (1971). Stable coexistence of R*fi*$^-$ factors in *Escherichia coli*. *Can. J. Microbiol.* **17**, 669-675.

Kim, K. & Meyer, R. (1986). Copy-number of broad host range plasmid R1162 is determined by the amounts of essential plasmid-encoded proteins. *J. Mol. Biol.* **185**, 755-767.

Kingsbury, D. T. & Helinski, D. R. (1970). DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor. *Biochem. Biophys. Res. Commun.* **4**, 1538-1544.

Kingsbury, D. T. & Helinski, D. R. (1973). Temperature sensitive mutants for the replication of plasmids in *Escherichia coli*: requirement for deoxyribonucleic acid polymerase I in the replication of the plasmid ColE1. *J. Bacteriol.* **114**, 1116-1124.

Kittell, B. L. & Helinski, D. R. (1991). Iteron inhibition of plasmid RK2 replication *in vitro*: Evidence for intermolecular coupling of replication origins as a mechanism for RK2 replication control. *Proc. Natl. Acad. Sci. USA.* **88**, 1389-1393.

Kline, B. C. (1985). A review of mini-F plasmid maintenance. *Plasmid.* **14**, 1-16.

Kline, B. C. (1988). Aspects of plasmid F maintenance in *Escherichia coli*. *Can. J. Microbiol.* **34**, 526-535.

Kline, B. C., Sandhu, G. S., Eckloff, B. W. & Aleff, R. A. (1992). Site-specific proteolysis of mini-F plasmid replication protein RepE destroys initiator function and generates an incompatibility substance. *J. Bacteriol.* **174**, 3004-3010.

Kline, B. C., Kogoma, T., Tam, J. E. & Shields, M. S. (1986). Requirement of the *Escherichia coli dnaA* gene product for plasmid F maintenance. *J. Bacteriol.* **168**, 440-443.

Kline, B. C. & Seelke, R. (1982). Genetic evidence that control of F replication is negative. *Mol. Gen. Genet.* **187**, 218-224.

Kline, B. C. & Trawick, J. (1983). Identification and characterisation of a second copy number control gene in mini-F plasmids. *Mol. Gen. Genet.* **192**, 408-415.

Kollek, R., Oertel, M. & Goebel, W. (1978). Isolation and characterisation of the minimal fragment required for autonomous replication of a copy mutant (PKN102) of the antibiotic resistance factor R1. *Mol. Gen. Genet.* **162**, 51-57.

Kolter, R. & Helinski, D. R. (1979) Regulation of initiation of DNA replication. *Ann. Rev. Genet.* **13**, 355-392.

Kolter, R. (1981). Replication properties of plasmid R6K. *Plasmid*, **5**, 2-9.

Komai, N., Nishizawa, T., Hayakawa, Y., Murotsu, T & Matsubara, K. (1982). Detection and mapping of six mini F-encoded proteins by cloning analysis of dissected mini-F segments. *Mol. Gen. Genet.* **186**, 193-203.

Konarska-Kozlowska, M. & Iyer, V. N. (1981a). The genetic and physical basis of variability in *Escherichia coli* strains carrying a reference IncN group plasmid. *Can. J. Microbiol.* **27**, 616-626.

Konarska-Kozlowska, M & Iyer, V. N. (1981b). Physical and genetic organisation of the IncN-group plasmid pCU1. *Gene*, **14**, 195-204.

Konarska-Kozlowska, M. & Iyer, V. N. (1983). Sequence homology between IncN group plasmids. *Plasmid*, **10**, 211-223.

Konarska-Kozlowska, M., Thatte, V. & Iyer, V. N. (1983). Inverted repeats in the DNA of plasmid pCU1. *J. Bacteriol.* **153**, 1502-1512.

Kontomichalou, P., Mitani, M. & Clowes, R. C. (1970). Circular R-factor molecules, controlling penicillinase synthesis, replicating in *Escherichia coli* under relaxed or stringent control. *J. Bacteriol.* **104**, 34-44.

Kornacki, J. A., Balderes, P. J. & Figurski, D. H. (1987). Nucleotide sequence of *korB,* a replication control gene of broad host-range plasmid RK2. *J. Mol. Biol.* **198**, 211-222.

Kornacki, J. A., Burlage, R. S. & Figurski, D. H. (1990). The *kil-kor* regulon of broad-host-range plasmid RK2: nucleotide sequence, polypeptide product and expression of regulatory gene *korC. J. Bacteriol.* **172**, 3040-3050.

Kornacki, J. A., West, A. H. & Firshein, W. (1984). Proteins encoded by the *trans*-acting replication and maintenance regions of broad host range plasmid RK2. *Plasmid*, **11**, 48-57.

Korneluk, R. G., Quan, F. & Gravel, J. F. (1985). Rapid and reliable dideoxy sequencing for double stranded DNA. *Gene*, **40**, 317-323.

Kozak, M. (1983). Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiol. Rev.* **47**, 1-45.

Kozlowski, M., Thatte, V., Lau, P. C. K., Visentin, L. P. & Iyer, V. N. (1987). Isolation and structure of the replicon of the promiscuous plasmid pCU1. *Gene*, **58**, 217-228.

Krämer, H., Niemöller, M., Amouyal, M., Revet, B., von Wilckern-Bergmann, B. & Müller-Hill, B. (1987). *Lac* repressor forms loops with linear DNA carrying two suitable spaced *lac* operators. EMBO *J.* **6**, 1481-1491.

Krishnan, B. R., Fobert, P. R., Seitzer, U. & Iyer, V. N. (1990). Mutations within the replicon of the IncN plasmid pCU1 that affect its *Escherichia coli polA*-independence but not its autonomous replication ability. *Gene*, **91**, 1-7.

Krishnan, B. R. & Iyer, V. N. (1988). Host ranges of the IncN group plasmid pCU1 and its minireplicon in Gram-negative purple bacteria. *App. Environ. Microbiol.* **54**, 2273-2276.

Krishnan, B. R. & Iyer, V. N. (1990). IncN Plasmid replicon, a deletion and subcloning analysis. *J. Mol. Biol.* **213**, 777-788.

Krishnapillai, V. (1986). Genetic analysis of bacterial plasmid promiscuity. *J. Genet.* **65**, 103-120.

Krishnapillai, V. (1988). Molecular genetic analysis of bacterial promiscuity. *FEMS Microbiol. Rev.* **54**, 223-238.

Kües, U. & Stahl, U. (1989). Replication of plasmids in Gram-negative bacteria. *Microbiol. Rev.* **53**, 491-516.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lane, H. E. D. (1981). Replication and incompatibility of F and plasmids in the IncFI group. *Plasmid*, **5**, 100-126.

Lane, D. & Gardner, R. C. (1979). Second *Eco*RI fragment of F capable of self-replication. *J. Bacteriol.* **139**, 141-151.

Lane, D., Hill, D., Caughey, P. & Gunn, P. (1984). The mini-F primary origin. Sequence analysis and multiple activities. *J. Mol. Biol.* **180**, 267-282.

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

Lanka, E., Lurz, R. & Fürste, J. P. (1983). Molecular cloning and mapping of *Sph*I restriction fragments of plasmid RP4. *Plasmid*, **10**, 303-307.

Leong, J. M., Nunes-Duby, S., Lesser, C. F., Younderian, P., Susskind, M. M. & Landy, A. (1985). The ϕ80 and P22 attachment sites. Primary structure and interaction with ***E. coli*** integration host factor. *J. Biol. Chem.* **260**, 4468-4477.

Lewin, B. (1990). *Genes IV.* 857 pp, Oxford University Press, Oxford.

Light, J. & Molin, S. (1983). Post-transcriptional control of expression of the *repA* gene of plasmid R1 mediated by a small RNA molecule. *EMBO J.* **2**, 93-98.

Lin, L.-S. & Meyer, R. J. (1986). Direct repeated, 20-bp sequence of plasmid R1162. *Nucleic Acids Res. Mol. Gen. Genet.* **194**, 423-431.

Lin, L.-S., Kim, Y.-J. & Meyer, R. J. (1987). The 20 bp, direct repeated DNA sequence of broad-host-range plasmid R1162 exerts incompatibility *in vivo* and inhibits R1162 replication *in vitro*. *Mol. Gen. Genet.* **208**, 390-397.

Linder, P., Churchward, G. Guixian, X., Yu, Y-Y. & Caro, L. (1985). An essential replication gene, *repA*, of plasmid pSC101 is autoregulated. *J. Mol. Biol.* **181**, 383-393.

Lodge, J., Williams, R., Bell, A., Chan, B. & Busby, S. (1990). Comparison of promoter activities in *Escherichia coli* and *Pseudomonas aeruginosa*: use of a new broad-host-range promoter-probe plasmid. *FEMS Microbiol. Lett.* **67**, 221-226.

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

López, J., Crespo, P., Rodríguez, J. C., Andrés, I. & Ortiz, J. M. (1989a). Analysis of IncF plasmids evolution: nucleotide sequence of an IncFIII replication origin. *Gene*, **78**, 183-187.

López, J., Rodríguez, J. C., Andrés, I. & Ortiz, J. M. (1989b). Characterisation of the repFVII replicon of the haemolytic plasmid pSU233: nucleotide sequence of an *incFVII* determinant. *J. Gen. Microbiol.* **135**, 1763-1768.

López, J., Delgado, D., Andrés, I., Ortiz, J. M. & Rodríguez, J. C. (1991). Isolation and evolutionary analysis of a repFVIB replicon of the plasmid pSU212. *J. Gen. Microbiol.* **137**, 1093-1099.

Lovett, M. A. & Helinski, D. (1976). Method for isolation of the replication of a bacterial replicon: construction of a mini-F*Km* plasmid. *J. Bacteriol.* **127**, 982-987.

Lundquist, P. D. & Levin, B. R. (1986). Transitory derepression and the maintenance of conjugative plasmids. *Genetics*, **113**, 483-497.

Maas, R. & Maas, W. (1962). Introduction of a gene from *Escherichia coli* B into Hfr and F⁻ strains of *Escherichia coli* K12. *Proc. Natl. Acad. Sci.* **48**, 1887-1893.

Maas, R., Saadi, S. & Maas, W. R. (1989). Properties and incompatibility behaviour of mini-plasmids derived from the bireplicon plasmid pCG86. *Mol. Gen. Genet.* **181**, 8-12.

Maki, S., Kuribayashi, M., Miki, T. & Horiuchi, T. (1983). An amber replication mutant of F plasmid mapped in the minimal replication region. *Mol. Gen. Genet.* **191**, 231-237.

Mandel, M. & Higa, A. (1970). Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**, 159-162.

Manis, J. J. & Kline, B. C. (1977). Restriction endonuclease mapping and mutagenesis of the F sex factor replicating region. *Mol. Gen. Genet.* **152**, 175-182.

Marians, K. J., Soeller, M. & Zipursky, S. L. (1982). Maximal limits of the *Escherichia coli* replication factor Y effector site sequences in pBR322 DNA. *J. Biol. Chem.* **257**, 5656-5662.

Martin-Parras, L.M., Hernández, P., Martínez-Robles, M. L. & Schvartzman, J. B. (1991). Unidirectional replication as visualised by two-dimentional agarose gel electrophoresis. *J. Mol. Biol.* **220**, 843-853.

Maruyama, T., Gojobori, T., Aota, S.-I. & Ikemura, T. (1986). Codon usage tabulated from GenBank genetic sequence data. *Nucleic Acids Res.* **14** (Supplement), r151-r197.

Masai, H. & Arai, K. (1987). RepA and DNA proteins are required for initiation of R1 plasmid replication *in vitro* and interact with the *oriR* sequence. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4781-4785.

Masai, H. & Arai, K. (1988). Initiation of lagging-strand synthesis for pBR322 plasmid DNA replication *in vitro* is dependent on primosomal protein i encoded by *dnaT*. *J. Biol. Chem.* **263**, 15016-15023.

Masai, H. & Arai, K. (1989). *Escherichia coli dnaT* gene function is required for pBR322 plasmid replication but not for R1 plasmid replication. *J. Bacteriol.* **171**, 2975-2980.

Masai, H., Nomura, N., Kubota, Y. & Arai, K. (1990). Roles of ϕX174 type primosome- and G4 type primase-dependent primings in initiation of lagging and leading strand synthesis of DNA replication. *J. Biol. Chem.* **265**, 15124-15133.

Masson, L. & Ray, D. S. (1986). Mechanism of autonomous control of the *Escherichia coli* F plasmid: different complexes of the initiator/repressor protein are bound to its operator and to an F plasmid origin. *Nucleic Acids Res.* **14**, 5693-5711.

Masson, L. & Ray, D. S. (1988). Mechanism of autonomous control of *Escherichia coli* F plasmid: purification and characterisation of the *repE* gene product. *Nucleic Acids Res.* **16**, 413-425.

Matsubara, K. (1976). Genetic structure of replication of plasmid *Kdv*. *J. Mol. Biol.* **102**, 427-439.

McClure, W. R. (1985). Mechanisms and control of transcription initiation in prokaryotes. *Ann. Rev. Biochem.* **54**, 171-204.

McEachern, M. J., Bott, M. A., Tooker, P. A. & Helinski, D. R. (1989). Negative control of plasmid R6K replication: possible role of intermolecular coupling of replication origins. *Proc. Natl. Acad. Sci. USA*, **86**, 7942-7946.

McEachern, M. J., Filutowicz, M., Yang, S., Greener, A., Mukhopadhyay, P. & Helinski, D. R. (1986). Elements involved in the copy number of the antibiotic resistance plasmid R6K. In *Antibiotic resistance genes: Ecology, transfer, and expression.* (Banbury Report 24), pp 195-204, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

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

Meyer, R. & Helinski, D. R. (1977). Unidirectional replication of the P-group plasmid RK2. *Biochim. Biophys. Acta.* **478**, 109-113.

Meyer, R. & Hinds, M. (1982). Multiple mechanisms for expression of incompatibility by broad-host-range plasmid RK2. *J. Bacteriol.* **152**, 1078-1090.

Meyer, R. J., Lin, L.-S., Kim, K. & Brash, M. A. (1985). Broad host-range plasmid R1162: replication incompatibility and copy number control. In *Plasmids in bacteria* (Helinski, D. R., Cohen, S. N., Clewell, D. B., Jackson, D. A. & Hollaender, A., eds), pp 173-188, Plenum Publishing Corporation, New York.

Meynell, E. & Datta, N. (1966). The relation of resistance transfer factors to the F-factor (sex-factor) of *Escherichia coli* K12. *Genet. Res.* **7**, 134-140.

Miki, T., Easton, A. M. & Rownd, R. H. (1980). Cloning of replication, incompatibility and stability functions of R plasmid NR1. *J. Bacteriol.* **141**, 87-99.

Miki, T., Yashioka, K.. & Horiuchi, T. (1984). 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.

Millman, G. & Hwang, E. S. (1987). Epstein Barr virus nuclear antigen (EBNA-1) forms a complex that binds with high concentration dependence to a single DNA binding site. *J. Virol.* **61**, 465-471.

Monti-Bragadin, C. & Samer, L. (1975). Compatibility of pTM89, a new F-like R-factor, and of derivative plasmids. *J. Bacteriol.* **124**, 1132-1136.

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

Mukherjee, S., Erickson, H. & Bastia, D. (1988). Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. *Cell*, **52**, 375-383.

Mukherjee, S., Patel, I. & Bastia, D. (1985). Conformational changes in a replication origin induced by an initiator protein. *Cell*, **43**, 189-197.

Muraiso, K., Tokino, T., Murotsu, T. & Matsubara, K. (1987). Replication of mini-F plasmid *in vitro* promoted by purified E protein. *Mol. Gen. Genet.* **206**, 519-521.

Murakami, Y. Ohmori, H., Yura, T. & Nagata, T. (1987). Requirement of *Escherichia coli dnaA* gene function for *ori-2*-dependent mini-F plasmid function. *J. Bacteriol.*, **169**, 1724-1730.

Murotsu, T. & Matsubara, K. (1980). Role of an autorepression system in the control of lambda *dv* plasmid copy number and incompatibility. *Mol. Gen. Genet.* **179**, 509-519.

Murotsu, T., Matsubara, K., Sugisaki, H. & Takanami, M. (1981). Nine unique repeating sequences in a region essential for replication and incompatibility of the mini-F plasmid. *Gene*, **15**, 257-271.

Murotsu, T., Tsutsui, H. & Matsubara, K. (1984). Identification of the minimal essential region for the replication origin of mini-F plasmid. *Mol. Gen. Genet.* **196**, 373-378.

Muto, A. & Osawa, S. (1987). The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc. Natl. Acad. Sci. USA*, **84**, 166-169.

Nash, J. & Krishnapillai, V. (1987). DNA sequence analysis of host range mutants of the promiscuous IncP-1 plasmids R18 and R68 with Tn*7* insertions in *ori*V. *Plasmid*, **18**, 35-45.

Nash, J. & Krishnapillai, V. (1988). Role of IncP1 plasmid primase in conjugation between *Pseudomonas* species. *FEMS Microbiol. Lett.* **49**, 257-260.

Nomura, N., Low, R. L. & Ray, D. S. (1982). Selective cloning of ColE1 DNA initiation sequences using the cloning vector M13 delta E101. *Gene*, **18**, 239-246.

Nomura, N., Masai, H., Inuzuka, M., Miyazaki, C., Ohtsubo, E., Itoh, T., Sasamoto, S., Matsui, M., Ishizaki, R. & Arai, K.-I. (1991). Identification of eleven single-strand initiation sequences (*ssi*) for priming of DNA replication in the F, R6K, R100 and ColE2 plasmids. *Gene*, **108**, 15-22.

Nordström, K. (1985). Control of plasmid replication: theoretical considerations and practical solutions. In *Plasmids in Bacteria* (Helinski, D., Cohen, S. N., Clewell, D. B., Jackson, D. A. & Hollaender, A., eds.), pp 189-214, Plenum Press, New York, London.

Nordström, K. (1990). Control of plasmid replication - how do DNA iterons set the replication frequency? *Cell*, **63**, 1121-1124.

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

Nordström, K., Molin, S. & Aagard-Hansen, H. (1980). Partitioning of plasmid R1 in *Escherichia coli*. I. Kinetics of loss of plasmid derivatives deleted of the par region. *Plasmid*, **4**, 215-227.

Nordström, K., Molin, S. & Light, J. (1984). Control of replication of bacterial plasmids: genetics, molecular biology, and physiology of the plasmid R1 system. *Plasmid*, **12**, 71-90.

Nordström, M & Nordström, K. (1985). Control of replication of IncFII plasmids: Comparison of the basic replicons and of the small *copB* systems of plasmids R100 and R1. *Plasmid*, **31**, 81-87.

Novick, R. P. (1969). Extrachromosomal inheritance in bacteria. *Bacteriol. Rev.* **33**, 210-263.

Novick, R. P. (1987). Plasmid Incompatibility. *Microbiol. Rev.* **51**, 381-395.

Novick, R. P., Clowes, R. C., Cohen, S. N., Curtiss III, R, Datta, N. & Falkow, S. (1976). Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol. Rev.* **40**, 168-189.

Nugent, M. E., Ellis, K., Datta, N. & Bradley, D. E. (1982). pHH502, a plasmid with IncP and IncIα characters, loses the latter by a specific *recA*-independent deletion event. *J. Gen. Microbiol.* **128**, 2781-2790.

Oka, A., Nomura, N., Morita, M., Sugisaki, H., Sugimoto, K. & Takanami, M. (1979). Nucleotide sequence of small ColE1 derivatives: structure of the regions essential for autonomous replication and colicin E1 immunity. *Mol. Gen. Genet.* **172**, 151-159.

Ogasawara, N, S., Moriya, K., von Meyenburg, F. G., Hanse, F. G. & Yoshikawa, H. (1985). Conservation of genes and their organisation in the chromosomal replication origin of *Bacillus subtilis* and *Escherichia coli*. *EMBO J.* **4**, 3345-3350.

Ogura, T. & Higara, S. (1983a). Mini-F genes that couple host cell division to plasmid proliferation. *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 4784-4788.

Ogura, T & Higara, S. (1983b). Partition mechanism of F plasmid: two gene-encoded products and a *cis*-acting region are involved in partition. *Cell*, **32**, 351-360.

Öhman, M. & Wagner, E. G. (1991). Regulation of replication of plasmid R1: An analysis of the intergenic region between *copA* and *repA*. *Mol. Gen. Genet.* **230**, 321-328.

Osawa, S., Jukes, T. H., Watanabe, K. & Muto, A. (1992). Recent evidence for evolution of the genetic code. *Microbiol. Rev.* **56**, 229-264.

Pabo, C. O. & Sauer, R. T. (1984). Protein-DNA recognition. *Ann. Rev. Biochem.* **17**, 203-321.

Pal, S. K. & Chattoraj, D. K. (1988). P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. *J. Bacteriol.* **170**, 3554-3560.

Pal, S., Mason, J. and Chattoraj, D. K. (1986). P1 plasmid replication: Role of initiator titration in copy number control. *J. Mol. Biol.*, **192**, 275-285.

Pansegrau, W. & Lanka, E. (1987). Conservation of a common "backbone" in the genetic organisation of the IncP plasmids RP4 and R751. *Nucleic Acids Res.* **15**, 2385.

Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA*, **85**, 2444-2448.

Persson, C. & Nordström, K. (1986). Control of replication of the host range plasmid RSF1010: the incompatibility determinant consists of direct repeated DNA sequences. *Mol. Gen. Genet.* **203**, 189-192.

Picken, R. N., Mazaitis, A. J., Saadi, S. & Maas, W. (1984). Characterisation of the basic replicons of the chimeric R/Ent plasmid pCG86 and the related Ent plasmid P307. *Plasmid* **12**, 10-18.

Pinkney, M., Diaz, R., Lanka, E. & Thomas, C. M. (1988). Replication of mini-RK2 plasmid in extracts of *Escherichia coli* requires plasmid-encoded protein TrfA and host-encoded proteins DnaA, B, G, DNA gyrase and DNA polymerase III. *J. Mol. Biol.* **203**, 927-938.

Pinkney, M., Theophilus, B. D. M., Warne, S. R., Tacon, W. C. A. & Thomas, C. A. (1987). Analysis of transcription from the *trfA* promoter of broad host range plasmid RK2 in *Escherichia coli, Pseudomonas putida* and *Pseudomonas aeruginosa. Plasmid*, **17**, 222-232.

Pohlman, R. F. & Figurski, D. H. (1983). Essential genes of plasmid RK2 in *Escherichia coli*: *trfB* controls a *kil* gene near *trfA*. *J. Bacteriol.* **156**, 584-591.

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

Pritchard, R. H. & Grover, B. (1981). Control of plasmid replication and its relationship to incompatibility. In *Molecular biology, pathogenicity and ecology of bacterial plasmids* (Levy, S. B., Clowes, S. N. & Koenig, E. L., eds.), pp. 271-278, Plenum Press, N.Y.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at distance. *Nature*, **322**, 697-701.

Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Millican, A., Eaton, M. & Humphreys, G. (1984). Codon usage can affect the efficiency of translation of genes in *Escherichia coli*. *Nucleic Acids Res.* **12**, 6663-6671.

Rokeach, L. A., Sϕgaard-Andersen, L. & Molin, S. (1985). Two functions of the *E* protein are key elements in the plasmid F replication control system. *J. Bacteriol.* **164**, 1262-1270.

Rosen, J., Ryder, R., Inokuchi, H., Ohtsubo, H. & Ohtsubo, E. (1980). Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. *Mol. Gen. Genet.* **179**, 527-537.

Rosenberg, M. & Court, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. *Ann. Rev. Genet.* **13**, 319-353.

Rosner, J. L. (1972). Formation, induction and curing of bacteriophage P1 lysogens. *Virology.* **48**, 679-689.

Ryder, T. B., Davidson, D. B., Rosen J. I., Ohtsubo, E. & Ohtsubo, H. (1982). Analysis of plasmid genome evolution based on nucleotide sequence. Comparison of two related plasmids of *Escherichia coli*. *Gene*, **17**, 299-310.

Ryder, T., Rosen, J., Armstrong, K., Davidson, D., Ohtsubo, E. Ohtsubo, H. (1981). Dissection of the replication region controlling incompatibility, copy number and initiation of DNA synthesis in the resistance plasmids R100 and R1. *ICN-UCLA Symp. Mol. Cell Biol.* **22**, 91-109.

Saadi, S., Maas, W. K. & Bergquist, P. L. (1984). RepFIC, a basic replicon of IncF plasmids that has homology with the basic replicon of IncFII plasmids. *Plasmid*, **12**, 61-64.

Saadi, S., Maas, W. K., Hill, D. F. & Bergquist, P. L. (1987). Nucleotide sequence analysis of repFIC, a basic replicon present in IncFI plasmid P307 and F, and its relation to the repA replicon of IncFII plasmids. *J. Bacteriol.* **169**, 1836-1846.

Saul, D., Spiers, A. J., McAnulty, Gibbs, M. G., Bergquist, P. L. & Hill, D. F. (1989). Nucleotide sequence and replication characteristics of repFIB, a basic replicon of IncF plasmids. *J. Bacteriol.* **171**, 2697-2707.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A laboratory manual* (Second edition), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

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

Scaife, J. & Gross, J. D. (1962). Inhibition of multiplication of an F-lac factor in Hfr cells of *Escherichia coli* K12. *Biochem. Biophys. Res. Commun.* **7**, 403-407.

Schmidhauser, T. J., Filutowicz, M. & Helinski, D. R. (1983). Replication of derivatives of the broad host range plasmid RK2 in two distantly related bacteria. *Plasmid*, **9**, 325-330.

Schmidhauser, T. J. & Helinski, D. R. (1985). Regions of broad-host-range plasmid RK2 involved in replication and stable maintenance in nine species of gram-negative bacteria. *J. Bacteriol.* **164**, 446-455.

Schreiner, H. C., Bechhofer, D. H., Pohlman, R. F., Young, C., Borden, P. A. & Figurski, D. H. (1985). Replication control in promiscuous plasmid RK2: *Kil* and *kor* functions affect expression of the essential replication gene *trfA*. *J. Bacteriol.* **163**, 228-237.

Schvartzman, J. B., Adolph, S., Martin-Parras, L.M. & Schildkraut, C. L. (1990). Evidence that replication initiates at only some of the potential origins in each oligomeric form of bovine papilloma virus type I DNA. *Mol. Cell. Biol.* **10**, 3078-3086.

Scott, J. R. (1984). Regulation of plasmid replication. *Microbiol. Rev.* **48**, 1-23.

Scott, J. R. & Vapnek, D. (1980). Regulation of replication of the P1 plasmid prophage. In *Mechanistic studies of DNA replication and genetic recombination* (Alberts, B., ed.), pp 335-345, Academic Press, New York.

Seelke, R. W. & Kline, B. C. (1984). A mutational hot spot in the incompatibility gene *incC* on mini-F plasmid. *Plasmid*, **12**, 37-40.

Seelke, R. W., Kline, B. C., Trawick, J. D. & Ritts, G D. (1982). Genetic studies of F plasmid maintenance genes involved in copy number control, incompatibility and partitioning. *Plasmid*, **7**, 163-179.

Seth, A. (1984). A new method for linker ligation. *Gene Anal. Tech.* **1**, 99-103.

Shafferman, A., Geller, T. & Hertman, I. (1979). Identification of the P1 compatibility and plasmid maintenance locus by mini-P1 lac$^+$ plasmid. *Virology*, **96**, 32-37.

Shafferman, A., Kolter, R., Stalker, D. M. & Helinski, D. R. (1982). Plasmid R6K DNA replication. III. Regulatory properties of the π initiation protein. *J. Mol. Biol.* **161**, 57-76.

Sharp, P. A., Hsu, M.-T., Ohtsubo, E. & Davidson, N. (1972). Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. I. Structure of F-prime factors. *J. Mol. Biol.* **71**, 471-479.

Sharp, P. M. & Li, W.-H. (1986). Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for "rare" codons. *Nucleic Acids Res.* **14**, 7737-7749.

Shine, J. & Dalgarno, L. (1974). The 3' terminal sequence of *E. col* 16S ribosomal RNA: Complementary to nonsense triplets and ribosomal binding site. *Proc. Natl. Acad. Sci. USA*, **71**, 1342-1346.

Shine, J. & Dalgarno, L. (1975). Determination of cistron specificity in bacterial ribosomes. *Nature*, **254**, 34-38.

Shingler, V. & Thomas, C. M. (1984). Analysis of the *trfA* region of broad-host-range plasmid RK2 by transposon mutagenesis and identification of polypeptide products. *J. Mol. Biol.* **175**, 229-249.

Shon, M., Germino, J. & Bastia, D. (1982). The nucleotide sequence of the replication origin of the plasmid R6K. *J. Biol. Chem.* **257**, 13823-13827.

Simons, R. W. (1988). Naturally occurring antisense RNA control - a brief review. *Gene*, **72**, 35-44.

Skovgaard, O. & Hansen, F. G. (1987). Comparison of *dnaA* nucleotide sequence of *Escherichia coli*, *Salmonella typhimurium* and *Serratia marcescens*. *J. Bacteriol.* **169**, 3976-3981.

Skurray, R. A., Nagaishi, H. & Clark, A. J. (1976). Molecular cloning of DNA from F sex factor of *Escherichia coli* K12. *Proc. Natl. Acad. Sci. USA*, **73**, 64-68.

Smith, C. A. & Thomas, C. M. (1984). Nucleotide sequence of the *trfA* gene of broad-host-range plasmid RK2. *J. Mol. Biol.* **175**, 251-262.

Smith, C. A., Shingler, V. & Thomas, C. A. (1984). The *trfA* and *trfB* promoter regions of broad host range plasmid RK2 share common potential regulatory sequences. *Nucleic Acids Res.* **12**, 3619-3630.

Smith, C. A. & Thomas, C. M. (1985). Comparison of the nucleotide sequences of the vegetative replication origin of broad host range IncP plasmids R751 and RK2 reveals conserved features of probable functional importance. *Nucleic Acid Res.* **13**, 557-572.

Smith, C. A. & Thomas, C. M. (1987). Comparison of the organisation of the genomes of phenotypically diverse plasmids of incompatibility group P: members of the IncPß subgroup are closely related. *Mol. Gen. Genet.* **206**, 419-427.

Smith, C. A. & Thomas, C. M. (1989). Relationships and evolution of IncP plasmids. In *Promiscuous plasmids of Gram-negative bacteria*. (Thomas, C. M., ed.), pp. 57-77, Academic Press. London, San Diego C. A.

Sϕgaard-Andersen, L., Rokeach, L. A. & Molin, S. (1984). Regulated expression of a gene important for replication of plasmid F in *Escherichia coli*. *EMBO J.* **3**, 257-262.

Sompayrac, L. & Maalϕe, O. (1973). Autorepressor model for control of DNA replication. *Nature (London) New Biol.* **241**, 133-135.

Stalker, D. M., Kolter, R. & Helinski, D. R. (1979). Nucleotide sequence of an origin of replication of the antibiotic resistance plasmid R6K. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1150-1154.

Stalker, D. M., Thomas, C. M. & Helinski, R. D. (1981). Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. *Mol. Gen. Genet.* **181**, 8-12.

Stenzel, T. T., Patel, P. & Bastia, D. (1987). The integration host factor of *Escherichia coli* binds to bent DNA at the origin of replication of the plasmid pSC101. *Cell*, **49**, 1401-1406.

Sternberg, N. (1979) A characterisation of bacteriophage P1 DNA fragments cloned in a lambda vector. *Virology*, **96**, 129-142.

Sternberg, A. & Austin, S. (1981). The maintenance of the P1 plasmid prophage. *Plasmid* , **5**, 20-31.

Sternberg, N., Enquist, L. & Weisberg, R. A. (1979). A simple technique for the isolation of deletion mutants of phage lambda. *Gene*, **8**, 35-51.

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

Sueoka, N. (1988). Directional mutation pressure and neutral molecular evolution. *Proc. Natl. Acad. Sci. USA*, **48**, 2653-2657.

Summers, D. K. & Sherratt, D. J. (1984). Multimerarisation of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerisation and stability. *Cell*, **36**, 1097-1103.

Swack, J. A., Pal, S. K., Mason, R. J., Abeles, A. L. & Chattoraj, D. K. (1987). P1 plasmid replication: measurement of initiator protein concentration *in vivo*. *J. Bacteriol.* **169**, 3737-3742.

Tait, R. C., Kado, C. I. & Rodriguez, R. L. (1983). A comparison of the origin of replication of pSa with R6K. *Mol. Gen. Gent.* **192**, 32-38.

Tardif, G. & Grant, R. B. (1983). Transfer of plasmids from *Escherichia coli* to *Pseudomonas aeruginosa*: characterisation of a *Pseudomonas aeruginosa* mutant with enhanced recipient ability for enterobacterial plasmids. *Antimicro. Agents Chemother.* **24**, 201-208.

Thatte, V., Bradley, D. E. & Iyer, V. N. (1985). N conjugative transfer system of plasmid pCU1. *J. Bacteriol.* **163**, 1229-1236.

Theophilus, B. D. M., Cross, M. A., Smith, C. A. & Thomas, C. M. (1985) Regulation of the *trfA* and *trfB* promoters of broad host range plasmid RK2: identification of sequences essential for regulation by *trfB/korA/korD*. *Nucleic Acids Res.* **13**, 8129-8142.

Theophilus, B. D. M. & Thomas, C. M. (1987). Nucleotide sequence of the transcriptional repressor gene *korB* which plays a key role in the regulation of the copy number of broad host range plasmid RK2. *Nucleic Acids Res.* **15**, 7443-7450.

Thomas, C. M. (1981). Molecular genetics of broad host range plasmid RK2. *Plasmid,* **5**, 10-19.

Thomas, C. M. (1983). Instability of a high copy number mutant of a miniplasmid derived from broad host range IncP plasmid RK2. *Plasmid*, **10**, 184-195.

Thomas, C. M. (1986). Evidence for the involvement of the *incC* locus of broad host range plasmid RK2 in plasmid maintenance. *Plasmid* **16**, 15-29.

Thomas, C. M. (1987a). Incompatibility group P plasmids: Genetics, evolution and use in genetic manipulation. *Ann. Rev. Microbiol.* **41**, 77-101.

Thomas, C. M. (1987b). Plasmid replication. In *Plasmids, a practical approach.* (Hardy, K. G., ed.), pp. 7-36, IRL Press, Oxford, Washington DC.

Thomas, C. M. (1988). Recent studies on the control of plasmid replication. *Biochim. Biophys. Acta.* **949**, 253-263.

Thomas, C. M., Cross, M. A., Hussain, A. A. K. & Smith , C. A. (1984). Analysis of copy number control elements in the region of the vegetative replication origin of the broad host range plasmid RK2. *EMBO J.* **3**, 57-63.

Thomas, C. M. & Helinski, D. R. (1989). Vegetative replication and stable inheritance of IncP plasmids. In *Promiscuous plasmids of Gram-negative bacteria.*(Thomas, C. M. ed.), pp. 1-25, Academic Press. London, San Diego C.A.

Thomas, C. M. & Hussain, A. A. K. (1984). The *korB* gene of broad host range plasmid RK2 is a major copy number control element which may act together with *trfB* by limiting *trfA* expression. *EMBO J.* **3**, 1513-1519.

Thomas, C. M., Hussain, A. A. K. & Smith, C. A. (1982). Maintenance of broad host range plasmid RK2 replicons in *Pseudomonas aeruginosa. Nature*, **298**, 674-676.

Thomas, C. M., Meyer, R. & Helinski, D. R. (1980). Regions of broad-host-range plasmid RK2 which are essential for replication and maintenance. *J. Bacteriol.* **141**, 213-222.

Thomas, C. M. & Smith, C. A. (1986). The *trfB* region of broad host range plasmid RK2: the nucleotide sequence reveals *incC* and key regulatory gene *trfB/korA/KorD* as overlapping genes. *Nucleic Acids Res.* **14**, 4453-4468.

Thomas, C. M. & Smith, C. A. (1987). Incompatibility group P plasmids: Genetics, evolution and use in genetic manipulation. *Ann. Rev. Microbiol.* **41**, 77-101.

Thomas, C. M., Stalker, D., Guiney, D. & Helinski, D. R. (1979). Essential regions for the replication and conjugal transfer of the broad host range plasmid RK2. In *Plasmids of medical, environmental and commercial importance* (Timmis, K. N. & Pühler, A., eds), pp 375-385, Elsevier/North-Holland Biomedical Press, Amsterdam.

Thomas, C. M., Stalker, D. M. & Helinski, D. R. (1981). Replication and incompatibility properties of segments of the origin of replication of the broad-host range plasmid RK2. *Mol. Gen. Genet.* **181**, 1-7.

Thompson, R. & Broda, P. (1973). DNA polymerase III and the replication of F and ColVBtrp in *Escherichia coli* K-12. *Mol. Gen. Genet.* **127**, 255-258.

Tilly, K. & Yarmolinsky, M. (1989). Participation of *Escherichia coli* heat shock proteins DnaJ, DnaK and GrpE in P1 plasmid replication. *J. Bacteriol.* **171**, 6025-6029.

Timmis, K., Cabello, F. & Cohen, S. N. (1975). Cloning, isolation and characterisation of replication regions of complex plasmid genomes. *Proc. Natl. Acad. Sci. USA*, **72**, 2242-2246.

Timmis, K. N., Cabello, F. & Cohen, S. (1978). Cloning and characterisation of *Eco*RI and *Hin*dII restriction endonuclease generated fragments of antibiotic resistance plasmids R6-5 and R6. *Mol. Gen. Genet.* **162**, 121-137.

Tokino, T., Murotsu, T. & Matsubara, K. (1986). Purification and properties of the mini-F plasmid-encoded E protein needed for autonomous replication control of the plasmid. *Proc. Natl. Acad. Sci. U.S.A.* **83**: 4109-4113.

Tolun, A & Helinski, D. R. (1981). Direct repeats of the F plasmid *incC* region express F incompatibility. *Cell*, **24**, 687-694.

Tolun, A. & Helinski, D. R. (1982). Separation of the minimal replication region of the F plasmid into a replication origin segment and a trans-acting segment. *Mol. Gen. Genet.* **186**, 372-377.

Tomizawa, J. (1986). Control of ColE1 plasmid replication: binding of RNA I and RNA II and inhibition of primer formation. *Cell*, **47**, 89-97.

Tomizawa, J. (1990). Control of ColE1 plasmid replication. *J. Mol. Biol.* **212**, 683-694..

Tomizawa, J., Itoh, T., Selzer, G. & Som, T. (1981). Inhibition of ColE1 RNA primer formation by plasmid specific small RNA. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1421-1425.

Tomizawa, J., Ohmori, H. & Bird, R. E. (1977). Origin of replication of colicin E1 plasmid DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1865-1869.

Tomizawa, J., Sakakibara, Y. & Kakefuda, T. (1974). Replication of colicin E1 plasmid DNA in cell extracts: origin and direction of replication. *Proc. Natl. Acad. Sci. USA*, **71**, 2260-2264.

Trawick, J. D. & Kline, B. C. (1985). A two-stage molecular model for control of mini-F replication. *Plasmid*, **13**, 59-69.

Tsutsui, H., Fujiyama, A., Murotsu, T. & Matsubara, K. (1983). Role of nine repeating sequences of the mini-F genome for expression of F-specific incompatibility phenotype and copy number control. *J. Bacteriol.* **155**, 337-344.

Tsutsui, T. & Matsubara, K. (1981). Replication control and switch-off function as observed with a mini-F factor plasmid. *J. Bacteriol.* **147**, 509-516.

Van Brunt, J., Waggoner, B. T. & Pato, M. L. (1977). Reexamination of F plasmid replication in a dnaC mutant of *Escherichia coli*. *Mol. Gen. Genet.* **150**, 285-292.

Veltkamp, E. & Stuitje, A. R. (1981). Replication and structure of the bacteriocinogenic plasmids CloDF13 and ColE1. *Plasmid*, **5**, 76-99.

Villafane, R. Bechhofer, D. H., Narayanan, C. S. & Dubnnau, D. (1987). Control genes of plasmid pE194. *J. Bacteriol.* **169**, 4822-4829.

Villarroel, R., Hedges, R. W., Maenhaut, R., Leemans, J. Engler, G., van Montagu, M. & Schell, J. (1983). Heteroduplex analysis of P-plasmid evolution: the role of insertion and deletion of transposable elements. *Mol Gen. Genet.* **189**, 390-399.

Vocke, C. & Bastia, D. (1983). Primary structure of the essential replicon of the plasmid pSC101. *Proc. Natl. Acad. Sci. USA*, **80**, 6557-6561.

Vocke, C. & Bastia, D. (1985). The replication initiator protein of plasmid pSC101 is a transcriptional repressor of its own cistron. *Proc. Natl. Acad. Sci. USA*, **82**, 2252-2256.

Wang, Q. & Kaguni, J. M. (1987). Transcriptional repression of the *dnaA* gene of *Escherichia coli* by dnaA protein. *Mol. Gen. Genet.* **209**, 518-525.

Watson, L. A., Phua, S. H., Bergquist, P. L & Lane, H. E. D. (1982). An M_r 29000 protein is essential for mini-F maintenance in *E. coli*. *Gene*, **19**, 173-178.

Wechsler, J. & Kline, B. C. (1980). Mutation and identification of the F plasmid locus determining resistance to acridine orange curing. *Plasmid*, **4**, 276-280.

Wickner, S. H. (1990). Three *Escherichia coli* heat shock proteins are required for P1 plasmid DNA replication: Formation of an active complex between *E. coli* DnaJ protein and the P1 initiator protein. *Proc. Natl. Acad. Sci. USA.* **87**, 2690-2694.

Wickner, S. H. & Chattoraj, D. K. (1987). Replication of mini-P1 plasmid DNA *in vitro* requires two initiation proteins encoded by the *repA* gene of phage P1 and the *dnaA gene of Escherichia coli. Proc. Natl. Acad. Sci. USA*, **84**, 3668-3572.

Willets, N. & Skurray, R. (1987). Structure and function of the F factor and mechanism of conjugation. In *Escherichia coli and Salmonella typhimurium, cellular and molecular*

biology (Neidhardt, F. C., Ingraham, J. L., Magasanik, B., Low, B. K., Schaechter, M. & Umbarger, H. E., eds), pp 1110-1133, American Society for Microbiology, Washington.

Willies, D. K., Uhlin, B. E., Amini, K. S. & Clark, A. J. (1981). Physical mapping of the *srl* and *recA* region of *Escherichia coli*: analysis of the Tn*10* generated insertions and deletions. *Mol. Gen. Genet.* **183**, 497-504.

Womble, D. D. & Rownd, R. H. (1987). Regulation of mini-F plasmid DNA replication. A quantitative model for control of plasmid mini-F replication in the bacterial cell division cycle. *J. Mol. Biol.* **195**,99-113.

Womble, D. D. & Rownd, R. H. (1988). Genetic and physical map of plasmid NR1: Comparison with other IncFII antibiotic resistance plasmids. *Microbiol. Rev.* **52**, 433-451.

Womble, D. D., Dong, X., Wu, R. P., Luckow, V. A., Martinez, A. F. and Rownd, R. H. (1984). IncFII plasmid incompatibility product and its target are both RNA transcripts. *J. Bacteriol.* **160**, 28-35.

Yakobson, E. & Guiney, G. (1983). Homology in the transfer origins of broad host range IncP plasmids: definition of two subgroups of P plasmids. *Mol. Gen. Genet.* **192**, 436-438.

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

Yamaguchi, K. & Masamune, Y. (1985). Autogenous regulation of synthesis of the replication protein in plasmid pSC101. *Mol. Gen. Genet.* **200**, 362-367.

Yamaguchi, K. & Yamaguchi, M. (1984a). The replication origin of pSC101: the nucleotide sequence and replication functions of the *ori* region. *Gene*, **29**, 211-219.

Yamaguchi, K & Yamaguchi, M. (1984b). The replication origin of pSC101: replication properties of a segment capable of autonomous replication. *J. Gen. Appl. Microbiol.* **30**, 347-358.

Yoon, K. H. & Pack, M. Y. (1987). Construction of a shuttle vector between *Escherichia coli* and *Zymomomnas anaerobia*. *Biotechnol. Lett.* **9**, 163-168.

Young, C., Bechhofer, D. H. & Figurski, D. H. (1984). Gene regulation in plasmid RK2: positive control by *korA* in the expression of *korC*. *J. Bacteriol.* **157**, 247-252.

Young, C., Burlage, R. S. & Figurski, D. H. (1987). Control of *kilA* gene of the broad-host-range plasmid RK2: involvement of *korA, korB* and a new gene *korE*. *J. Bacteriol.* **169**, 1315-1320.

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

Yun, T. & Vapnek, D. (1977). Electron microscopic analysis of bacteriophage P1, phage P1Cm and phage P7. Determination of genome sizes, sequence homology and location of antibiotic resistance determinants. *Virology*, **77**, 376-385.

Zabeau, M. & Stanley, K. K. (1982). Enhanced expression of cro-ß-galactosidase fusion proteins under the control of the P_R promoter of bacteriophage lambda. *EMBO J.* **1**, 1217-1224.

Zeuthen, J. & Pato, M. L. (1971). Replication of the F'lac sex factor in the cell cycle of *Escherichia coli*. *Mol. Gen. Genet.* **111**, 242-255.

Zünd, P. & Lebek, G. (1980). Generation time-prolonging R plasmids: correlation between increases in the generation time of *Escherichia coli* caused by P-plasmids and their molecular size. *Plasmid*, **3**, 65-69.

APPENDIX

APPENDIX 1

A. Bacterial strains[a]

Bacterial strain	Genotype	Reference
DK-1	*rec*A derivative of MC1061, *ara*DL39, del(*ara, leu*), 7697, del *lac*X74, *gal*U$^-$ *gal*K$^-$, *hsr*$^-$, *hsm*$^+$, del (*srl-rec*A)306, *str*A, Sm^R	Willies *et al.*, 1981
AB1157	*rec*A, *thr*-1, *leu*-6, *thi*-1, *lacY1*, *galK2*, *ara-14*, *xyl-5*, *mtl1-proA*2, *his5*, *arg*E3, *str*-31	Bachman, 1987
GW125a	*pol*A1 derivative of AB1157	Dorrington & Rawlings, 1989
LKIII	derivative of K54 *lacI*$^-$, *lacZ* M15, *lac*Y$^+$	Zabeau & Stanley, 1982
J53	K12 derivative, F$^-$, *pro*, *met*, Na^r	Meynell & Datta, 1966
APP (*Klebsiella pneumoniae*)	wild type, clinical isolate, Ak^r, Ap^r, Cm^r, Gm^s, Km^r, Nm^r, Nt^r, Sxt^r, Tc^r, Tb^r, W^r	This work
J53-2(RM98)	Ap^r, Sm^r, Sp^r, Tra$^+$, Ike^s, IncN	Konarska-Kozlowska & Iyer, 1981a
J53(RP4)	Ap^r, Kan^r, Tc^r IncP	Datta *et al.*, 1971

[a]All strains are *E. coli*, except were indicate.

B. Plasmids

Plasmid	Description	Antibiotic Marker	Reference
pUC19	cloning vector	Ap	Yanisch-Perron *et al.*, 1985
pCU1	39 Kb, Tra^+, Ike^s, isolated from J53-2(RM98) IncN	Ap, Sm, Sp	Konarska--Kozlowska & Iyer, 1981a Kozlowski *et al.*, 1987
pNZ116	9.1 Kb *Eco*RI fragment of F (mini-F, f5) in pBR322	Ap	Lovett & Helinski, 1976 Bergquist *et al.*, 1986

C. Bacteriophages

Plasmid specific phages	Description
Ike	RM98 IncN ssDNA filamentous chloroform sensitive *Eco* K12 J53 N3 50124 Khatoon & Iyer, 1971 Coetzee *et al.*, 1988
PR772	R772 IncN, IncP, IncW ds DNA icosahedral 99% chloroform sensitive *Eco* K12 J53 N3 50124 Coetzee *et al.*, 1979 Coetzee *et al.*, 1988
X	Inc M N P-1 U W X ssDNA filamentous chloroform sensitive STM LT2 M827 R6K 50497 Bradley *et al.*, 1981 Coetzee *et al.*, 1988

APPENDIX 2

A. Plasmids isolated in this study

Plasmid	Description	Antibiotic Marker	Reference
pGSH500	107 Kb, Tra^+ from a clinical isolate of *K. pneumoniae*	Ak, Ap, Cm, Km, Nm, Nt, Pn, Tb, Tc, Sxt, W	This study
pGSH510	76 Kb, Tra^- spontaneous deletion of pGSH500	Ak, Ap, Cm, Km, Nm, Pn, Tb, Sxt, W	This study
pGSH520	33 Kb, Tra^- spontaneous deletion of pGSH500	Ak, Ap, Cm, Km, Nm, Pn, Tb, Sxt, W	This study
pGSH530	22 Kb, Tra^- spontaneous deletion of pGSH500	Km, Nm, Sxt W	This study
pGSH540	10 Kb , Tra^- spontaneous deletion of pGSH500	Km	This study

B. Plasmid recombinants constructed in this study

Plasmid	Description	Antibiotic Marker
pFDT100	3.80 Kb *Hind*III fragment of pGSH500 in pUC19	Ap
pFDT200	3.30 Kb *Pst*I fragment of pGSH500 in pUC19	Ap
pFDT210	2.90 Kb exonuclease III deletion of pFDT200 (nucleotides 1-2889)	Ap
pFDT211	0.20 Kb *Hpa*I-*Pst*I fragment of pFDT210 in pUC19 (nucleotides 2684-2889)	Ap
pFDT212	0.66 Kb *Bst*XI-*Pst*I fragment of pFDT210 in pUC19 (nucleotides 2230-2889)	Ap
pFDT213	1.00 Kb *Hpa*I fragment of pFDT210 in pUC19 (nucleotides 1688-2684)	Ap
pFDT214	1.00 Kb *Hinc*II fragment of pFDT210 in pUC19 (nucleotides 660-1688)	Ap
pFDT215	0.66 Kb *Eco*RI-*Hinc*II fragment of pFDT210 in pUC19 (nucleotides 1-660)	Ap

Plasmid	Description	Antibiotic Marker
pFDT216	1.80 Kb *Eco*RI-*Bst*XI fragment of pFDT210 in pUC19 (nucleotides 1-1810)	Ap
pFDT220	2.40 Kb *Eco*RI-*Ava*II fragment of pFDT210 in pUC19 (nucleotides 1-2362)	Ap
pFDT2201	2.20 Kb exonuclease III deletion of pFDT220 (nucleotides 186-2362)	Ap
pFDT2202	1.90 Kb exonuclease III deletion of pFDT220 (nucleotides 478-2362)	Ap
pFDT2203	1.70 Kb exonuclease III deletion of pFDT220 (nucleotides 644-2362)	Ap
pFDT2204	1.50 Kb exonuclease III deletion of pFDT220 (nucleotides 840-2362)	Ap
pFDT2205	1.36 Kb exonuclease III deletion of pFDT220 (nucleotides 1003-2362)	Ap
pFDT2206	1.10 Kb exonuclease III deletion of pFDT220 (nucleotides 1262-2362	Ap

Plasmid	Description	Antibiotic Marker
pFDT2207	0.93 Kb exonuclease III deletion of pFDT220 (nucleotides 1438-2362)	Ap
pFDT2208	0.67 Kb exonuclease III deletion of pFDT220 (nucleotides 1693-2362)	Ap
pFDT220201	1.80 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-2279)	Ap
pFDT220202	1.62 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-2100)	Ap
pFDT220203	1.55 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-2031)	Ap
pFDT220204	1.30 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-1773)	Ap
pFDT220205	1.10 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-1574)	Ap
pFDT220206	0.85 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-1328)	Ap

Plasmid	Description	Antibiotic Marker
pFDT220207	0.58 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-1054)	Ap
pFDT220208	0.36 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-840)	Ap
pFDT220209	0.25 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-723)	Ap
pFDT220210	0.23 Kb exonuclease III deletion of pFDT2202 (nucleotides 478-710)	Ap
pFDT200Gm^r	3.30 Kb *Pst*I fragment of pFT200 (mini--replicon)	Gm (2.10 Kb *Eco*RI-*Hin*dIII fragment with AAC(3))
pFDT220204Gm^r	1.30 Kb fragment of pFDT220204 (mini--replicon)	Gm (2.10 Kb *Eco*RI-*Hin*dIII fragment with AAC(3))

Antibiotic marker abbreviations: Ak, amikacin; Ap, ampicillin; Cm, chloramphenicol; Gm, gentamicin; Km, kanamycin; Nm, neomycin; Nt, netilmicin; Pn, penicillin; Sxt, cotrimoxazole; Tb, tobramycin; Te, tetracycline; W, trimethoprim.

APPENDIX 3

Molecular weights of fragments generated after digesting pGSH500 with various restriction enzymes

Restriction enzyme	M_rs of fragments (Kb)	Total M_r (Kb)
*Hind*III	24.9; 14.2; 13.2; 9.0; 8.5; 5.7; 5.4; 4.7; 3.8 (x2); 3.1; 2.8; 2.1; 1.9; 1.7; 1.5; 1.1	107.4
*Eco*RI	22.1; 16.7; 13.7 (x2); 13.3; 9.5; 5.2; 3.5; 2.9; 2.6; 2.4	105.6
*Sma*I	24.6 (x2); 14.2; 8.5; 7.2; 6.2; 5.9; 4.7; 2.7; 2.6; 2.3; 1.6; 1.2; 1.0	107.3
*Sal*I	16.7 (x2); 14.2; 12.3; 10.1; 8.5; 8.1; 5.2; 3.6; 2.8; 2.2; 1.9; 1.6; 1.3	105.2

Restriction enzyme	**M_rs of fragments (Kb)**	**Total M_r (Kb)**
*Nco*I	22.1; 20.0; 16.7; 14.2; 10.8; 7.6; 5.7; 4.7; 4.1; 1.3; 1.2	108.4
Average molecular weight		106.8 $\pm$ 2

APPENDIX 4

MEDIA

Luria broth/agar

broth:	Yeast extract	5 gm
	Tryptone	10 gm
	NaCl	5 gm

(NB. quantities for one litre)

agar:	Yeast extract	5 gm
	Tryptone	10 gm
	NaCl	5 gm
	Agar	12 gm

(NB. quantities for one litre)

MacConkey agar:

	MacConkey agar base	40 gm
	Davis agar	8 gm

(NB. quantities for one litre)

BUFFERS AND SOLUTIONS

Agarose gel running buffer:

0.4M Tris-acetate

0.001M EDTA

pH 8.0

Sequencing gel running buffer:

0.133M Tris

0.04M Boric acid

0.0025M EDTA

1x SSC:

0.15M NaCl

0.15M Sodium citrate

pH 7.0

N.B. Other solutions were made as described in Sambrook *et al.* (1989).

APPENDIX 5

Modified QIAGEN plasmid Mini-preparation

QIAGEN kits, obtained from DIAGEN GmbH (Düsseldorf, Germany) were used to prepare plasmid DNA, by a modified QIAGEN Mini-preparation method. This modified method is described here. The buffers used are abbreviated as in the kit and their composition is detailed under the section "Composition of buffers" (see below).

As described in the instruction, *E. coli* cultures containing the desired plasmid were grown overnight at 37^{o}C. A 3 ml aliquot of the cultures was removed and was centrifuged (3000xg) at room temperature for 10 minutes. The bacterial pellet was suspended in 0.3 ml buffer P1 (1.5 ml eppendorf tube). This was followed by addition of 0.3 ml of buffer P2, gentle mixing by inversion, and incubation at room temperature for 5 minutes. A 0.3 ml aliquot of buffer P3 was added to the tube, mixed gently by inversion once again and centrifuged at 4^{o}C for 15 minutes (15000xg). The suppernatant was removed promptly after centrifugation, transferred to a fresh tube and centrifuged again under the same conditions. According to the QIAGEN protocol, this suppernatant should be loaded onto a "QIAGEN tip/column" to elute a pure plasmid DNA preparation. Since the yields from these columns were variable and often the DNA would not go into solution following precipitation of the column eluent, it was decided not to use the "QIAGEN tip/columns". To obtain pure plasmid DNA without the tip/columns, the modification described below was introduced.

Following the chromosomal precipitating step (after addition of P3/centrifugation), the suppernatant (approximately 0.9 ml) was divided into two aliquots (0.4 ml). To each aliquot, an equal volume of buffered phenol (Tris HCl, pH 8.0) was added and gently mixed. A further 0.4 ml aliquot of chloroform/isoamyl alcohol (24:1) was added. The two phases were gently mixed by inversion of tubes (no vortex). The tubes were then

centrifuged at 4^{o}C for 2 minutes (15000xg). The aqueous phase containing the DNA was removed and was re-extracted with phenol:chloroform as described before. The aqueous phase was re-extracted for a third time with chloroform/isoamyl alcohol. The plasmid DNA was precipitated from the aqueous phase with 0.6 volume isopropanol for 30 minutes at room-temperature followed by centrifugation at room-temperature for 30 minutes (15000xg). The DNA pellet was washed with 70 % ethanol, dried briefly (2 minutes in a Speed-vacum, Savant) and was dissolved in 30 μl ultra-pure water (Merck). The DNA was once again reprecipitated, with 1/10 volume 4M LiCL and 2.5 volumes absolute alcohol (- 20^{o}C). The precipitated DNA was kept at -20^{o}C until required.

Composition of buffers:

Buffer P1: RNase A (100 μg/ml) in 50 mM Tris HCl, 10 mM EDTA, pH 8.0, 4^{0}C.

Buffer P2: 20 mM NaOH, 1% SDS, room-temperature.

Buffer P3: 2.55 M KAc, pH 4.8, room-temperature.