REPLICATION INITIATION STUDIES OF A FAMILY OF SMALL STAPHYLOCOCCAL PLASMIDS.

by

Deborah Fiona Balson.

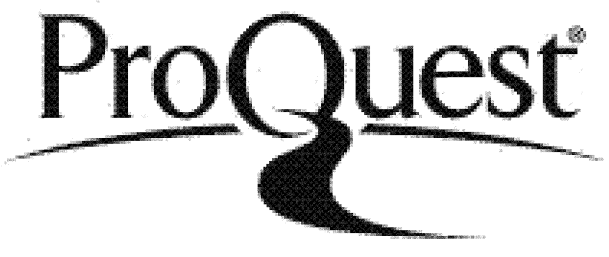
A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

OF THE UNIVERSITY OF LEICESTER.

DEPARTMENT OF BIOCHEMISTRY

UNIVERSITY OF LEICESTER

JANUARY 1989



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

CONTENTS.

ACKNOWLEDGEMENTS.

I should like to thank my supervisor, Professor Bill Shaw, for giving me the opportunity of working on this project and for his advice and support. I should also like to thank Chris Thomas, my co-worker on this project, for helpful suggestions and discussion, and, also, for providing me with purified REP D protein.

Special thanks go to the members of my committee, Eli Orr, whose help with the *in vitro* replication work was especially appreciated, and Tony Maxwell. I should also like to thank Iain Murray, Ann Lewenden, Philip Day, Jackie Ellis, Craig Livingstone, Dale Wigley and Jim Turner for their help and for making my time in the lab more enjoyable. Also thanks to Jim Turner and John Keyte for oligonucleotide synthesis.

Finally, I should like to thank my parents, my grandmother and last, but not least, my fiance, Steve Holmes, for their support and encouragement over the years, especially during the writing of this thesis.

ABBREVIATIONS.

A_x	absorbance at x nm.
bp	base pairs (of nucleic acid).
BSA	bovine serum albumin.
cat	gene encoding chloramphenicol acetyl transferase.
CAT	chloramphenicol acetyl transferase.
cAMP	cyclic AMP (adenosine 3' 5'-cyclic monophosphate).
CAP	catabolite gene-activator protein.
Ci	curies.
cop	plasmid copy control region.
cpm	counts per minute.
DNA	deoxyribonucleic acid.
DNase	deoxyribonuclease.
dNTP	deoxyribonucleotide 5' triphosphate.
DTT	dithiothreitol.
E. coli	*Escherichia coli.*
EDTA	ethylene diaminetetra-acetic acid.
Em	erythromycin.
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid.
ICR	inverted complementary repeat sequence.
inc	incompatability group.
IPTG	isopropyl-β-D-thiogalactoside.
kb	kilobases (of nucleic acid).
KDa	molecular weight in kiloDaltons.
mins	minutes.
mRNA	messenger ribonucleic acid.

NAD nicotinamide adenine dinucleotide.

OH hydroxyl

ORF open reading frame.

ori plasmid origin of replication.

PEG polyethylene glycol.

rep gene encoding REP protein.

REP plasmid replication initiator protein.

RNA ribonucleic acid.

RNase ribonuclease.

rNTP ribonucleotide 5' triphosphate.

S. aureus *Staphylococcus aureus.*

SDS sodium dodecylsulphate.

Sm streptomycin.

Tc tetracycline.

TCA trichloroacetic acid.

TEMED N, N, N', N' tetramethyl ethylenediamine.

tet gene encoding resistance to tetracycline.

TRIS tris (hydroxymethyl) amino methane.

UV ultraviolet.

X-GAL 5-bromo-4-chloro-indolyl-β-D-galactosidase.

ATP adenosine 5' triphosphate.

CTP cytidine 5' triphosphate.

GTP guanosine 5' triphosphate.

UTP uridine 5' triphosphate.

cAMP cyclic AMP [adenosine 3' 5'-cyclic monophosphate].

dATP 2'-deoxyadenosine 5'-triphosphate.

dCTP 2'-deoxycytidine 5'-triphosphate.

dGTP 2'-deoxyguanosine 5'-triphosphate.

dTTP thymidine 5'-triphosphate.

ddATP 2', 3'-dideoxyadenosine 5'-triphosphate.

ddCTP 2', 3'-dideoxycytidine 5'-triphosphate.

ddGTP 2', 3'-dideoxyguanosine 5'-triphosphate.

ddTTP 2', 3'-dideoxythymidine 5'-triphosphate.

$\alpha[^{35}S]$dATP deoxyadenosine 5'-[α-^{35}S] thiotriphosphate (sp isomer).

$\alpha[^{32}P]$dATP deoxyadenosine 5'-[α-^{32}P] triphosphate.

$\gamma[^{32}P]$ATP adenosine 5'-[γ-^{32}P] triphosphate.

[Me-^{3}H]dTTP [Methyl-^{3}H] thymidine 5'-triphosphate.

CHAPTER 1.

GENERAL INTRODUCTION.

1.1. - A family of small staphylococcal antibiotic resistance plasmids with a common replication mechanism.

1.1.1. - Plasmid pT181, a 4.4kb tetracycline resistance plasmid.

1.1.2. - Small staphylococcal plasmids related to pT181.

1.2. - The replication of pT181 and related plasmids.

1.2.1. - The pT181 origin of replication and the action of REP C *in vitro*.

1.2.2. - The replication mechanism of pT181 and related plasmids.

1.2.3. - Sequence comparisons of five rep genes.

1.2.4. - The control of replication of pT181 and related plasmids.

1.3. - The study of the pC221 replication mechanism.

1.3.1. - The replication of pC221 and REP D.

1.3.2. - The aims of this project.

CHAPTER 1.

GENERAL INTRODUCTION.

1.1 :- A family of small, staphylococcal antibiotic resistance plasmids with a common replication mechanism.

1.1.1 :- Plasmid pT181 - a 4.4kb tetracycline resistance plasmid.

Plasmid pT181 is a 4.4kb plasmid which confers tetracycline resistance. It has been sequenced (Khan & Novick, 1983) and was found to have three open reading frames (ORF) (Figure 1.1). ORF (a) is the inducible tetracycline resistance gene, tet (Iordanescu, 1979). ORF (b) specifies PRE (Gennaro *et al*, 1987), a 45kDa protein which appears to be required for site-specific recombination at recombination site RS_A, a 75bp site which is 5' to the pre coding sequence (Novick *et al*, 1984a). Recombination at this site can lead to the formation of cointegrates between pT181 and other plasmids containing this site, eg pE194.

ORF (c) encodes REP, a 38kDa protein which appears to be required for the replication of pT181. This is called REP C, as pT181 is in staphylococcal incompatibility group 3 (Inc3). Temperature-sensitive mutations affecting plasmid replication of pT181 have been mapped within this gene (Novick *et al*, 1982). Cell free extracts of *Staphylococcus aureus* will only support the replication of pT181 *in vitro* when the extracts are prepared from a strain containing pT181 (Khan *et al*, 1981). This suggested the requirement of a plasmid encoded product, presumably REP C.

REP C has been purified by cloning of the repC gene into a thermosensitive high copy number vector to enable overproduction of the protein in *Escherichia coli* (Koepsel *et al*, 1985a). Addition of purified REP C to extracts prepared from plasmid-free *S. aureus* allows the replication of pT181 in these extracts showing that REP C is required for pT181 replication *in vitro*.

The site of action of REP C, the pT181 origin of replication (oriC) has been found to lie within the repC coding sequence itself (Khan *et al*, 1982). REP C has been shown to have Mg^{++}-dependent, sequence-specific endonuclease and "topoisomerase-like" activities at oriC (Koepsel *et al*, 1985b). The pT181 copy control region (cop), which is responsible for the regulation of pT181 replication, has been found to lie immediately 5' to the repC coding sequence (Novick *et al*, 1984; Carleton *et al*, 1984). Another feature of the pT181 genome is a 150bp palindromic sequence called palA, which is also thought to be required for pT181 replication (Gruss *et al*, 1987).

1.1.2 :- Small staphylococcal plasmids related to pT181.

Plasmid pT181 is the prototype of a family of small staphylococcal plasmids which have a common replication mechanism. They all have a copy number of between 20 and 40 per chromosome equivalent and have very similar size and genomic organisation. The 4.6kb staphylococcal chloramphenicol resistance plasmid pC221 (Inc4) belongs to this family of plasmids. It has been sequenced (Brenner & Shaw, 1984) and was found to have four open reading frames (Figure 1.2). ORF (a) encodes chloramphenicol acetyltransferase (CAT), a 215 amino acid protein which confers resistance to chloramphenicol (Shaw, 1983; Shaw *et al*, 1985). ORF (b) and ORF (c) overlap and are thought to be involved in the formation of relaxation complexes, which may be required for the

mobilization of pC221 by conjugative plasmids (Archer & Projan, in preparation). ORF (d) was found to have approximately 80% homology with the pT181 repC gene, both at the level of the nucleotide sequence and of the amino acid sequence. This gene was therefore called repD.

Of other small, staphylococcal plasmids sequenced, pS194 (Inc5), pUB112 (Inc9) and pC223 (Inc10) (Projan & Novick, in press) have very similar genomic organisation to plasmids pT181 and pC221. All three have antibiotic resistance determinants (streptomycin in pS194 and chloramphenicol in pUB112 and pC223). They also have open reading frames responsible for either site-specific recombination (pUB112) or relaxation complex formation (pS194 and pC223) and open reading frames homologous to pT181 repC. The overall organisation of all five plasmid genomes is very similar. Very little of each of the plasmid genomes is non-functional and, in some cases, the open reading frames overlap. All of the open reading frames are co-directional.

Heteroduplex mapping between some small staphylococcal plasmids has been carried out (Skinner, 1984). Heteroduplex mapping between pC221 and the chloramphenicol resistance plasmid pCW7 (incompatibility group unknown) showed homology between them, not only over the region containing the cat gene, as expected, but also over the region containing the repD gene of pC221. This plasmid is of similar size and copy number to pT181 and related plasmids. Therefore it seemed likely that pCW7 belongs to this family of plasmids.

Experiments have been carried out in which the origin regions of pT181, pC221, pS194, pUB112 and pC223 were cloned into pRN5101, a temperature-sensitive derivative of the staphylococcal plasmid pE194 (Projan & Novick, in press). This plasmid is compatible with all members of the pT181 family (Iordanescu *et al*, 1978). Only the plasmid from which the cloned origin was derived would support replication of the cloned origin *in trans* at the non-

permissive temperature and there was no cross-reaction with any other plasmid origin in the family. Therefore, the cloned origins are recognised uniquely by a *trans*-acting product encoded by the parent plasmid, presumably the REP protein. A particular REP protein appears to be absolutely specific for its cognate origin *in vivo*.

Not all staphylococcal plasmids belong to the pT181 "family". Table (1.1) shows a list of known staphylococcal plasmid incompatibility groups with an example of each. Some are much larger than the pT181 family, have multiple resistance determinants and have different methods of replication.

1.2 :- The replication of pT181 and related plasmids.

1.2.1 :-The pT181 origin of replication and the action of REP C *in vitro*.

The origin of replication of pT181 (oriC) has been shown to lie within a 168bp region which forms part of the repC coding sequence itself (Figure 1.3) (Khan *et al*, 1982). Inverted complementary repeat sequences (ICR I, II and III) within the origin give rise to the possibility of secondary structure in the form of hairpins or cloverleaf structures.

REP C has been found to have a sequence-specific nicking /closing activity at oriC *in vitro* (Koepsel *et al*, 1985[b]). A *cis*-acting plasmid locus, called cmp has been identified which increases the interaction between the replication origin and the initiator protein (Gennaro & Novick, 1986). Mutations within this region appear to reduce the superhelical density of the plasmid DNA, suggesting that supercoiling of the DNA promotes the interaction between REP C and the origin.

The protein introduces a single-stranded nick at a fixed position within the loop region of the putative hairpin at ICR II, becoming covalently attached to the 5' phosphate group generated at the nick site. Much of the energy from the original phosphodiester bond is stored within this DNA-protein linkage and may, therefore, presumably be used in the religation reaction.

REP C therefore has a sequence-specific "topoisomerase-like" activity which can relax supercoiled, circular DNA containing the pT181 origin sequence. A topoisomerase is an enzyme which modifies the topological state of a DNA molecule by transient breakage of one strand (type I topoisomerase) or both

strands (type II topoisomerase), passage of one or both strands through the break and religation of the strands. (Reviewed by Gellert, 1981; Wang, 1985).

1.2.2 :- The replication mechanism of pT181 and related plasmids.

REP C has been shown to have a sequence-specific nicking /closing activity at the pT181 origin (Koepsel *et al*, 1985). Another well characterised initiation protein which acts in this manner is the gene A protein of bacteriophage ØX174 (Ikeda *et al*, 1976, Eisenberg & Kornberg, 1979). This protein is involved in the initiation of double-stranded phage replicative form (RF) DNA replication via a rolling circle mechanism which produces single-stranded (ss) phage DNA (Mok *et al*, 1987).

The gene A protein introduces a single-stranded nick at the origin of the double-stranded RF form of the DNA, a specific site within the gene A cistron itself. The protein becomes covalently attached to the 5' phosphate group generated at the nick site via a tyrosyl-dAMP phosphodiester bond (Eisenberg *et al*, 1980; Sanhueza *et al*, 1984) and complexes with the host helicase enzyme, which moves around the phage RF DNA, unwinding the double helix. This allows elongation at the free 3' hydroxyl group at the nick site by the host DNA polymerase. Cleavage and religation of the newly formed strand by gene A protein gives rise to single-stranded covalently closed circles which may be used to produce more RF DNA or packaged into phage particles (Figure 1.4). Complementary strand synthesis occurs via a separate mechanism, requiring the synthesis of RNA primers by the host primosome complex (Eisenberg *et al*, 1976).

By analogy with phage ØX174, it is thought that pT181 may also replicate via a rolling circle mechanism, with the free 3' hydroxyl group

generated by cleavage of the origin by REP C acting as a primer for the replication of one strand. Cleavage and religation of the newly synthesised strand by REP C after one round of replication leads to the formation of single-stranded circular pT181, the presence of which has been noted *in vivo* (Te Reile *et al*, 1986). A large palindromic sequence, referred to as palA, has been identified as the site of initiation of second strand replication (Gruss *et al*, 1987). Mutations within this sequence lead to a build up of single-stranded pT181 *in vivo*. By analogy with ØX174 replication, the initiation of second strand synthesis may require RNA primer synthesis.

Therefore, it would appear that the replication of pT181 is more likely to be closely related to the replication of circular DNA of some bacteriophages than to other well-studied mechanisms of plasmid replication, such as Col E1, which require the synthesis of an RNA primer for initiation of their replication (Figure 1.5). The proposed mechanism for pC221 and related plasmids is shown in figure (1.6) (Thomas, 1988).

1.2.3 :-Sequence comparison of five rep genes.

Comparison of the sequences of five rep genes (pT181 repC, pC221 repD, pS194 repE, pUB112 repI, and pC223 repJ) shows a very high degree of homology (approximately 75-85%) between them (Figure 1.7) (Projan & Novick, in press). However, the major differences between them appear to be clustered into two major regions, one at the 3' end of the coding sequence, corresponding to part of the origin region, and the other towards to 5' end of the coding sequence.

The major differences between the five origin sequences appear to occur within ICR III, with a few differences within ICR I. ICR II is almost

perfectly conserved in the five plasmids, with just one base change in pS194 and pUB112. All five plasmids have two Hpa II sites on either side of ICR II.

DNA footprinting experiments with REP C on the oriC sequence show that REP C binds mainly at ICR III, along with the 3' side of ICR II (Koepsel *et al*, 1986a). Therefore it would appear that the divergent region, ICR III, may be responsible for specificity determination at the origin.

The second region of divergence, towards the 5' end of the gene may, therefore, encode the region of the REP protein which is responsible for sequence-specific recognition at the origin. Evidence for this comes from work carried out on recombinants between pT181 and pC221 (Iordanescu, 1979) (Figure 1.8). *In vivo* recombinants in which the crossover had occurred within the conserved region in the middle of the rep gene were isolated and found to be defective for replication. They required the presence of the parental plasmid from which the origin was derived in order to provide a REP protein which would support replication *in vivo*. They are also able to complement replication-defective mutants of the other parental plasmid. This suggests that the recognition specificity of the hybrid protein is situated at the C-terminal end, whereas the specific recognition site for the protein corresponds to the divergent region within the origin. Presumably, the active site of the protein is encoded by the conserved regions of the gene.

1.2.4 :- The control of replication of pT181 and related plasmids.

REP is rate-limiting for the replication of these plasmids *in vivo* (Manch-Citron *et al*, 1986), and regulation of replication is mediated via the regulation of expression of the rep gene (Kumar *et al*, 1985). Control of the rep gene appears to be at the level of translation, via an RNA

countertranscript mechanism which is similar to some *E. coli* plasmid replication regulation mechanisms, such as plasmids ColE1 and R1 (reviewed by Scott, 1984).

The regions of pT181, pC221 and pS194 responsible for replication control have been localised to a 300bp region immediately 5' to the rep coding sequence (Carleton *et al*, 1984; Projan *et al*, 1985). For pT181, a Taq I-Rsa I restriction fragment which contains this region, cloned into another plasmid, displays an incompatibility towards pT181 known as Inc3A (Novick *et al*, 1984b). This is due to two small, overlapping RNA molecules which are encoded in this region (Kumar *et al*, 1985). These lie within the region which encodes the rep mRNA leader sequence and are transcribed in the opposite direction to the rep mRNA. These two small antisense RNAs (or RNA countertranscripts), called RNA I and RNA II, serve jointly as negative regulators of replication by controlling the translation of rep mRNA (Figure 1.9).

The control of replication of pT181 is mediated via the binding of antisense RNA to the rep mRNA leader sequence. This favours the formation of a hairpin structure at an inverted complementary repeat sequence, which sequesters the rep ribosome binding site. The interaction between the RNA countertranscripts and the rep mRNA leader sequence is probably not continuous (Kumar *et al*, 1985). Computer analysis of the repC mRNA leader sequence and antisense RNA I has shown the possibility of secondary structure with a large central loop (Figure 1.10). As many of the copy number mutants sequenced have base changes within this loop it seems likely that the initial interactions between RNA I and the repC mRNA involves base pairing between the corresponding loops of the complementary transcripts, in much the same way as has been found in ColE1 (Tomizawa *et al*, 1982; Lacatena *et al*, 1981) and the Inc FII plasmids (Nordsrom *et al*, 1984; Brady *et al*, 1983).

The copy control regions of pT181, pC221, pS194, pC223 and pUB112 have been compared (Projan & Novick, in press). Each has either one or two RNA countertranscripts and the possibility of forming the secondary structures required for regulation of copy number. However, slight changes in the sequence of this region lead to a change in the plasmid incompatibility group, as the RNA countertranscript from one of these plasmids can only act in a regulatory manner on the rep mRNA derived from the same plasmid or plasmids of the same incompatibility group. Changes within inverted repeat sequences tend to be such that there is a complementary change on both sides of the inverted repeat in order to maintain the same potential for RNA-RNA base pairing. As the rep mRNA leader sequence and the RNA countertranscripts are encoded in the same region but on opposite strands, any changes within the RNA countertranscript will be automatically reflected within the rep mRNA leader sequence.

There are two very important mutations within the copy control regions of pT181 and pC221 (Khan *et al*, 1982; Projan *et al*, 1985). These are called pT181 cop608 and pC221 cop903 respectively. These mutations involve deletions within the copy control region, over the sequence encoding the RNA countertranscripts, which results in increased expression of the rep gene. This leads to an enlarged copy number of between 800 and 1000 per chromosome equivalent. These mutants were used to prepare large amounts of pT181 and pC221 for *in vitro* work during this project.

As well as control of expression of the pT181 repC gene at the translational level by the countertranscript mechanism, there is also some evidence for control at the transcriptional level by an attenuation mechanism (Khan *et al*, 1983). The G-C rich stem loop structure which may form at the

ribosome binding site, followed by a stretch of U residues, resembles rho-dependent mRNA termination structures found in *E. coli* (Rosenberg *et al*, 1979).

1.3 :- The study of pC221 replication.

1.3.1 :-The replication of pC221 and REP D.

The sequencing of pC221 (Brenner & Shaw, 1984) led to the discovery of the pC221 repD gene, homologous with the pT181 repC gene. The repD gene has been cloned into a thermosensitive, dual-origin, inducible, high copy number vector, pHD (Figure 1.11) (Thomas, 1988). This allows over-expression of repD in *E. coli*. Purified REP D has a sequence-specific nicking /closing activity at the pC221 origin (oriD) with cleavage occurring at the same position as with REP C at the pT181 origin. REP D can, therefore, relax supercoiled DNA containing oriD by sequence-specific cleavage followed by religation to give the relaxed, covalently closed circular form. Energy from the transient DNA-protein linkage and probably also from the negative supercoiling of the DNA is used in this religation reaction. REP D appears to act stoichiometrically, i.e. it appears to become inactivated by some means after one nicking and religation event. At higher ratios of REP D to DNA the relaxed, covalently closed circular DNA is cleaved again by another molecule of REP D, but is not religated. This is presumably because there is no energy contribution from DNA supercoiling for the religation reaction. In the presence of Ba^{++} or Ca^{++}, instead of Mg^{++}, nicking and religation are separated, giving nicked, open circular plasmid DNA as the final product.

Use of synthetic oligonucleotides based on the oriD sequence has enabled the amino acid via which REP D becomes covalently attached to the DNA to be identified (Thomas, 1988). The oligonucleotide, labelled at the nick site with ^{32}P, was treated with REP D to give cleavage and covalent attachment. Complete acid hydrolysis of the labelled protein showed that a tyrosine residue was labelled. In other, similar topoisomerase-like enzymes, such as ØX174 gene

A protein (Eisenberg, 1980; Sanhueza *et al*, 1984), the FLP recombinase of the *Saccharomyces cerevisiae* 2μm plasmid (Gronostajski *et al*, 1985), the *Ustilago maydis* topoisomerase (Brougham *et al*, 1986) and the *Micrococcus luteus* topoisomerase and gyrase (Tse *et al*, 1980), the amino acid responsible for covalent attachment to the DNA was tyrosine, except for the resolvase enzyme of transposon γδ, where there is some evidence for an interaction between a serine residue and the resolvase site (Hatfull *et al*, 1986). Covalent attachment presumably occurs by the formation of a phosphodiester bond between the free 5' phosphate group of the DNA and the terminal oxygen of tyrosine or serine.

Proteolytic cleavage of REP D labelled with the oligonucleotide, followed by amino acid sequencing of the labelled polypeptide showed that the labelled amino acid residue was tyrosine 188, which is conserved in all known rep sequences. Site-directed mutagenesis of REP D in order to change this tyrosine residue to phenylalanine yields a mutant REP D (Y188F) which can bind to oriD, as shown by its retardation of DNA fragments containing oriD on agarose gels, but which can no longer cleave it.

1.3.2 :- The aims of this project.

The aims of this project were:

(i) to investigate the replication of pC221 *in vitro*, with particular emphasis on the specificity of REP D for pC221 and other plasmid origins.

(ii) to investigate the binding of REP D to the pC221 origin using DNA "footprinting" techniques, and to compare the results with published data on REP C and pT181.

(iii) in the light of the results of experiments carried out on pC221 replication *in vitro*, investigation of the importance of hairpin formation at ICR II of oriD in recognition and /or nicking at the pC221 origin.

(iv) to sequence the pCW7 repN gene to enable further comparison between it and other known rep genes of the pT181 family.

In summary, the original goal of the project was to gain a more precise view of the interactions of REP proteins with their respective ori targets, a necessary first step in understanding the biochemical apects of the replication initiation event.

Table 1.1: Properties of some S. aureus drug resistance plasmids.

The table shows the known incompatibility groups of *S. aureus* plasmids with an example of each group. Many of the smaller plasmids have a common replication mechanism involving initiation of replication by the rep gene product. The rep genes of these plasmids are named according to their incompatibility groups, e.g. pT181 is in incompatibility group 3 and its rep gene is known as repC.

GROUP	EXAMPLE	SIZE (kb)	RESISTANCE PHENOTYPE	REPLICATION	REFERENCE
1	pI258	28.7	Pc Asa Asi Sb Hg Om Cd Bi Pb Em	Unknown	Ruby & Novick (1975)
2	pII147	32.6	Pc Asa Asi Hg Om Cd Pb	Unknown	Ruby & Novick (1975)
3	pT181	4.4	Tc	REP C	Iordanescu *et al* (1978)
4	pC221	4.6	Cm	REP D	Ruby & Novick (1975)
5	pS194	4.6	Sm	REP E	Iordanescu *et al* (1978)
6	pK545	22.5	Km Nm	Unknown	Ruby & Novick (1975)
7	pUB101	21.2	Pc Hg Cd Fu	Unknown	Lyon & Skurray (1987)
8	pC194	2.8	Cm	Non-REP	Iordanescu *et al* (1978)
9	pUB112	4.6	Cm	REP I	Iordanescu *et al* (1978)
10	pC223	4.6	Cm	REP J	Iordanescu *et al* (1978)
11	pE194	3.5	Em	Non-REP	Iordanescu & Surdeanu (1979)
12	pT48	2.8	Em	Non-REP	Catchpole *et al* (1988)
13	pUB110	4.6	Nm Tm Km	Non-REP	Iordanescu *et al* (1980)
14?	pCW7	4.6	Cm	REP N	Skinner (1984)

Asa: Arsenate
Asi: Arsenite
Bi: Bismuth
Cd: Cadmium
Cm: Chloramphenicol
Em: Erythromycin
Fu: Fusidic acid
Hg: Mercury
Km: Kanamycin
Nm: Neomycin
Om: Organomercurials
Pb: Lead
Pc: Penicillin
Sb: Antimony
Sm: Streptomycin
Tc: Tetracycline
Tm: Tobramycin

Figure 1.1: Schematic diagram of the pT181 genome.

The loci are described in more detail in the text.

RS_A Site of interplasmid cointegrate formation.

pre Gene encoding a site-specific recombinase.

palA Palindromic sequence which is the origin of lagging strand replication in *S. aureus*.

cop Copy control region.

repC Gene encoding REP C, a *trans*-active replication initiator protein.

oriC pT181 origin of replication.

tet Tetracycline resistance gene.

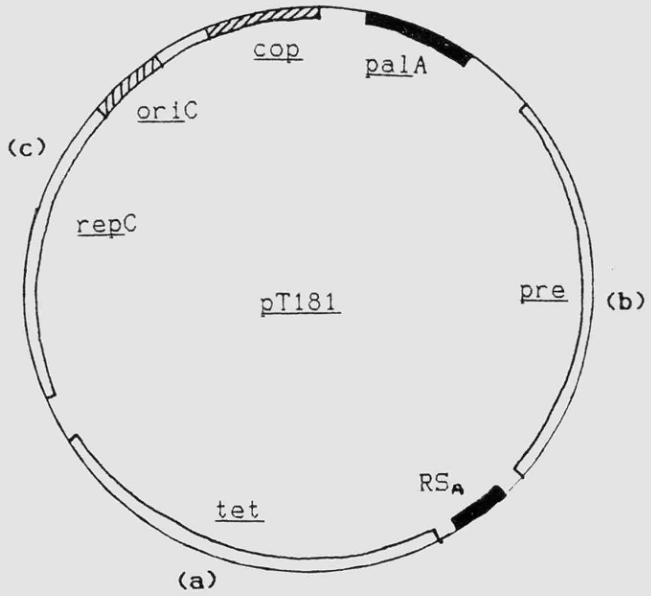

cop
palA
oriC
(c)
repC
pT181
pre
(b)
RS$_A$
tet
(a)

Figure 1.2: Schematic diagram of the pC221 genome.

The loci are described in more detail in the text.

mob Region containing two overlapping genes which encode proteins required for relaxation complex formation; may be involved in pC221 mobilisation by conjugative plasmids.

cop Copy control region.

repD Gene encoding REP D, a *trans*-active replication initiator protein.

oriD pC221 origin of replication.

cat Gene encoding CAT, an enzyme which confers resistance to chloramphenicol.

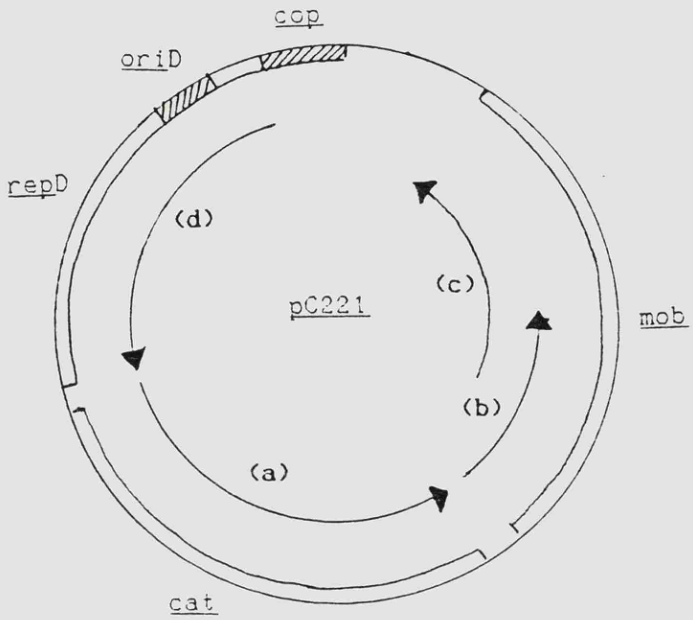
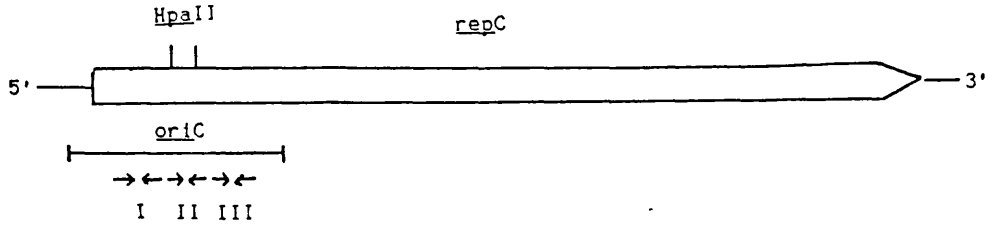
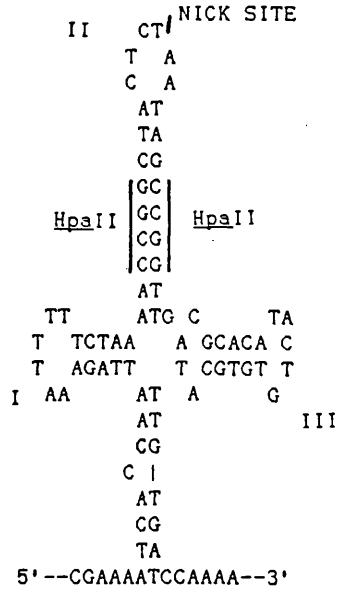

cop
oriD
repD
(d)
pC221
(c)
mob
(b)
(a)
cat

Figure 1.3: The rep and ori sequences of pT181.

The diagram shows the position of the pT181 origin region (oriC) within the repC gene. Inverted complementary repeat sequences (ICR I, II and III) give rise to the possibility of secondary structure within oriC, in the form of a cloverleaf, as shown.


HpaII
repC
5'
3'
oriC
I II III


NICK SITE
II
CT
T A
C A
AT
TA
CG
GC
GC
CG
CG
HpaII
HpaII
AT
TT
ATG C
TA
T TCTAA
A GCACA C
T AGATT
T CGTGT T
I AA
AT A
G
AT
III
CG
C
AT
CG
TA
5'--CGAAAATCCAAAA--3'

pT181 /oriC

Figure 1.4: Diagrammatic representation of proposed rolling circle mechanism for bacteriophage ØX174.

The diagram shows the proposed rolling circle mechanism for replication of bacteriophage ØX174 replicative form (RF) DNA. The phage-encoded gene A protein introduces a single-stranded nick into the (-) strand at a specific site within the A gene itself. The gene A protein becomes covalently attached to the 5' phosphate at the nick site and complexes with the host helicase enzyme. This complex moves around the DNA, unwinding the double helix. The host single-stranded binding protein (ssb) is also required. The free 3' hydroxyl group generated at the nick site is utilised as a primer by the *E. coli* DNA polymerase III enzyme. After one round of synthesis, the endonuclease and ligase activities of the gene A protein are required for termination and production of circular, single-stranded molecules which are required for packaging into phage particles. The gene A protein is transferred to the 5' end of the newly synthesised strand, allowing re-initiation of viral strand synthesis.

As well as single-stranded DNA, this mechanism can also give rise to double-stranded RF DNA by initiation of second strand synthesis. This requires the synthesis of an RNA primer by the host primosome complex.

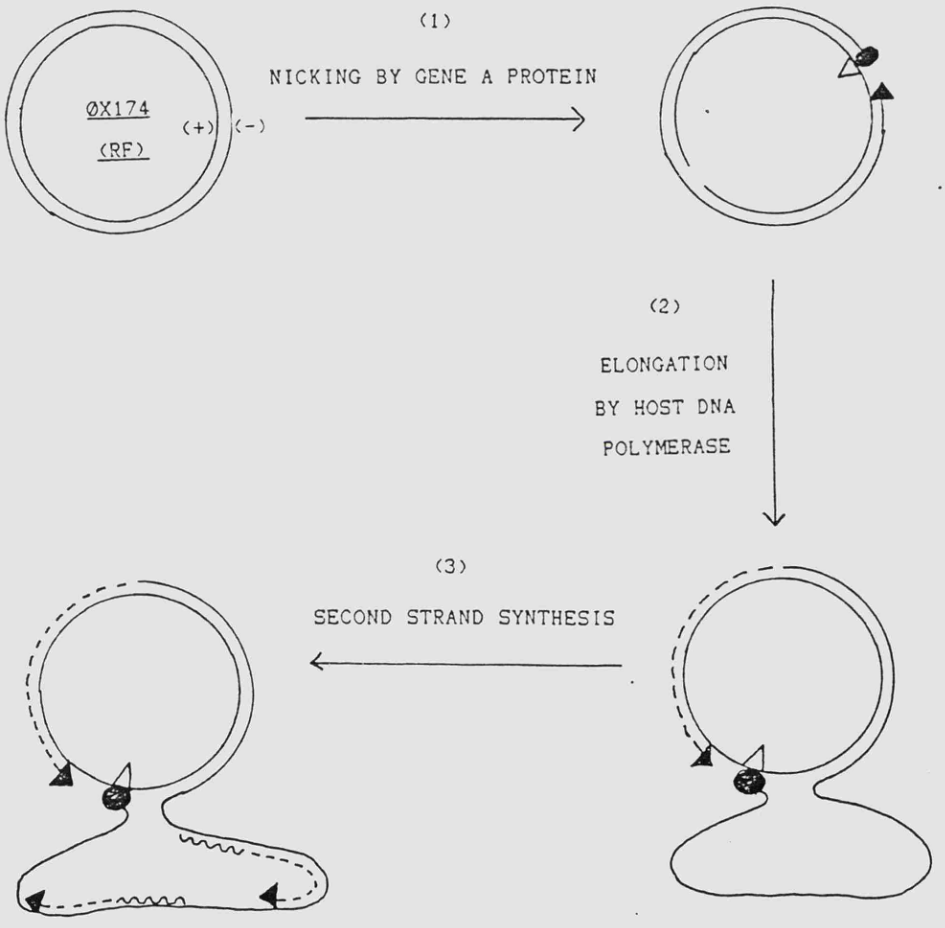

(1)
NICKING BY GENE A PROTEIN
ØX174
(RF)
(+)
(-)
(2)
ELONGATION
BY HOST DNA
POLYMERASE
(3)
SECOND STRAND SYNTHESIS
GENE A PROTEIN
HELICASE
DNA POLYMERASE
PRIMER SYNTHESIS

Figure 1.5: The proposed replication mechanism of *E. coli* plasmid Col E1.

A diagrammatic representation of the proposed replication mechanism of plasmid Col E1 is shown.

For Col E1 replication, Rifampicin-sensitive RNA synthesis is required in order to produce a primer. Primer transcription initiates at a point approximately 555bp upstream of the origin. The preprimer RNA is processed to give the mature primer by cutting at the origin by RNase H. Primer formation is inhibited by a regulatory RNA countertranscript (RNA 1) which prevents the hybridisation of the DNA and the primer RNA by hybridisation with the primer RNA in the form of a cloverleaf, as shown.

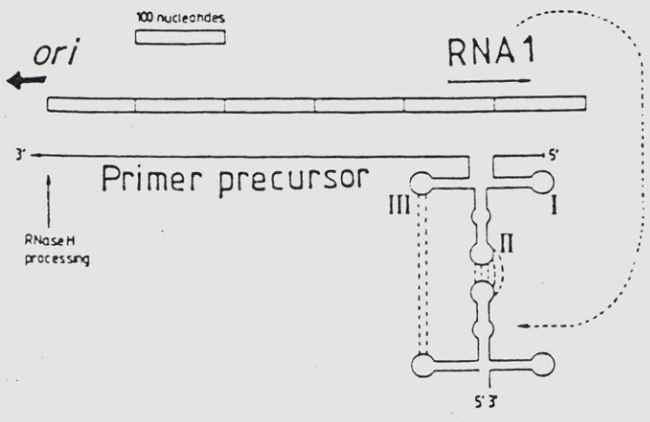

100 nucleotides
ori
RNA 1
3'
5'
Primer precursor
III
I
II
RNaseH
processing
5'3'

Figure 1.6: The proposed replication mechanism of *S. aureus* plasmid pC221.

A diagrammatic representation of the proposed replication mechanism of pC221 is shown.

Recogniton of its cognate origin sequence by the REP protein is followed by cleavage of the (+) strand and covalent linkage of REP to the 5' phosphate group generated. By analogy with bacteriophage ØX174 replication, host replication factors then extend the new (+) strand from the free 3' hydroxyl group generated. Single-stranded binding proteins may protect the displaced (-) strand. REP may associate with the host replisome so as to be present at the origin after a whole round of replication. Cleavage of the new strand from the old by REP, followed by religation of new and old strands would give rise to one daughter plasmid and a single-stranded (+) strand DNA circle. Second strand synthesis then proceeds from a large palindromic sequence, palA, probably involving RNA primer synthesis.

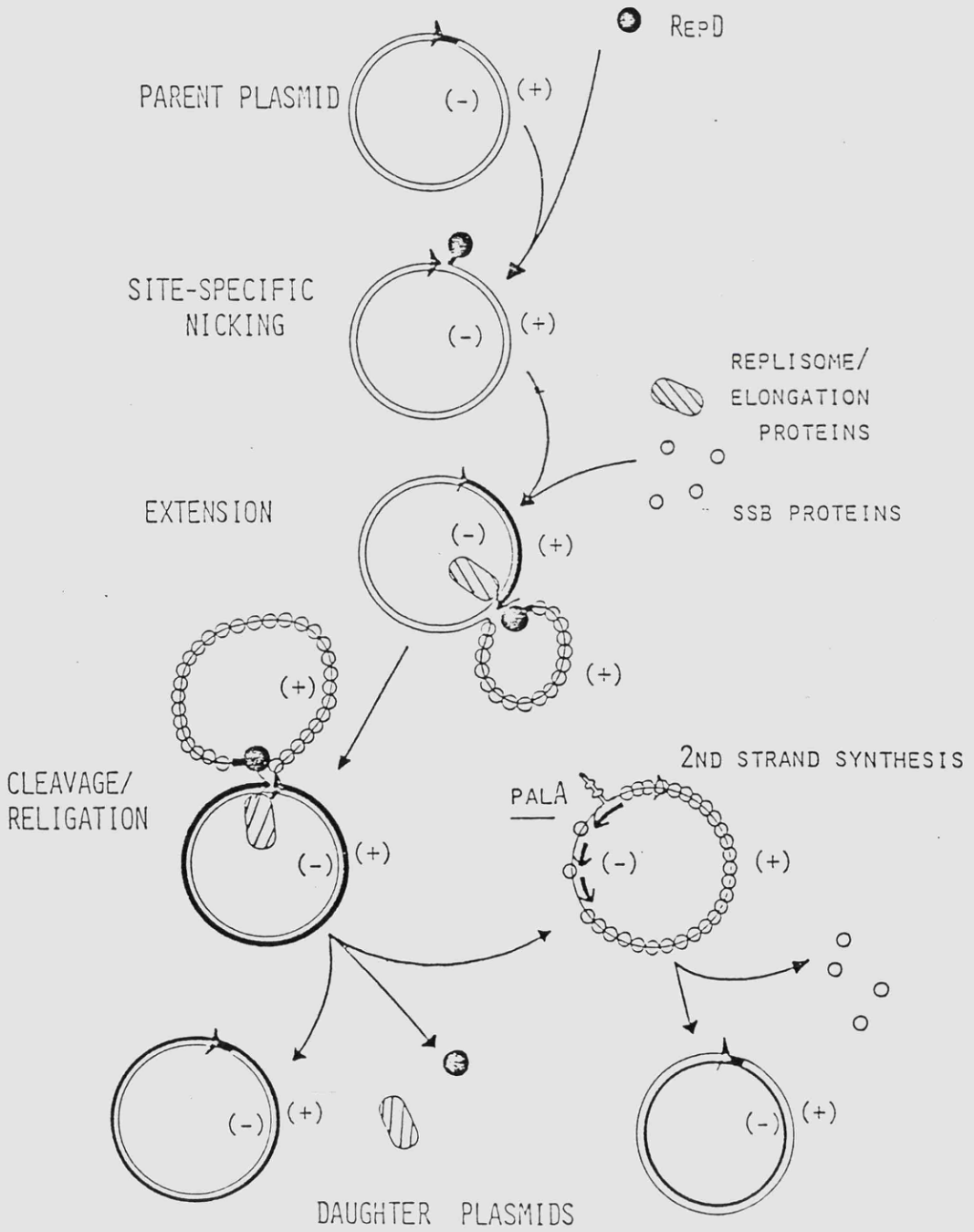

RepD
PARENT PLASMID
(-)
(+)
SITE-SPECIFIC
NICKING
REPLISOME/
ELONGATION
PROTEINS
SSB PROTEINS
EXTENSION
CLEAVAGE/
RELIGATION
2ND STRAND SYNTHESIS
PALA
DAUGHTER PLASMIDS

Figure 1.7: Comparison of the DNA sequences of five known rep genes.

The DNA sequences of the pT181 repC, pC221 repD, pS194 repE, pUB112 repI and pC223 repJ genes are presented (Projan *et al*, in press). Where three or more of the five bases are identical, capital letters are used.

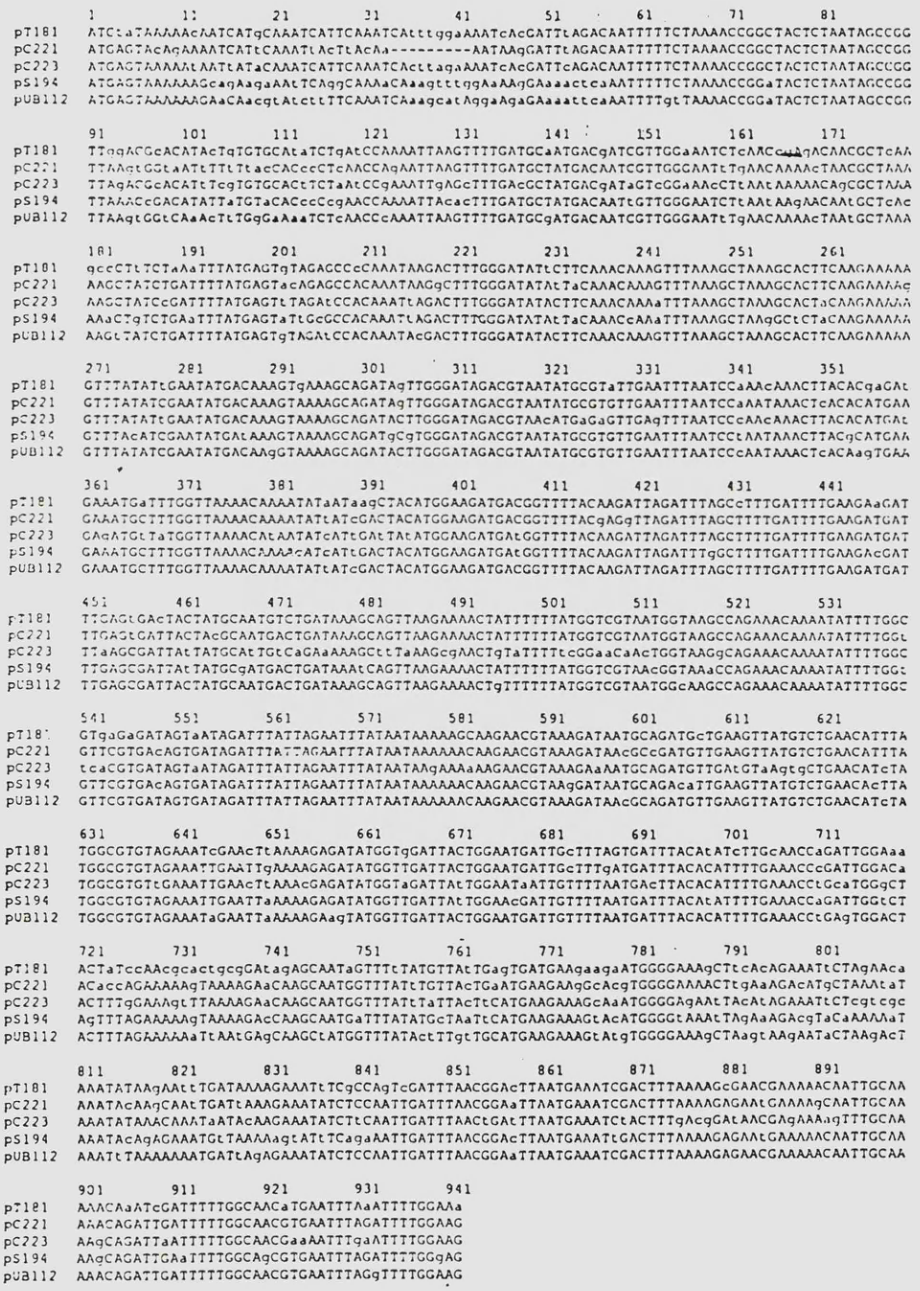


Figure 1.8: *In vivo* recombinants between pC221 and pT181.

In vivo recombinants between pC221 and pT181 were isolated which were defective in replication (Iordanescu, 1979). These were found to be hybrid plasmids in which the crossover event had occured within the conserved region in the middle of the rep genes. Therefore, these plasmids had the ori region and 5' end of the rep gene derived from one plasmid and the 3' end of the rep gene derived from the other plasmid. The recombinant plasmids were found to require the presence of the parental plasmid from which the ori sequence was derived in order to provide REP *in trans* for initiation of replication *in vivo*. They were also able to complement replication-defective mutants of the other parental plasmid *in trans*.

These results suggested that the domain of the protein responsible for sequence-specific recognition at the origin is situated towards the C-terminal end of the protein, whereas the specific recognition site for the protein corresponds to the divergent region within the origin. Presumably, the highly conserved regions of the gene encode the active site of the protein, responsible for the nicking and religation reactions at the origin.

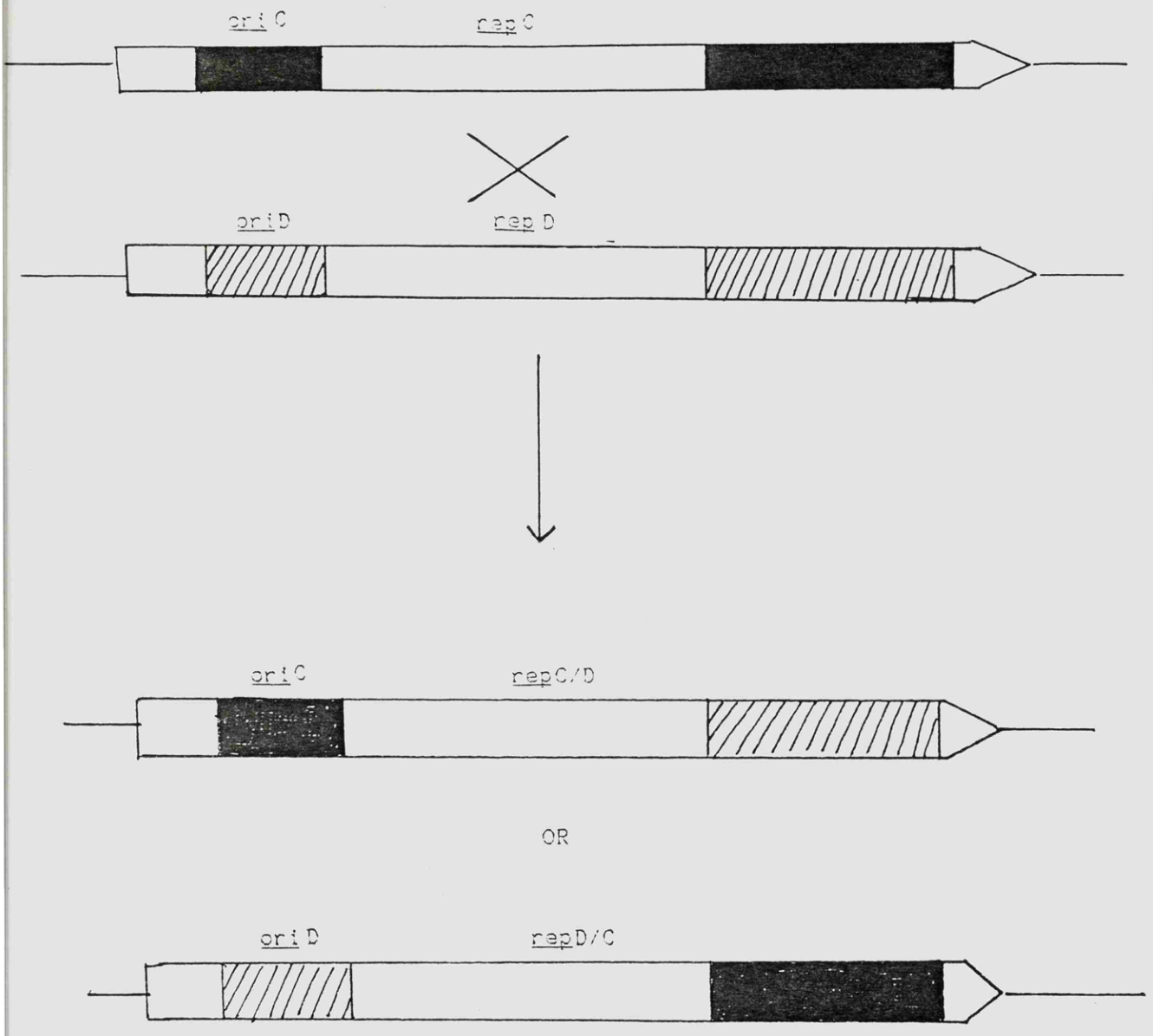
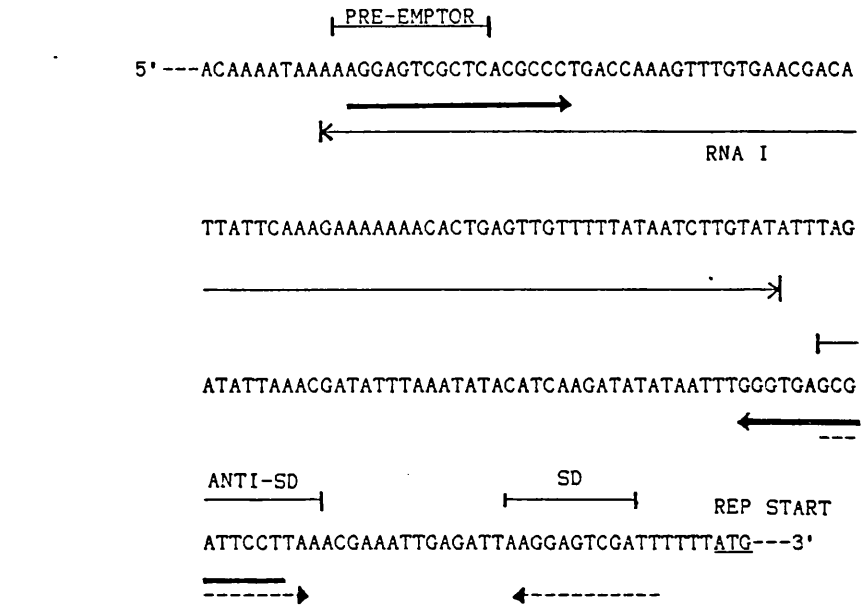

ori C
rep C
ori D
rep D
ori C
repC/D
OR
ori D
repD/C

Figure 1.9: The pT181 copy control region (cop).

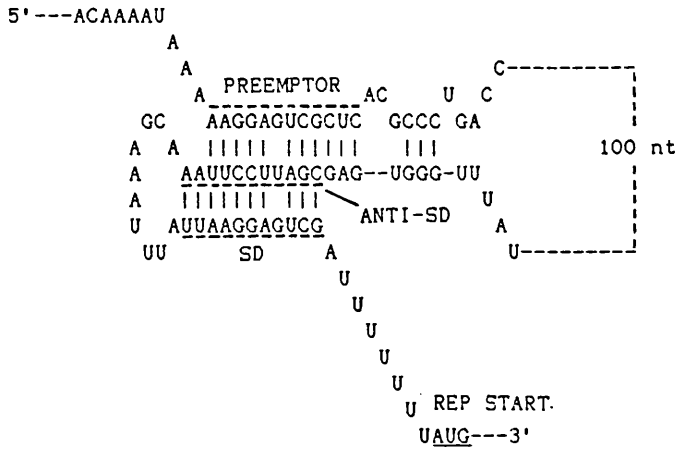
Diagram (a) shows the position of the repC start site, the ribosome binding site (SD), the anti-SD sequence, the preemptor sequence and the antisense RNA (RNA I). Diagram (b) shows the putative secondary structure within the repC mRNA leader sequence. Formation of a hairpin between SD and anti-SD sequences sequesters the ribosome binding site of the mRNA, preventing its translation. However, formation of a hairpin at the preemptor sequence, which overlaps with the anti-SD sequence, prevents formation of the hairpin at the ribosome binding site, allowing translation of the mRNA to occur. Binding of the antisense RNA to its complementary sequence on the rep mRNA leader sequence prevents formation of a hairpin at the preemptor sequence, allowing the hairpin at the ribosome binding site to form. This prevents translation of the rep mRNA. Therefore, the antisense RNA appears to have an inhibitory effect on rep expression at the level of translation.

The G-C rich hairpin which may form at the ribosome binding site resembles rho-dependant termination structures seen in *E. coli*. Therefore, this may also be involved in an attenuation regulation mechanism at the level of rep gene transcription.

(a)



(b)



(SD = SHINE-DALGARNO RIBOSOME BINDING SITE)

Figure 1.10: Predicted secondary structures of the pT181 repC mRNA leader sequence and antisense RNA I.

The diagram shows the computer predicted secondary structures of RNA I, the repC mRNA leader sequence and the complex between the two. The latter structure was obtained by omitting the first 90 bases of the rep mRNA from the computer analysis. The ribosome binding site (SD) is shown boxed. Formation of a complex between the two RNAs appears to favour the sequestering of the SD sequence within a double-stranded hairpin structure, preventing repC mRNA translation. The interaction between the repC mRNA leader sequence and the second antisense RNA (RNA II) has been shown to be the same.

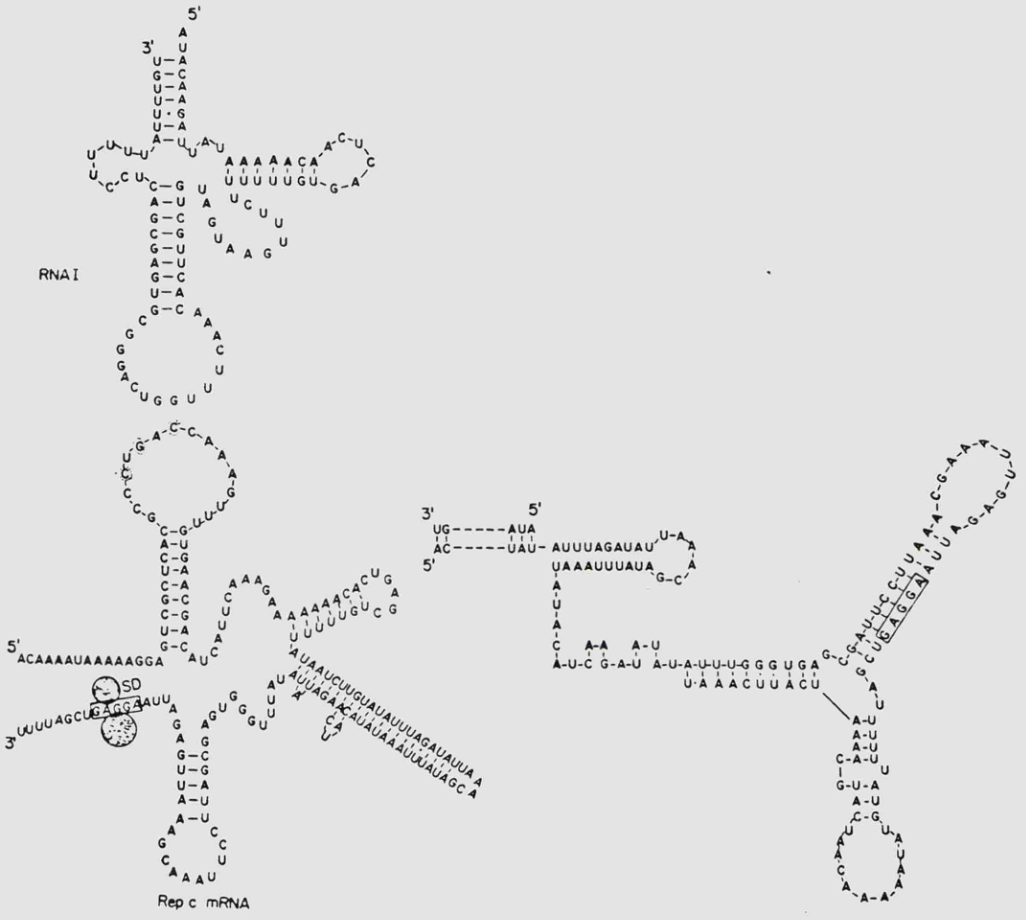
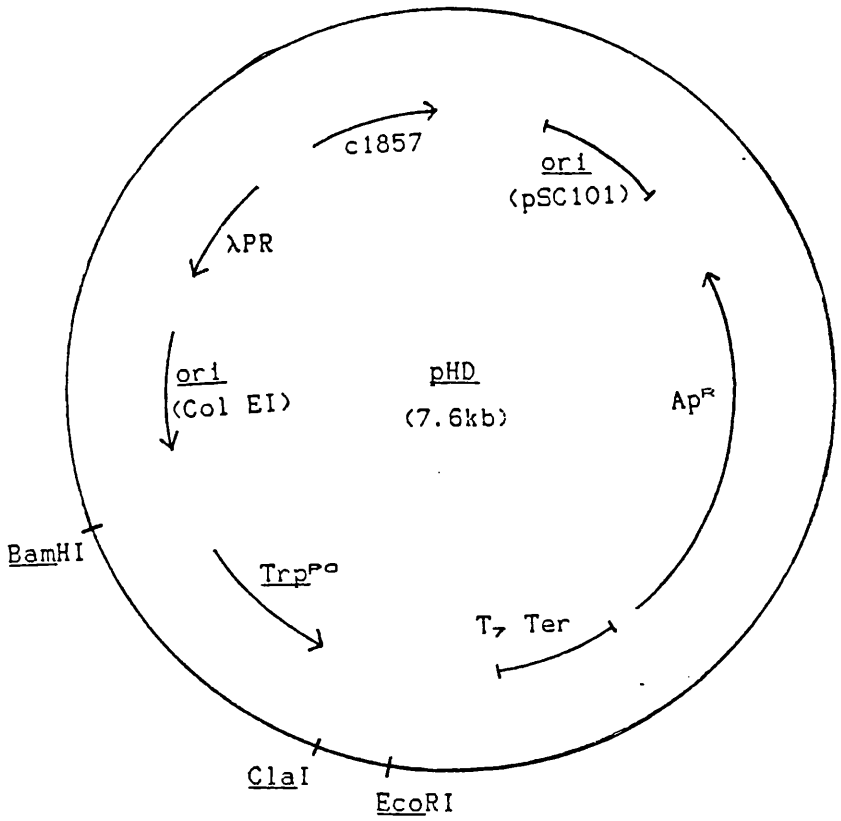

RNA I
Rep c mRNA
SD

Figure 1.11: Schematic diagram of pHD, a thermosensitive, high copy number vector (Thomas, 1988).

Plasmid pHD is a dual-origin, thermosensitive, high copy number vector for use in *E. coli*. Copy number is controlled by two origins, one with a low copy number from plasmid pSC101 and one temperature-sensitive origin with a high copy number consisting of the ColE1 origin under the control of the λ rightward promoter, λP_R, and the associated temperature-sensitive repressor, cI857. At higher temperatures, cI857 is inactive, λP_R becomes active and copy number increases as large amounts of the ColE1 primer RNA are produced. At high copy number, the tryptophan promoter, trpPO, escapes repression due to the host trp represor encoded by the E. coli chromosome, leading to enhanced transcription through inserts bounded by the unique Cla I or Eco RI sites. The ampicillin resistance gene, Ap^r, gives a selectable drug resistance marker.


c1857
ori
(pSC101)
λPR
ori
(Col EI)
pHD
(7.6kb)
Ap^R
BamHI
Trp^PO
T7 Ter
ClaI
EcoRI

CHAPTER 2.

MATERIALS AND METHODS.

2.1. - Materials.

2.1.1. - Chemicals and enzymes.

2.1.2. - Bacterial strains.

2.1.3. - Plasmids and vectors.

2.1.4. - Buffers and stock solutions.

2.1.5. - Microbial growth media.

2.1.6. - Oligonucleotides.

2.2. - Methods.

2.2.1. - Storage of bacterial strains.

2.2.2. - Preparation of plasmid DNA.

2.2.3. - DNA manipulations.

2.2.4. - DNA cloning and sequencing.

2.2.5. - *S. aureus in vitro* replication system.

2.2.6. - DNase I footprinting of the pC221 origin (oriD) DNA with REP D.

2.2.7. - *In vitro* mutagenesis of oriD by the Kunkel method.

2.2.8. - Studies of the REP D /oriD interaction using synthetic oligonucleotides.

CHAPTER 2.

MATERIALS AND METHODS.

2.1: - Materials.

2.1.1: - Chemicals and enzymes.

Most chemicals were purchased from Fisons Ltd. and, unless otherwise stated, were of analytical grade purity. Agarose was purchased from Miles laboratories. Tris (Trizma) base, DTT, lysozyme, lysostaphin, nucleic acids, nucleotides, ammonium persulphate, bromophenol blue, BSA, ethidium bromide, 8-hydroxy quinolone, IPTG, PEG 6000, RNase (pancreatic, type I-AS) and topoisomerase (type I, from calf thymus) were from Sigma. Deoxy and dideoxyribonucleotide triphosphates, Klenow fragment of DNA polymerase I (FPLC pure), DNase I and Sephadex G-25 and G-50 (DNA grade) were from Pharmacia. Bacto tryptone, Bacto yeast extract, Bacto agar, Bacto tryptose blood agar base and penassay broth were from Difco. $\alpha[^{35}S]$ dATP (triethyl ammonium salt, 400Ci/mMol, 10mCi/ml), $\alpha[^{32}P]$ dATP (triethyl ammonium salt, 3000Ci/mMol, 10mCi/ml), $\gamma[^{32}P]$ dATP (triethyl ammonium salt, 3000 Ci/mMol, 10mCi/ml) and [methyl-^{3}H] dTTP (triethyl ammonium salt, 40-60Ci/mMol, 1mCi/ml) were from Amersham. Restriction enzymes, Klenow fragment of DNA polymerase I, T_4 DNA ligase and T_4 polynucleotide kinase were from Bethesda Research Laboratories (BRL) /Gibco. Alkaline phosphatase was from Boehringer Mannheim. Pronase was from Calbiochem Behring. Amberlite MB1 resin, dichloromethyl silane solution (in

1.1.1.trichloroethane) and xylene cyanol (FF) were from BDH. X-gal was from Anglian biotechnology. 3mm filter paper and GF/C glass fibre discs (25mm) were from Whatman. Acrylamide and bis acrylamide were from Serva.

2.1.2: - Bacterial strains.

STRAIN	NOTES	REFERENCE OR SOURCE
E. coli JM101.	d lacpro, supE, thi1, F', tra D36, proAB, lacI^{q}, ZdM15. Routine host for bacteriophage M13 based recombinants.	Messing, 1983.
E. coli RZ1032.	dut, ung, deficient in thymidine biosynthesis. Used for misincorporation of deoxyuracil into DNA for *in vitro* mutagenesis.	Obtained from I. Murray, University of Leicester.
S. aureus WS2001. (RN450).	Cured of prophages.	Novick, 1976.
S. aureus RN1786.	RN450, nuclease free.	Koepsel et al, 1985a.

2.1.3: - Plasmids and vectors.

Plasmid or vector	Source	Size	Selection marker	Notes	Reference
Phage M13 (mp18&19)	*E. coli.*	7.2kb.	lacZ	Cloning of DNA fragments to allow preparation of single-stranded DNA for in vitro mutagenesis and DNA sequencing.	Messing; 1983.
pHD.	*E. coli.*	7.6kb.	Amp^R.	Temperature sensitive dual origin vector for over-expression of cloned genes.	Thomas, 1988.
pC221.	*S. aureus.*	4.6kb.	Cm^R.	Inc4. Replication via REP D. Copy number ≃20 per chromosome.	Ruby & Novick, 1975.
pC221 / cop903.	*S. aureus.*	4.6kb.	Cm^R.	Inc4. Replication via REP D. Deletion mutation within cop counter-transcript region, (Δ806-1192). Copy number ≃800 per chromosome.	Projan et al, 1985.

Plasmid or vector	Source	Size	Selection marker	Notes	Reference
pT181	*S. aureus*	4.4kb	Tet^R	Inc 3. Replication via REP C. Copy number ≈20 per chromosome.	Novick, 1976.
pT181 / cop 608.	*S. aureus.*	4.4kb.	Tet^R.	Inc3. Replication via REP C. Deletion within cop counter-transcript region, (Δ183-326). Copy number ≈800 per chromosome.	Khan et al, 1982.
pCW7	*S. aureus.*	4.4kb	Cm^R.	Inc14. Replication via REP N. Copy number ≈20 per chromosome.	Wilson & Baldwin, 1980. Wilson *et al*, 1981.
pC194.	*S. aureus.*	2.8kb.	Cm^R.	Inc8. No sequences homologous with pT181.	Iordanescu, 1975.

2.1.4: - Buffers and stock solutions.

Tris /EDTA buffer (TE) X100 :- 1M Tris.HCl (pH8),
0.1M EDTA.
Sterilised by autoclaving (15 mins, 15psi).

Solutions required for preparation of plasmid DNA :-

Lysostaphin stock solution :- Lysostaphin was dissolved in TE to 10 mg/ml and stored in small aliquots at -20°c.

Lysozyme stock solution :- Lysozyme was dissolved in TE to 50mg/ml and stored in small aliquots at -20°c.

Lysis buffer :- 50mM Sucrose,
10mM EDTA,
25mM Tris.HCl (pH8).

Alkaline SDS :- 0.2M NaOH,
1% SDS.

Potassium acetate (pH 4.8) :- 3M Potassium acetate,
11.5% glacial acetic acid.

Ethidium bromide stock solution :- Ethidium bromide was dissolved in water to 5mg/ml and stored in the dark at 4°c.

Isopropanol (saturated with CsCl) :- Equal volumes of isopropanol and saturated CsCl solution were equilibrated and the layers allowed to separate.

Preboiled RNase :- Pancreatic RNase was dissolved to 10mg/ml in 10mM Tris.HCl (pH7.5) and heated to 100°c for 15 mins. It was allowed to cool slowly, then stored in small aliquots at -20°c.

Self-digested pronase :- Pronase was dissolved to 2mg/ml in 0.01M Tris.HCl (pH7.8), incubated at 37°c for 2 hours and stored in small aliquots at -20°c.

Solutions required for routine DNA manipulations :-

Sterile water :- All water used for DNA manipulations was deionised, distilled, filtered through millipore filtration apparatus and sterilised by autoclaving (15 mins, 15psi).

Tris/Borate/EDTA (TBE) electrophoresis buffer X10 :- 109g Tris base,
55g Boric acid,
9.3g EDTA.
The solution was made to 1l with distilled water.

Tris/acetate/EDTA (TAE) electrophoresis buffer X10 :- 242g Tris base,
57.1g Glacial acetic acid,
100ml 0.5M EDTA.
The solution was made to 1l with distilled water.

Agarose gel loading buffer :- 0.25% Bromophenol blue,
0.25% Xylene cyanol,
15% Ficoll (type 400).

Phenol/Chloroform/Isoamyl alcohol stock solution :- Freshly distilled phenol (with 0.1% 8-Hydroxy quinoline) was equilibrated with an equal volume 1M Tris.HCl (pH7.5) 2-3 times, until the pH was approximately neutral. The upper aqueous layer was dicarded and an equal volume of 24:1 Chloroform /Isoamyl alcohol was added. The mixture was again equilibrated with an equal volume of 1M Tris.HCl (pH7.5) 2-3 times. Most of the upper aqueous layer was discarded leaving a thin layer to protect the phenol against oxidation. The mixture was stored in the dark at 4°c.

3M Sodium acetate :- Sodium acetate (408.1g) was dissolved in 800ml of distilled water. The pH was adjusted to 5.2 with glacial acetic acid. The solution was made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).

Sephadex G-25 or G-50 resin :- Sephadex G-25 or G-50 resin (30g) was added to TE (X1) (250ml) and sterilized by autoclaving (15 mins, 15psi).

<u>Restriction enzyme reaction buffers(obtained from BRL)</u> :-

Core buffer X10 :- 500mM Tris.HCl (pH8),
100mM $MgCl_2$,
500mM NaCl.

React 1 buffer :- 50mM Tris.HCl (pH8),
10mM $MgCl_2$,

React 2 buffer :- 50mM Tris.HCl (pH8),

10mM $MgCl_2$,

50mM NaCl.

React 3 buffer :- 50mM Tris.HCl (pH8),

10mm $MgCl_2$,

100mM NaCl.

Solutions required for isolation of fragments by electroelution from gel slices :-

Electroelution tank buffer X10 :- 100mM Tris.HCl (pH8),

10mM EDTA,

5mM NaCl.

High salt buffer :- 3M Sodium acetate (pH7.9),

Bromophenol blue (to colour).

Solutions required for DNA cloning :-

Calf intestinal phosphatase storage buffer :- 30mM Triethanolamine,

3M NaCl,

1mM $MgCl_2$,

1mM $ZnCl_2$.

(pH7.6).

The buffer was sterilised by autoclaving

(15 mins, 15psi).

Calf intestinal phosphatase reaction buffer :- 500mM Tris.HCl (pH7.5),
10mM $MgCl_2$,
1mM $ZnCl_2$.
The buffer was sterilised by autoclaving (15 mins, 15psi).

Ligase reaction buffer X10 :- 500mM Tris.HCl (pH7.5),
100mM $MgCl_2$,
100mm DTT,
0.1% BSA.
The buffer was filter sterilised and stored in small aliquots at -20°c.

M13 phage buffer :- 3g KH_2PO_4,
7g Na_2HPO_4,
5g NaCl.
The buffer was made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).
The following were made and autoclaved separately,
(a) 1M $MgSO_4$,
(b) 0.1M $CaCl_2$,
(c) 1% (w/v) Gelatin.
1ml each of (a), (b) and (c) were added to 1l of phosphate buffer mix, making sure that all solutions were cold to prevent precipitation of $Ca_3(PO_4)_2$ and $Mg_3(PO_4)_2$.

Solutions required for dideoxy sequencing :-

PEG /NaCl :- 2.5M NaCl,
10% PEG 8000.
The buffer was sterilised by autoclaving
(15 mins, 15psi).

TM buffer :- 100mM Tris.HCl (pH8.5),
50mM $MgCl_2$.
The buffer was sterilised by autoclaving
(15 mins, 15psi).

Dideoxy sequencing mixes :-

	T	C	G	A
0.5mM dTTP	6.25µl	125µl	125µl	125µl
0.5mM dCTP	125µl	6.25µl	125µl	125µl
0.5mM dGTP	125µl	125µl	6.25µl	125µl
4mM ddTTP	31.18µl	-	-	-
4mM ddCTP	-	5µl	-	-
4mM ddGTP	-	-	10µl	-
4mM ddATP	-	-	-	0.75µl
TE buffer(x1)	231.3µl	247µl	244µl	125µl

The mixes were stored in small aliquots at -20°c.

Chase mix :- 0.25mM dNTPs in TE buffer (X1).

Ribo & Deoxyribonucleic acid stock solutions :- These were dissolved in water to 10mM and the pH adjusted to 7 using a 0.005M solution of Tris base. The stock solutions were filter sterilised and stored in small aliquots at -20°c.

40% acrylamide stock solution :- 83g Acrylamide,

2g Bis Acrylamide.

The solution was made to 100ml with water, deionised using Amberlite ion exchange resin and stored in the dark at 4°c.

20% Acrylamide /Urea gel mix (for oligonucleotide purification) :-

100g Urea,

20ml X10 TBE,

100ml 40% Acrylamide.

The mix was made to 200ml with distilled water and stored in the dark at 4°c.

6% Acrylamide /Urea gradient gel mixes (For sequencing gels):-

0.5X MIX :- 430g Urea,

50ml X10 TBE,

150ml 40% Acrylamide,

The mix was made to 1l with water and stored in the dark at 4°c.

2.5X MIX :- 430g Urea,

250ml X10 TBE,

150ml 40% Acrylamide,

50g Sucrose,

50mg Bromophenol blue.

The mix was made to 1l with water and stored in the dark at 4°c.

Oligonucleotide gel loading buffer :-8ml Formamide (deionised),

100µl 0.5ml EDTA (pH7),

2mg Bromophenol blue,

5mg Xylene cyanol.

The mixture was made to 10ml with distilled water.

M13 sequencing gel loading buffer X5 :- 100ml Formamide (deionised),
0.1g Bromophenol blue,
0.1g Xylene cyanol
The mixture was made to 10ml with distilled water.

Solutions required for *S. aureus* transformations :-

Hypertonic buffered medium (HBM) :- 0.7M Sucrose,
0.02M Maleate,
0.02M $MgCl_2$.
The buffer was adjusted to 6.2 with NaOH and penassay broth added (43g/l). The sediment was filtered away and the broth sterilised by autoclaving (15 mins, 15psi).

Chloramphenicol stock solution :- Chloramphenicol was made to 35mg/ml in absolute ethanol, filter sterilised and stored in small aliquots at -20°c.

Solutions required for the *S. aureus in vitro* replication system :-

S. aureus cell free extract buffer :- 25mM HEPES (pH8),
100mM KCl,
1mM DTT.
The buffer was filter sterilised and stored at -20°c.

S. aureus in vitro replication mixes :-

(a) dNTP mix :- 1M HEPES (pH8) 200µl,
1M Magnesium acetate 100µl,
10mM dATP 12.5µl,
10mM dCTP 12.5µl,
10mM dGTP 12.5µl,
10mM dTTP 12.5µl,
10mM NAD 25µl,
10mM cAMP 300µl.

(b) Energy mix :- 0.1M dATP 200µl,
0.5M Creatine phosphate 300µl,
10mg/ml Creatine kinase 100µl.

(c) rNTP mix :- 40mM rCTP 30µl,
40mM rGTP 30µl,
40mM rUTP 30µl,
Distilled water 210µl.

100% TCA solution :- 227ml of distilled water was added to 500g TCA.

Topoisomerase I reaction buffer X10 :- 500mM Tris.HCl (pH7.5),
500mM KCl,
100mm $MgCl_2$,
5mM DTT,
1mM EDTA,
300µg/ml BSA.

The buffer was filter sterilised and stored in small aliquots at -20°c.

Novobiocin stock solution :- Novobiocin was dissolved to 10mg/ml in distilled water, filter sterilised and stored in small aliquots in the dark at -20°c.

Solutions required for DNase I footprinting :-

Klenow storage buffer :- 0.5 M Tris.HCl (pH7.5),
1M Ammonium sulphate,
0.1M 2-Mercaptoethanol,
10mg/ml BSA.
The buffer was filter sterilised and stored in small aliquots at -20°c.

Klenow reaction buffer X10 :- 0.5M Tris.HCl (pH7.2).
0.1M $MgSO_4$,
1mM DTT.
The buffer was filter sterilised and stored at in small aliquots at -20°c.

T_4 kinase reaction buffer X10 :- 1M Tris.HCl (pH8.3),
0.1M $MgCl_2$.
The buffer was sterilised by autoclaving (15 mins, 15psi).

DNase I storage buffer :- 10mM Tris.HCl (pH7.5),
10mM $CaCl_2$,
1mM DTT,
500µg/ml BSA.
The buffer was filter sterilised and stored in small aliquots at -20°c.

DNase I reaction buffer X10 :- 400mM Tris.HCl (pH7.5),
60mM $MgCl_2$.
The buffer was sterilised by autoclaving (15 mins, 15psi).

Solutions required for *in vitro* mutagenesis :-

Tetracycline stock solution :- Tetracycline was made to 12.5mg/ml in 50% ethanol, filter sterilised and stored in small aliquots in the dark at -20°c.

DTT stock solution :- DTT was dissolved to 1M in distilled water, filter sterilised and stored in small aliquots at -20°c.

Mutagenesis buffer A X10 :- 0.2M Tris.HCl (pH7.5),
0.1M $MgCl_2$,
0.5M NaCl.
The buffer was sterilised by autoclaving (15 mins, 15psi) and mixed 9:1 with 1M DTT before use.

Mutagenesis buffer B X10 :- 0.2M Tris.HCl (pH7.5),
0.1M $MgCl_2$.
The buffer was sterilised by autoclaving (15 mins, 15psi) and mixed 9:1 with 1M DTT before use.

REP D *in vitro* nicking /closing reaction buffer :- 100mM Tris.HCl (pH8),
0.1M $MgCl_2$,
1M KCl,
10mM EDTA.
The buffer was sterilised by autoclaving (15 mins, 15psi).

Solutions required for studies on the REP D /oriD interaction using synthetic oligonucleotides :-

REP D oligonucleotide nicking buffer :- 40mM Tris.HCl (pH8),
40mM $MgCl_2$.
The buffer was sterilised by autoclaving (15 mins, 15psi).

2.1.5: - Microbiological growth media.

L.B. broth :- 10g Bacto tryptone,

5g Bacto yeast extract,

10g NaCl.

The broth was made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).

2YT broth :- 16g Bacto tryptone,

10g Bacto yeast extract,

5g NaCl.

The broth was made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).

M9 minimal medium :- 6g Na_2HPO_4,

3g KH_2PO_4,

0.5g NaCl,

1g NH_4Cl.

The broth was made to 1l with distilled water, adjusted to pH7.4 and sterilised by autoclaving (15 mins, 15psi). The following were added :-

(a) 2ml 2M $MgSO_4$,

(b) 10ml 20% Glucose,

(c) 0.1ml 1M $CaCl_2$.

These were sterilised separately by filtration (glucose) or autoclaving. (15 minutes, 15psi).

Agar media :- Media were made up as normal and 15g (for plates) or 7g (for soft agar) of Bacto agar was added. The media were made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).

DNase test agar :- 33g Difco tryptose blood agar base,

2g Salmon testes DNA,

0.8g $CaCl_2$.

The broth was made to 1l with distilled water and sterilised by autoclaving (15 mins, 15psi).

DM3 wall regenerating medium (for *S. aureus* transfomations) :-

200ml 4% agar (2% for soft agar),

500ml Sodium succinate,

125ml distilled water.

The broth was sterilised by autoclaving (15 mins, 15psi) and the following was added:-

(a) 25ml 20% Difco yeast extract,

(b) 50ml 10% Difco casamino acids,

(c) 50ml 7% K_2HPO_4,

(d) 25ml 20% Glucose,

(e) 20ml 1M $MgCl_2$,

(f) 5ml 2% BSA.

These were made and sterilliised (by filtration or autoclaving) separately.

2.1.7: - Oligonucleotides.

OLIGONUCLEOTIDE	SIZE (bases)	APPLICATION	SEQUENCE 5'----------3'
DFB - 1	17	repN sequence	GGGTGAGCGACTTCTTA
DFB - 2	17	"	TAAGAAGTCGCTCACCC
DFB - 3	17	"	TATGAGTGTCGAGCCAC
DFB - 4	17	"	GTGGCTCGACACTCATA
DFB - 5	17	"	GGAGGAATAATCCTCTG
DFB - 6	17	"	CGCATGAAGAAATGCTT
DFB - 7	17	"	AAGCATTTCTTCATGCG
DFB - 8	20	"	GGTGTTCGTGACAGTGATAG
DFB - 9	20	"	CTATCACTGTCACGAACACC
DFB - 10	20	"	CATCGTTCCAATAATCAACC
DFB - 11	20	"	ATCGACTTTAAAAGCGAACG
DFB - 12	20	"	CGTTCGTTTTAAAGTCGTA
DFB - 13	20	"	AATGATGAAGAAGAATGGGG
DFB - 14	20	"	GACCTGATCTAATTTAGACC
DFB - 15	20	"	GGTCTAAATTAGATCAGGTG
DFB - 16	20	"	CCCGTACCGACATTTTAATT
DFB - 17	21	oriD mutagenesis	CTCTAATAGCGGGTTAAGTGG
DFB - 18	21	"	TTCTAAAACCCGCTACTCTAA
DFB - 19	24	REP D specificity studies.	AACCGGCTACTCTAATAGCCGGTT
DFB - 20	24	"	AACCGGCTACTCTGATAGCCGGTT
DFB - 21	24	"	AACCGGCTACTCTCATAGCCGGTT
DFB - 22	24	"	AACCGGCTACTCTTATAGCCGGTT
DFB - 23	24	"	AACCGGCTACTCCAATAGCCGGTT

OLIGONUCLEOTIDE	SIZE (bases)	APPLICATION	SEQUENCE 5'----------3'
DFB - 24	24	REP D specificity studies.	AACCGGCTACTCAAATAGCCGGTT
DFB - 25	24	"	AACCGGCTACTCGAATAGCCGGTT
DFB - 26	24	"	ACTGCAGACCTCTAATAGCCGGTT
DFB - 27	24	"	AACCGGCTACTCTAAGTCTGCAGT
DFB - 28	24	"	ACTGCAGACCTCTAAGTCTGCAGT

2.3: - Methods.

2.3.1: - Storage of bacterial strains.

For short term storage, strains were streaked out on agar plates, with appropriate selection if necessary, grown up overnight and stored at 4°c wrapped in parafilm. Single colonies were picked for growth and plates were re-streaked every few weeks.

For long term storage, strains were stored in liquid medium containing 15% glycerol at -70°c. A few cells were scraped from the surface of the glycerol stock, grown overnight in liquid medium and streaked on agar plates to give single colonies before use.

2.3.2: - Preparation of plasmid DNA.

Isolation of plasmid DNA from *E. coli* (Birnboim & Doly, 1979).

Cultures were grown overnight with shaking at the appropriate temperature. Cells were harvested by centrifugation (8,000rpm, 10 minutes) and resuspended in cold lysis buffer. Cells were lysed by incubation with lysozyme (1mg/ml) on ice for 5 minutes, followed by treatment with 2 volumes of alkaline SDS and 0.5 volumes of potassium acetate (pH 4.8) on ice. Cell lysates were cleared by centrifugation (15,000rpm, 10 minutes) and the supernatant filtered

through glass wool. DNA was precipitated from the supernatant with isopropanol and pelleted by centrifugation (19,000rpm, 30 minutes). The DNA pellet was rinsed with cold 70% ethanol, dried under vacuum and resuspended in TE buffer (X1).

Isolation of plasmid DNA from *S. aureus*.

Plasmid DNA was isolated from *S. aureus* essentially via the method of Birnboim & Doly, exept that, before lysis, cells were washed with cold 10mM EDTA and resuspended in the same. The cells were then treated with 2 volumes of cold 1:1 acetone /ethanol on ice. Cells were pelleted by centrifugation (8,000rpm, 10 minutes), resuspended in lysis buffer and incubated with lysostaphin (50mg/ml) at 37°c for 30 minutes. The cells were then lysed, as before, by treatment with alkaline SDS and potassium acetate (pH 4.8) on ice. Cell lysates were cleared by centrifugation (15,000rpm, 10 minutes) and DNA was precipitated from the supernatant with isopropanol. The DNA was pelleted by centrifugation (19,000rpm, 30 minutes), rinsed with cold 70% ethanol, dried under vacuum and resuspended in TE buffer (X1).

Small scale plasmid isolation from *E. coli* (Plasmid minipreps).

Cultures (5ml) were grown overnight at the appropriate temperature and 1.5ml were transferred to an Eppendorf tube. Cells were pelleted by centrifugation, resuspended in 100µl of lysis buffer containing lysozyme (4mg/ml) and incubated on ice for 5 minutes. Cells were lysed by treatment with alkaline SDS and potassium acetate (pH 4.8) and cell lysates were cleared by

centrifugation. The supernatant was extracted with phenol and the DNA precipitated with ethanol (see section 2.3.3.). The DNA was pelleted by centrifugation, rinsed with cold 70% ethanol, dried under vacuum and resuspended in TE buffer (X1) containing RNase (50μg/ml).

Isolation of supercoiled plasmid DNA by CsCl /Ethidium bromide density gradient centrifugation.

Plasmid DNA isolated from *E. coli* or *S. aureus* was treated with preboiled RNase, then self-digested pronase (both at 50μg/ml) for 1 hour at 37°c to remove RNA and protein (this was particularly important for material prepared from *S. aureus*, as this tends to have a very high RNA and protein content). The DNA was then precipitated with ethanol, pelleted by centrifugation, washed with cold 70% ethanol and resuspended in TE buffer (X1). To each ml of DNA solution, 1g of solid CsCl and 80μl of 10mg/ml ethidium bromide were added and the refractive index of the solution was adjusted to 1.3860 (Maniatis *et al*, 1982). The solution was centrifuged at 48,000rpm for 18 hours in a Beckman vertical rotor (VTi50 or VTi65) with the brake off. Supercoiled plasmid DNA (lower band) was separated from relaxed plasmid DNA and chromosomal DNA (upper band). The bands were visualised under UV light and the lower band was drawn out of the tube using a needle and syringe. Ethidium bromide was removed by extraction with isopropanol saturated with CsCl (2 volumes) 2-3 times. CsCl was removed by dialysis against TE buffer (X1), followed by ethanol precipitation of the DNA. It was usually necessary to carry out two CsCl /Ethidium bromide density gradient spins in order to obtain mostly the supercoiled form of the DNA.

2.3.3: - DNA manipulations.

The removal of protein from DNA samples by phenol extraction.

The DNA samples were vortexed with an equal volume of Phenol / Chloroform /Isoamyl alcohol to give an emulsion. Organic and aqueous layers were separated by centrifugation in a microfuge for 10 minutes and the upper, aqueous layer was transferred to another tube using a pipette, with care taken not to disturb the denatured protein at the interface. The organic layer and the denatured protein were discarded. An equal volume of 24:1 chloroform /isoamyl alcohol was added to the aqueous phase and the above repeated. The DNA was then recovered by ethanol precipitation.

Ethanol precipitation of DNA.

DNA was precipitated from aqueous solution by addition of 1/10 volumes of 3M Sodium acetate and 2.5 volumes of ethanol, followed by cooling (-70°c for 30 minutes). DNA was pelleted by centrifugation (19,000 rpm, 30 minutes, 4°c) and the pellet was washed with cold 70% ethanol, dried under vacuum and resuspended in TE buffer (1X).

Determination of DNA concentration in aqueous solution.

The concentration of DNA in aqueous solution was determined by measuring the absorbance at 260nm.

1 A_{260} unit = 50μg of double-stranded DNA.

For single-stranded DNA :-

1 A_{260} unit = 31μg of single-stranded DNA.

For smaller amounts of DNA, which could not be measured by absorbance determination, the concentration was estimated by visualisation of a small sample on an agarose gel and comparison with known amounts of standard DNA.

Agarose gel electrophoresis of DNA.

Agarose was added to TBE or TAE electrophoresis buffer (X1) to a concentration of 0.8-1% and heated gently until boiling to dissolve the agarose. The solution was allowed to cool for a short while and ethidium bromide was added to 0.1μg/ml (if required). The gel was poured onto a plate with the comb in position and allowed to set. The comb was removed and the gel was placed into an electrophoresis tank containing electrophoresis buffer (X1) with, if required, ethidium bromide (0.1μg/ml). Agarose gel loading buffer (X6, 1/6 volume) was added to the DNA samples, which were then applied to the wells. The gel was run at 100V until the bromophenol blue dye was approximately 2/3 of

the way along the gel. If necessary, the gel was stained for 30 minutes in 0.1µg/ml ethidium bromide and the DNA bands were visualised under UV light.

Silanising of glassware for DNA work.

All glassware for use with DNA, glass plates for acrylamide gel electrophresis and glass wool were cleaned thorougly and silanised by treatment with dimethyl dichloro silane solution (2% in CCl_4). This was allowed to dry on the glass (68°c) and any excess, along with HCl produced by the silanising reaction, was rinsed off with distilled water , then ethanol. The glassware and glass wool was then allowed to dry and sterilised by autoclaving (15 minutes , 15psi) before use.

Restriction digestion of DNA.

DNA (typically 0.01 - 1µg) was incubated with up to 5 units of the appropriate restriction endonuclease in a total volume of 10µl of the appropriate react buffer. (React 1 buffer was used for AccI, AluI and HpaII, React 2 buffer was used for BstEII, HindIII, HinfI and TaqI and React 3 buffer was used for BamHI, BglII and EcoRI). The samples were incubated for 4-12 hours at the appropriate temperature.

Isolation of DNA Fragments from agarose gels by the "freeze-squeeze" method (Tautz & Renz, 1983).

The required DNA fragment was excised from an agarose gel and the gel slice was equilibrated with 0.3M Sodium acetate (pH7) /1mM EDTA in the dark for 30 minutes. The bottom of a small Eppendorf tube was pierced and plugged with silanised glass wool. The gel slice was placed inside and the tube was immersed in liquid nitrogen to freeze the gel. The tube was then inserted into a large, capless Eppendorf tube and centrifuged for 10 minutes in a microfuge. DNA was precipitated from the solution collected in the large tube with ethanol and the DNA pellet was washed with cold 70% ethanol, dried under vacuum and resuspended in TE buffer (X1).

Isolation of DNA fragments from agarose gels by electroelution.

The required DNA fragment was excised from an agarose gel and the gel slice placed into a well of an electroelution tank. The tank was filled with electroelution tank buffer (X1) until the wells and V-shaped channels were just full. High salt buffer (100µl) was placed in the bottom of the V-shaped channel and the tank connected to a power pack run at 150V for 45 minutes. The tank was disconnected from the power pack and the DNA (300µl) was collected from the bottom of the V-shaped channel using a drawn-out capillary tube. The DNA was recovered from the high salt buffer by passage through a G-25 spun column followed by ethanol precipitation.

Purification of DNA samples using Sephadex spun columns.

A sterile 1ml plastic syringe was plugged with silanised glass wool and placed inside a sterile test tube. Sterile Sephadex G-25 or G-50 resin in TE buffer (1ml) was added to the syringe and the tube was spun in a bench centrifuge until the speed reached 3,000rpm and the centrifuge was then switched off. More Sephadex was applied to the syringe and the tube centrifuged again, as before, until the syringe was almost full. Sterile water or TE buffer (X1) (100µl) was applied to the column and the tube centrifuged, as before, to rinse the column. This was repeated 2-3 times. The syringe was then placed into a new sterile test tube, and the DNA solution (100µl) was applied to the top of the column. The tube was centrifuged, as before, and the desalted DNA solution was collected from the test tube with a pipette. The solution was then centrifuged to remove any traces of Sephadex.

2.3.4: - DNA cloning and sequencing.

Preparation of vector DNA (M13).

Vector DNA (2µg) was cut with the appropriate restriction endonuclease (1 unit) in a total volume of 20µl. 5' phosphate groups were removed from the cut vector to prevent it's re-circularisation by treatment with calf intestinal phosphatase (CIP). The cut vector was made to a volume of 45µl with sterile water. Phosphatase buffer (X10, 5µl) and CIP (0.01 units per pmol 5' ends) were added and the solution incubated at 37°c for 30 minutes. A futher aliquot of CIP was added, followed by further incubation at 37°c for 30 minutes.

The vector DNA was then phenol extracted twice, ethanol precipitated and resuspended in TE buffer (X1).

Preparation of insert DNA.

DNA (1μg) was cut with the appropriate restriction endonuclease and, if necessary, the required DNA fragment was isolated from an agarose gel, passed through a Sephadex G-25 spun column, phenol extracted, ethanol precipitated and resuspended in TE buffer (X1).

Ligation reactions.

Approximately 10ng of vector DNA were used per ligation reaction and between 10:1 and 50:1 molar ratios of insert to vector. Ligase buffer (X10, 1μl), ATP (0.1M, 1μl) and T_4 DNA ligase (200 units) were added. The ligation mixtures were made to a total volume of 10μl with sterile water and incubated at 15°c overnight.

Transformation of *E. coli* JM101 with M13.

Overnight *E. coli* JM101 cultures (0.1ml) were used to innoculate 10mls of 2YT broth. This was incubated at 37°c with shaking until growth was at mid-exponential phase (OD_{600} approximately 0.77). Cells were harvested by centrifugation and resupended in cold 50mM $CaCl_2$ (5ml). Cells were incubated on ice for 30 minutes and then pelleted by centrifugation. Cells were resuspended

in 1ml of 50mM $CaCl_2$. These competent cells were divided into 200µl aliquots. Each aliquot was incubated on ice for 30 minutes with 5µls of ligation mix. The cells were heat shocked at 42°c for 2 minutes and then put on ice. Transformed cells were added to 3mls of molten 2YT top agar (42°c) with 40µl of 100mM IPTG and 40µl of 24% X-gal (in Dimethyl formamide) and poured onto dried 2YT agar plates. The plates were allowed to set, inverted and incubated at 37°c overnight.

Picking of M13 plaques.

M13 grow to form plaques within a lawn of *E. coli* after transformation. Wild type M13 give rise to blue plaques due to breakdown of the added X-gal by the M13 encoded β galactosidase to produce a blue dye. M13 containing inserts give rise to white plaques due to inactivation of the M13 lacZ gene by the insertion.

White plaques were picked by plunging a sterile toothpick into the plaque and transferring the phage into 1ml of sterile M13 phage buffer. In this way, phage could be stored at 4°c for several months.

Preparation of phage M13 single-stranded DNA.

Overnight *E. coli* JM101 cultures (1.5µl) were used to innoculate 1.5mls of 2YT broth. Phage stock (0.1ml) was added and cells were grown overnight with shaking at 37°c. Cultures were transferred to Eppendorf tubes and the cells were pelleted by centrifugation. The supernatant (1ml) was collected using a pipette, avoiding disturbance of the cell pellet. Phage was

precipitated by adding 1/3 volumes of sterile 2.5M NaCl /10% PEG 8000 and pelleted by centrifugation in a microfuge for 10 minutes. The supernatant was removed using a drawn out Pasteur pipette. The pellet was resuspended in sterile water and an equal volume of Phenol /Chloroform /Isoamyl alcohol was added. The mixture was vortexed to give an emulsion and incubated at room temperature for 15 minutes. The mixture was vortexed again , centrifuged for 5 minutes and 3/4 of the supernatant was collected with care taken not to disturb the denatured protein at the interface. The supernatant was then extracted again with 24:1 Chloroform /Isoamyl alcohol to remove traces of phenol and the DNA was precipitated with ethanol. The DNA was pelleted by cenrifugation and the DNA pellet was rinsed with cold 70% ethanol, dried under vacuum and resuspended in 25µl of sterile water.

Polyacrylamide gel purification of oligonucleotides.

Oligonucleotides, synthesised manually using phosphotriester chemistry with paper discs as a support (Brenner & Shaw, 1985), or by automated synthesis using an Applied Biosystems 380B DNA synthesiser, were ethanol precipitated and the concentration determined by measurement of the absorbance of the solution at 260nm. The oligonucleotides (2-4µg) in a volume of 2-4µl were taken and 2µl of oligonucleotide gel loading buffer was added.

20% Acrylamide /50% Urea gel mix was put into a side arm flask and a 1/150th volume of 10% Ammonium persulphate was added. The solution was degassed under vacuum and a 1/1500th volume of TEMED was added. The mixture was swirled and quickly poured between silanised glass plates which had been taped together with 0.4mm spacers, with care being taken not to introduce air

bubbles into the gel. A sharks-tooth comb was inserted and the gel was clamped and allowed to set. The tape was then removed from the bottom of the gel, the comb inverted and the gel clamped to a vertical gel tank containing TBE electrophoresis buffer (X1). The oligonucleotide samples containing dye were boiled for 1 minute and loaded into the wells. The gel was run at 1400V, 30mA and 30W for 2-3 hours. The plates were prised apart and the gel covered with cling film on both sides. The oligonucleotides were visualised by UV shadowing onto a silica TLC plate and their positions marked.

The oligonucleotides were cut from the gel and eluted into 100µl of sterile water overnight. The oligonucleotides were desalted by passage through a Sephadex G-25 spun column and the concentration was determined by measuring the absorbance at 260nm.

Priming of M13 single-stranded templates.

The following were mixed in a 1.5ml Eppendorf tube :-

Primer oligonucleotide (5ng/ml)	1µl
TM buffer	1µl
Template DNA (0.1µg/ml)	5µl
Sterile water	3µl

The above mixture was heated in a boiling water bath for 5 minutes. The water bath was then switched off and the priming mixture allowed to cool in the water bath for 10 minutes. The tube was removed from the water bath and allowed to cool at room temperature for a further 10 minutes. Any condensate on

the side of the tube was spun to the bottom by centrifugation for a few seconds.

Dideoxy sequencing reactions.

For each primed template, four capless Eppendorf tubes were labelled A,C,G and T respectively. The following was added to each tube :-

TUBE	A	C	G	T
PRIMED TEMPLATE	2µl	2µl	2µl	2µl
A MIX	2µl	-	-	-
C MIX	-	2µl	-	-
G MIX	-	-	2µl	-
T MIX	-	-	-	2µl
KLENOW/^{35}SdATP	2µl	2µl	2µl	2µl

(Klenow/[^{35}S] dATP consists of 1:1:10 Klenow fragment (4units/µl), α[^{35}S] dATP (400ci/mM, 10mCi/ml) and sterile water).

The reactions were incubated at room temperature for 20 minutes. Chase mix (2µl) was added and the reactions incubated for a further 30 minutes at room temperature. The reactions were then stored at -20°c until use.

Acrylamide /Urea gradient gels for DNA sequencing.(Biggin *et al*, 1983).

2.5X TBE /Acrylamide /Urea mix (15ml) and 0.5X TBE /Acrylamide /Urea mix (60ml) were put into separate side arm flasks. 10% Ammonium persulphate (90 and 360µl respectively) was added and the solutions degassed under vacuum.

TEMED (5 and 15µl respectively) was added, the flasks swirled and, using a 20ml graduated pipette, 8ml of 0.5X mix, followed by 12ml of 2.5X mix were taken up. Approximately four air bubbles were introduced into the pipette to mix the solutions, which were then carefully poured between two silanised glass plates (30cm × 40cm) which had been taped together with 0.4mm spacers. The remaining 0.5X mix was then carefully poured between the plates to give a gradient, with care being taken not to introduce air bubbles into the gel.

Sharks-tooth combs were inserted and the gel was clamped and allowed to set. The tape was then removed from the bottom of the gel and the combs inverted. The gel was clamped to a vertical electrophoresis tank containing X1 TBE buffer in the bottom tank and X½ TBE buffer in the top tank. The gel was pre-run at 1600V, 30mA and 45W for 30 minutes before loading.

Immediately before loading, M13 sequencing gel loading buffer (2µl) was added to the sequencing samples, which were then boiled for 3 minutes and loaded directly onto the gel. The gel was run for approximately 2 hours, until the Bromophenol blue dye was just running off the end of the gel. The plates were prised apart and the gel fixed in 10% Methanol /10% Acetic acid for 20 minutes, blotted onto Whatman 3MM paper, dried on a Bio-rad gel drier at 80°c and exposed to Fuji RX X-ray film overnight.

2.3.5: - *S. aureus in vitro* replication system.

Test for nuclease deficient *S. aureus*.

S. aureus strain RN1786 was tested for nuclease deficiency by streaking onto DNase test agar and incubation at 37°c overnight. The test agar contains $CaCl_2$ and DNA and if flooded with 1M HCl, the DNA precipitates and the agar becomes cloudy. When "wild type" *S. aureus* was grown on the test agar, there was a clear region around the bacteria on flooding with 1M HCl due to degradation of the DNA by Staphylococcal nucleases. Strain RN1786, which is nuclease deficient, showed no such clear region on the plate.

Transformation of *S. aureus* by the protoplast method (Lindberg, 1981).

S.aureus transformation was carried out by the protoplast method of Lindberg (1981), *S.aureus* RN1786 was grown overnight with shaking at 37°c in 2YT broth (50ml). Cells were harvested by centrifugation and resuspended in hypertonic buffered medium (HBM) (10ml). Lysostaphin was added to 20µg/ml and the cells were incubated at 37°c for 3 hours with gentle shaking. The cells were centrifuged at 6,000rpm for 15 minutes to remove cell debris. The supernatant was then centrifuged at 19,000rpm for 15 minutes to pellet the protoplasted cells. The protoplasts were resuspended in HBM (1.6ml) and divided into 0.2ml aliquots. DNA (10ng in TE buffer) was added to the aliquots for transformation with gentle shaking. 40% PEG 6000 (1.8ml) was added, followed by more HBM (10ml) after 2 minutes. The protoplasts were pelleted by centrifugation at 19,000rpm for 30 minutes and resuspended in HBM (1ml). Dilutions were plated

out on DM3 wall regenerating medium agar plates and incubated at 37°c for 3 hours. Selective top agar (3ml) was then poured over the plate and allowed to set. The plates were inverted and incubated overnight at 37°c. For transformation with pC221, top agar containing chloramphenicol at 30µg/ml was used. Chloramphenicol resistant colonies were picked and streaked out on nutrient (2YT) agar plates, containing chloramphenicol at 10µg/ml, to give single colonies.

Preparation of *S. aureus* cell free extracts.

Extracts were prepared from a strain of *S. aureus* deficient in nuclease (RN1786). 2YT broth (1l in two 2.5l flasks) was inoculated with 1ml of an overnight culture of *S. aureus* RN1786. The cultures were grown at 37°c, with gentle shaking, until mid-exponential phase was reached. The cells were harvested by centrifugation and washed with cold buffer A (25mM HEPES (pH8) /100mMKCl /1mMDTT).

The cells were weighed and resuspended in an equal volume of buffer A. The cells were quickly frozen in liquid nitrogen and thawed. Lysostaphin was added to 200µg/ml and the cells were incubated on ice for 1-2 hours. The cells were then quickly frozen again in liquid nitrogen and thawed to lyse. The lysates were cleared by centrifugation (45,000rpm, 1 hour, 4°c) and the supernatant decanted. Ammonium sulphate was added to the supernatant to 70% saturation (0.472g/ml) and the solution was cooled on ice for 30 minutes with stirring. The ammonium sulphate precipitate was pelleted by centrifugation

(19,000rpm, 30 minutes, 4°c) and the pellet washed with saturated ammonium sulphate.

The pellet from 5g original cell weight was resuspended in 1ml of buffer A (+ 10% ethylene glycol) the extract was dialysed against 500ml of the same for 1½ hours (with two changes of buffer). Small aliquots were quickly frozen in liquid nitrogen and stored at -70°c.

In vitro replication of pC221 and related plasmids using *S. aureus* cell free extracts.

The *S. aureus* cell free extracts were assayed by their ability to incorporate radioactively labelled deoxyribonuceotide triphosphates into acid precipitable material in the presence of added plasmid DNA and purified REP D protein. Typical assay mixtures contained the following :-

S. aureus cell free extract	5µl
Plasmid DNA	0.5 µg
REP D protein	10-50ng
Replication cocktail	5µl
PEG 4000 (25% w/v)	2.5µl

Replication cocktail (per 5µl) :-

dNTP mix	1.75µl
energy mix	1.5µl
rNTP mix	0.75µl
[Me-^{3}H] dTTP	1.0µl (1µCi)

The mixture was made to a total volume of 25µl with buffer A and incubated for, typically, 1 hour at 30°c. The reaction was stopped by addition of 1ml of cold 10% Trichloroacetic acid solution (TCA) and the mixtures were cooled on ice for 20 minutes. Acid precipitable material was collected on Whatman GF/C filters, which were then rinsed with cold 10% TCA, followed by ethanol, dried and placed into 4ml of a toluene based scintillation fluid (Fisons optiscint T) in a plastic scintillation vial. Acid precipitable radioactivity was determined by liquid scintillation counting.

Restriction digestion, gel electrophoresis and autoradiography of plasmid DNA labelled with α[^{32}P] dATP using the *in vitro* replication system.

Assays were carried out as normal, exept that the added label was 1µCi of α[^{32}P] dATP instead of [Me-^{3}H] dTTP. After incubation at 30°c for 1 hour, the reaction was stopped by phenol extraction and ethanol precipitation. The pellet was resuspended in 16µl of TE buffer and X5 agarose gel loading buffer (4µl) was added. The samples were run on a 0.8% agarose gel, which was then stained in 0.1µg/ml ethidium bromide, dried between two sheets of Whatman 3MM paper for 2 hours at 37°c and wrapped in cling-film. The bands were visualised under UV light and the positions of the bands marked. The gel was then exposed to Fuji RX X-ray film overnight.

If necessary, the labelled DNA was digested with 1 unit of the appropriate restriction endonuclease for 4-12 hours before agarose gel electrophoresis.

Relaxation of supercoiled plasmid DNA using DNA topoisomerase I.

Supercoiled plasmid DNA (5μg) was treated with 10 units of calf thymus DNA topoisomerase I in a total volume of 50μl of DNA topoisomerase reaction buffer for 2-4 hours at 37°c. A small sample of the DNA was subjected to agarose gel electrophoresis to check that the relaxation of the DNA had gone to completion. The relaxed plasmid DNA was phenol extracted, ethanol precipitated and resuspended in TE buffer.

2.4.6: - DNase I footprinting of the pC221 origin (oriD) with REP D protein.

Isolation of a 976bp restriction fragment of pC221 containing the origin sequence

pC221 DNA (approximately 200μg) was digested with HinfI (20 units) at 37°c overnight. The digested DNA was run on a 0.8% agarose gel and the largest, 976bp fragment, which contains oriD, was excised from the gel. The restriction fragment was isolated from the gel slice by electroelution, followed by ethanol precipitation and then resuspended in sterile distilled water.

Removal of 5' phosphate groups from the 976bp fragment.

The 976bp HinfI fragment (approximately 20μg) was dissolved in 45μl of sterile water. Calf intestinal phosphatase (200 units) and X10 phosphatase buffer (5μl) were added and the mixture incubated for 30 minutes at 37°c. A

further aliquot of phosphatase was added, followed by further incubation for 30 minutes at 37°c. The mixture was then phenol extracted and the DNA ethanol precipitated.

5' end-labelling of the 976bp fragment.

The 976bp HinfI fragment (approximately 20μg), which had been previously treated with phosphatase, was dissolved in 15μl of sterile water. Kinase buffer (X10, 2.5μl), γ [^{32}P] ATP (7.5ul, 75μCi) and T_4 kinase (10-20 units) were added and the mixture incubated at 37°c for 30 minutes.

3' end-labelling of the 976bp fragment.

The 976bp HinfI fragment (approximately 20μg) was dissolved in 14.5μl of sterile water. Nick translation buffer (X10, 2.5μl), dNTPs (dCTP, dGTP and dTTP) (1mM, 3μl), α[^{32}P] dATP (5μl, 50μCi) and Klenow (1 unit) were added and the mixture was incubated at room temperature for 30 minutes.

Purification and restriction digestion of the end-labelled 976bp fragment.

End-labelled (5' or 3') DNA was made to a volume of 100μl with sterile TE buffer and unincorporated radioactive deoxyribonucleotides were removed by passage through a Sephadex G-50 spun column. The DNA was subjected to phenol extraction, ethanol precipitation and resuspended in 18μl of sterile water. The DNA was digested with AluI (20 units) at 37°c overnight. The digested

DNA was run on a 0.8% agarose gel and the end-labelled 186bp fragment, containing oriD, was cut from the gel. The DNA was isolated from the gel slice by electroelution, ethanol precipitated and resuspended in 10-20µl of distilled water. The amount of labelled DNA present was estimated by liquid scintillation counting of a small sample in comparison with a known standard amount of the radioactive label counted in the same manner.

DNase I footprinting reactions with the end-labelled 186bp fragment and REP D.

The end-labelled (5' or 3') 186bp restriction fragment (approximately 10ng) was partially digested by DNaseI. DNase I reaction buffer (X10, 1µl) and DNase I (4ng) were added to the DNA and the volume made to 10µl with sterile water. The mixture was incubated for 5 minutes at room temperature and the reaction stopped by addition of 2µl of M13 sequencing gel loading buffer followed by boiling for 5 minutes to denature the sample.

In order to carry out DNase I footprinting reactions in the presence of REP D, the DNA was preincubated with a 100:1 ratio of REP D to DNA on ice for 15 minutes, then at room temperature for 15 minutes before digestion with DNase I.

All samples were loaded onto a 6% Acrylamide /Urea gradient gel alongside dideoxy sequencing tracks, using known templates and primers, as size markers. The gel was run for approximately 2 hours at 1600V, 30mA and 45W. The gel was fixed, blotted, dried and exposed to Fuji RX X-ray film overnight.

2.3.7. - *In vitro* mutagenesis of oriD by the Kunkel method (Kunkel. 1985).

Cloning of oriD into M13.

A 278bp TaqI restriction fragment of pC221, containing oriD, was cloned into M13 /mp19 via the AccI site (see section 2.3.4.). Cloned DNA was transformed into *E. coli* JM101 and white plaques were picked into phage buffer. Single-stranded and replicative forms of the DNA were prepared and checked by running on a 0.8% agarose gel. HindIII /BglII double digests were carried out on the double-stranded replicative forms of the cloned DNA in order to check the size of the inserts.

Preparation of single-stranded DNA with deoxyuridine incorporated.

E. coli RZ1032 was grown overnight with shaking at 30°c in 2YT broth (10ml) containing tetracycline (12.5µg/ml). A small aliquot of this (500µl) was used to innoculate 2YT broth (500ml) with tetracycline (12.5mg/ml) and deoxyuridine (0.25µg/ml). Phage stock (100µl) was added and the cells grown overnight with shaking at 30°c. Cells were pelleted by centrifugation (10,000rpm, 30 minutes). The supernatant (400ml) was removed, with care being taken not to disturb the cell pellet, and PEG /NaCl (140ml) was added. Phage was precipitated and pelleted by centrifugation (10,000rpm, 30 minutes). The supernatant was discarded and the dregs removed using a drawn out Pasteur pipette. Phage was resuspended in 100ml of sterile water, phenol extracted and the DNA was then precipitated with ethanol. Single-stranded phage DNA was resuspended in sterile

water and the concentration determined by measuring the absorbance of the solution at 260nm. Single-stranded DNA with deoxyuridine incorporated was transformed into *E. coli* JM101 alongside normal single-stranded DNA in order to test for a difference in viability between the two types of DNA in this host. The DNA with deoxyuridine showed an approximately 100 fold decrease in viability in JM101 compared to the normal DNA.

Preparation of mutagenic oligonucleotides.

Oligonucleotides were purified by polyacrylamide gel electrophoresis followed by passage through a Sephadex G-25 spun column. The concentration of oligonucleotide in solution was determined by measurement of the absorbance of the solution at 260nm. The mutagenic oligonucleotide (50pmol) was dried down under vacuum and resuspended in 7µl of distilled water. Kinase buffer (X10, 1µl), ATP (10mM, 1µl) and DTT (10mM, 1µl) were added, followed by 4 units of T_4 kinase. The mixture was incubated at 37°c for 30 minutes, then at 65°c for 15 minutes to inactivate the kinase.

Extension /ligation reactions.

Single-stranded M13 /oriD DNA (0.5pmol), containing deoxyuridine, was dried down under vacuum. Sterile water (7µl), mutagenesis buffer A (X10, 1µl) and kinased oligonucleotide (10mM, 2µl) were added and the mixture boiled for 3 minutes and then incubated at room temperature for 30 minutes to anneal the oligonuclotide and template DNA. Extension /ligation reactions were carried out by addition of mutagenesis B buffer (X10, 1µl), ATP (10mM, 1µl), dNTPs (2mM, 4µl)

and sterile water (4μl). Klenow (2 units) and T_4 DNA ligase (3 units) were added and the mixture incubated at 15°c overnight.

Dilutions of the extension /ligation reaction mixtures were transformed into *E. coli* JM101 and plaques were picked into phage buffer. Single-stranded DNA was prepared from the phage and dideoxy sequencing reactions were carried out using M13 universal primer with these templates to test for the required mutations.

The assay of mutant oriD sequences using the *in vitro* nicking /closing activity of REP D.

Double-stranded replicative forms of M13 /mp19, "wild type" M13 /oriD and the mutant forms of M13 /oriD were prepared on a large scale by the Birnboim and Doly method. DNA (0.5μg, 0.1pmol) was incubated with 0, 0.05, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, and 12.8pmol of REP D respectively in a total volume of 20μl of REP D nicking /closing buffer on ice for 15 minutes , then at 30°c for 15 minutes. Agarose gel loading buffer (X5, 4μl) was added and the samples loaded onto a 0.8% agarose gel in TBE buffer (X1) containing ethidium bromide (0.1 μg/ml).

2.3.8. - Studies on the REP D /oriD interaction using synthetic oligonucleotides.

Purification and 5' end-labelling of oligonucleotides.

Oligonucleotides based on the ICR II sequence of oriD were purified by polyacrylamide gel electrophoresis, followed by passage through a Sephadex G-25 spun column. The oligonucleotides (0.625µg, 100pmol) were dried down under vacuum and resuspended in 50µl of sterile water. Kinase buffer (X5, 13.5µl), $\gamma[^{32}P]$ ATP (2µl, 2µCi) and T_4 DNA kinase (10 units) were added and the mixture incubated at 37°c for 30 minutes, then at 65°c for 15 minutes. End-labelled oligonucleotides were passed through a Sephadex G-25 spun column to remove unincorporated label. The amount of labelled oligonucleotide present was estimated by liquid scintillation counting of small samples of the oligonucleotide in comparison with a known standard sample of the radioactive label counted in the same manner.

Preparation and 5' end-labelling of oligonucleotide size markers.

A mixture of oligonucleotides of 10, 15 and 17 bases were used as size markers. These were not gel purified before use, as the traces of oligonucleotides of other sizes gave rise to a continuous ladder of size markers. A solution (10µl) containing the three oligonucleotides at a concentration of 10µM was prepared and kinase buffer (X5, 5µl), $\gamma[^{32}P]$ ATP (12µl, 12µCi) and T_4 DNA kinase (10 units) were addded. The mixture was incubated at 37°c for 45 minutes, then at 65°c for 15 minutes. The end-labelled size markers were passed through a Sephadex G-25 spun column to remove

unincorporated label. Small samples were subjected to liquid scintillation counting, in comparison with a known standard amount of the radioactive label counted in the same manner, to determine the concentration of labelled oligonucleotide size markers in solution.

Treatment of end-labelled oriD oligonucleotides with REP D.

5' end-labelled oligonucleotides (2µl, 0.01-0.02pmol) were incubated with a 1:1 ratio of REP D to oligonucleotide in a total volume of 8µl of REP D oligonucleotide nicking buffer at 30°c for 10 minutes. Oligonucleotide gel loading buffer (4µl) was added and the samples were boiled for 5 minutes before loading onto a 20% Acrylamide /Urea gel. Size markers (2µl) were also loaded. The gel was run for 2-3 hours at 1400V, 30mA and 30W. The plates were prised apart and the gel covered with cling-film on both sides before exposure to Fuji RX X-ray film overnight.

CHAPTER 3.

THE SEQUENCING OF THE pCW7 repN GENE.

3.1. - Introduction.

3.2. - Results and discussion.

3.2.1. - Cloning and sequencing of the pCW7 repN gene.

3.2.2. - Comparison of the pCW7 repN sequence with five other rep sequences.

CHAPTER 3.

THE SEQUENCING OF THE pCW7 repN GENE.

3.1 - Introduction.

Plasmid pCW7 is a 4.4kb plasmid of *S. aureus* which specifies resistance to chloramphenicol (Figure 3.1) (Wilson et al, 1978). Since it does not appear to belong to the same incompatibility group as any other known staphylococcal plasmid, it was provisionally put into a new staphylococcal plasmid incompatability group, inc 14. Heteroduplex mapping has been carried out between pCW7 and pC221 (Figure 3.2). This showed homology between them, not only over the region encoding chloramphenicol acetyl transferase (CAT), as expected, but also over the region corresponding to the pC221 repD gene. This new rep gene discovered in pCW7 has been called repN (Skinner, 1984).

At the time of discovery of the repN gene, the DNA sequences of only two other rep genes, pT181 repC and pC221 repD, were known. It was therefore thought that it might be useful to have other rep sequences for comparison with those already known, in order to yield further information on conserved regions of the genes, putative secondary structure within the genes, and on the upstream copy control regions. More recently, the sequences of three other rep genes, pS194 repE, pUB112 repI and pC223 repJ, have been determined (Projan *et al*, in press). The following chapter describes the sequencing of the repN gene and its comparison with five other rep genes.

3.2. - Results and discussion.

3.2.1. - Cloning and sequencing of the pCW7 repN gene.

Fragments of pCW7 were cloned into bacteriophage M13 to sequence the pCW7 repN gene using the dideoxynucleotide triphosphate method (Sanger *et al*, 1977). All rep genes studied have two HpaII sites at the origin of replication, within inverted complementary repeat sequence (ICR) II. The restriction map of pCW7 (Skinner, 1984) shows that pCW7 has *at least* one HpaII site at this position, as the two sites are too close together to be resolved by restriction analysis. There are no other HpaII sites in pCW7, however, there is a single HindIII site. Digestion of pCW7 with HpaII and HindIII gave two fragments visible on an agarose gel, of 1.98 and 2.42kb respectively. Both of these fragments were cloned into M13 /mp19 (Messing, 1983) cut with HindIII and AccI (Figure 3.3).

Ligation mixtures were transformed into *E. coli* JM101 and white plaques were picked. Single-stranded and double-stranded (replicative form) phage DNA was prepared. The double-stranded DNA was checked for inserts by digestion with HindIII and HpaII. Both of the expected insert fragments were obtained from this digest. Dideoxy sequencing using the M13 universal primer on the single-stranded templates of these two clones gave approximately 200bp of repN sequence on either side of the origin.

From this sequence, it was possible to design oligonucleotide primers to enable further sequencing to be carried out. The whole of pCW7 was cloned into M13 in both orientations (Figure 3.4) to give two universal templates, allowing sequencing to be carried out on both strands using synthetic

oligonucleotide primers (Brenner & Shaw, 1985). A primer was also designed from the pC221 *cat* gene promoter region, which was thought likely to be highly conserved between pC221 and pCW7, in order to sequence back into the 3' end of the *repN* gene. From the new sequences obtained, more oligonucleotide primers were designed, until the entire *repN* sequence could be obtained on both strands (Figure 3.5).

3.2.2. - Comparison of the pCW7 repN sequence with five other known rep sequences.

When the deduced REP N amino acid sequence was compared with those of five other REP proteins (Figure 3.6), it was found to have between 75 and 83% homology with them (Table 3.1). REP N appears to be most similar to REP C. All of the major differences between the REP N amino acid sequence and the others were situated, as expected, in two major regions, one at the 5' end of the coding sequence, corresponding to part of the origin sequence, and the other towards the 3' end of the coding sequence.

The six *rep* sequences also show approximately 80% homology at the DNA sequence level. The well defined regions of conservation and relative divergence (see figure 1.5) suggest the existence of distinct functional domains within the protein, with the C-terminal region probably involved in sequence-specific recognition at the divergent region of the origin, situated towards the 5' end of the *rep* coding sequence.

The six genes also show a high degree of structural conservation. All six origin sequences have three inverted complementary repeat sequences (ICR I,

II and III) which may form secondary structure in the form of hairpins. Were hairpin structures to form at all of the inverted repeat sequences, a "cloverleaf" structure might be formed (Figure 3.7). ICR II is highly conserved between the six plasmids, with just one base substitution in pS194 and pUB112. There are also some base changes in ICR I. However, the major changes between the six origin sequences occur within ICR III. From REP C footprinting data on the pT181 origin (Koepsel *et al*, 1986), it has been suggested that REP binds to the origin mainly over the region of ICR III, as well as the 3' side of ICR II (Figure 3.8). As the major differences between the origin sequences occur within ICR III, then it is plausible that this region is the specificity determinant for each origin. The other major divergent region, situated towards the 3' ends of the coding sequences, presumably encodes the region of the REP protein which is responsible for sequence-specifc recognition at ICR III.

The above proposal may, however, be an over-simplification. Although ICR III shows the greatest divergence between the six sequences, it is quite similar in pS194 and pCW7, with only one base difference. The 3' ends of the coding sequences of their rep genes are, however, quite different. It would, therefore, be interesting to test the specificity of the REP proteins of these plasmids, both *in vivo* and *in vitro*.

REP may bind to ICR III in the form of a hairpin, or, alternatively, in the linear duplex conformation. In the latter case, the dyad symmetry of the inverted repeat sequence may provide a binding site for REP as a dimer, which appears to be the native conformation of REP D protein (Thomas, 1988). This might be similar to λ repressor protein, λ cro protein, lac repressor protein and cAMP receptor protein (CAP), in which the two DNA contacting domains of these dimeric proteins bind two adjacent major groove regions of the DNA

centred in a region of dyad symmetry (reviewed by Pabo & Sauer, 1984). Binding of REP to the origin may or may not be accompanied by changes in secondary structure of the DNA, such as hairpin formation at ICR I, II or III.

ICR III of pT181, pC221 and pC223 differ from the other three in that the two halves of these inverted repeats are separated by a spacer of 4 or 5bp. ICR III of the other three plasmids has a spacer of only 1bp, which makes the formation of a hairpin at these inverted repeat sequences unlikely, since the formation of a hairpin with a loop of only one base is energetically unfavourable due to steric hindrance (Tinoco *et al*, 1971). Therfore, it seems likely that REP binds to ICR III in the linear duplex conformation. If so, and if the inverted repeat sequence acts as a binding site for the REP dimer, then the difference in size of the spacers may reflect differences in the geometry of binding of different REP proteins on these different origins, as the difference in size of the spacer will change the position of the inverted repeat sequence on the DNA helix. The consequence would presumably be a major displacement of the binding of the REP dimers at the origins (Figure 3.9).

The upstream copy control regions (cop) of the six rep genes studied also show a high degree of homology. This homology breaks down at approximately -240bp upstream of the rep translational start site. The pCW7 cop region has all the essential regulatory features present. The repN mRNA leader sequence has putative secondary structure in the form of inverted complementary repeat sequences which may form hairpin structures (Figure 3.10), along with the approximately 100 base complementary RNA countertranscript, which form the basis of the copy control mechanism.

Slight differences within the countertranscript region of the cop DNA sequence may alter the incompatibility group into which the particular plasmid falls, as the RNA countertranscript can only act in a regulatory manner on the rep mRNA derived from the same plasmid or from plasmids of the same incompatibility group. The cop region of pCW7 appears to be sufficiently different from the others within the countertranscript region to suggest that it belongs to a new incompatibilty group. However, incompatibility tests between pCW7 and the other known plasmids of the pT181 "family" *in vivo* would have to be carried out to confirm this.

Table 3.1: Percentage homologies between REP amino acid sequences.

The percentage homologies between known REP amino acid sequences are shown in the table. The six REP proteins studied show between approximately 73 and 85% homology at the amino acid sequence level.

	C	D	E	I	J	N
repC	100	80.3	73.8	77	78.3	82.8
repD	80	100	82.6	84.2	78.4	82.3
repE	73.8	82.8	100	81.5	73.6	80
repI	77.2	84.5	81.6	100	76.1	78.5
repJ	78.3	78.7	73.6	76.3	100	78.4
repN	82.8	82.5	80	78.3	74.8	100

Fig 3.1: Restriction map of plasmid pCW7.

The map shows the major restriction sites of pCW7 as determined by restriction endonuclease mapping (Skinner, 1984).

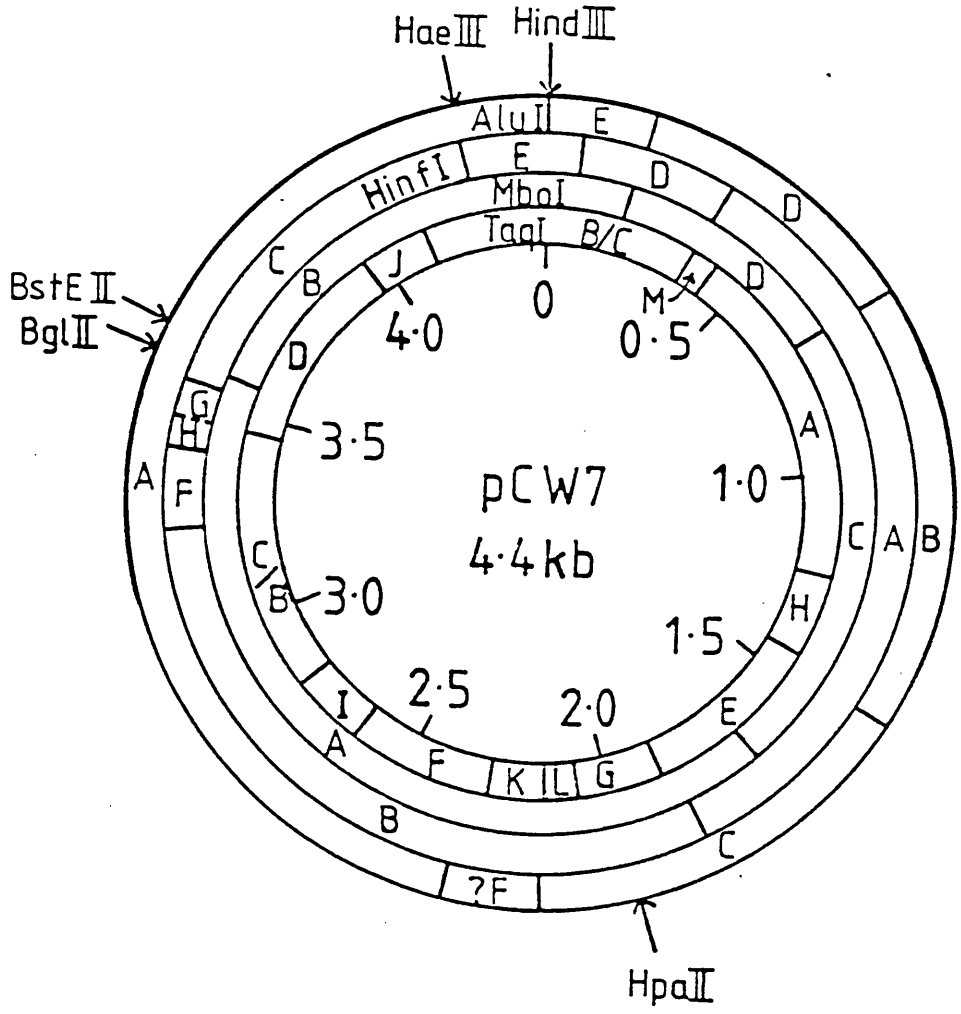
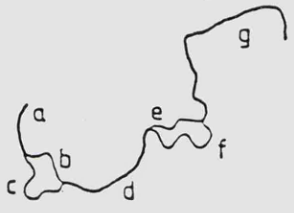

HaeIII
HindIII
AluI
E
HinfI
E
D
MboI
TaqI
B/C
D
BstEII
BglII
C
B
J
0
M
4·0
0·5
D
G
H
3·5
A
F
pCW7
4·4kb
1·0
C
A
B
C
B
3·0
H
1·5
I
2·5
2·0
E
A
F
K
IL
G
B
C
?F
HpaII

Fig 3.2: Heteroduplex mapping between plasmids pC221 and pCW7.

The diagram shows the results of heteroduplex mapping between pC221 and pCW7 (Skinner, 1984). Both plasmids were cloned into pBR322 via the unique HindIII for these studies. The cloned plasmids were linearised with PstI and heteroduplexes were formed which were then studied by electron microscopy, as shown in diagram (a). Regions a and g correspond to the pBR322 sequences. regions b/c and e/f correspond to regions of non-homology between pC221 and pCW7. Region d corresponds to the region of homology between pC221 and pCW7, which extends over the region corresponding to the pC221 cat gene, as expected, as well as the repD gene. In approximately 39% of cases, the heteroduplex map has two additional single-stranded "bubbles", d_2 and d_4, within this region, indicating regions where this homology breaks down. One of these is at the 5' end of the rep coding sequence, at approximately the region corresponding to the origin. The other is situated towards the 3' end of the rep gene. These two regions are in good agreement with the areas of non-conserved DNA sequence between the five known rep sequences.

(a)

Fig. 1



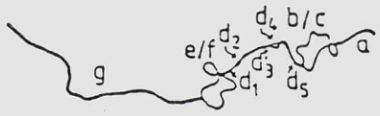


Fig. 2

Fig. 1

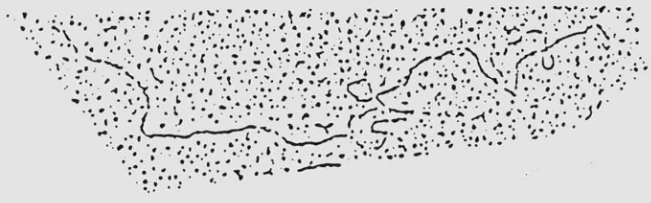


Fig. 2

(b)

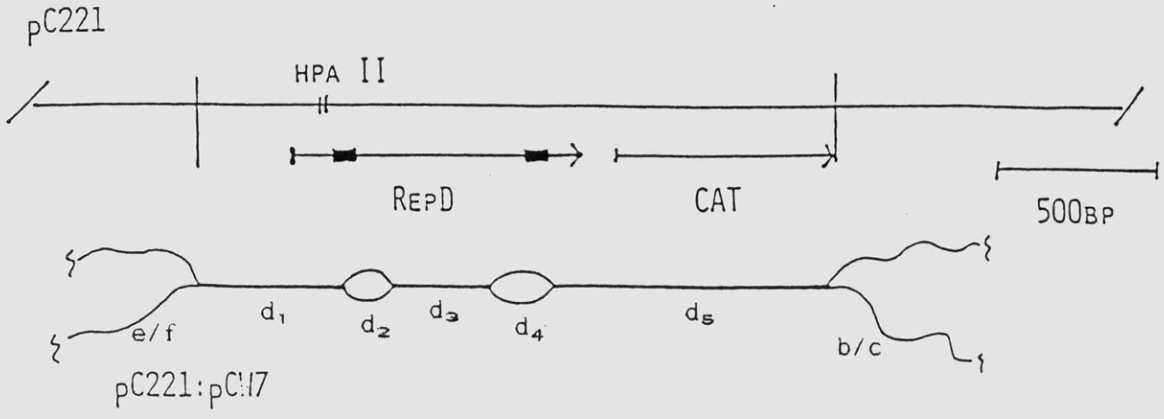


Fig 3.3: Cloning of two HpaII /HindIII fragments into M13 /mp19.

The diagram shows the route to the cloning of two sticky ended HpaII /HindIII fragments of pCW7 into M13 /mp19 digested with AccI and HindIII. The pCW7 rep and cop regions are shown. The position of the M13 universal primer is marked by the arrow (U). Dideoxy sequencing on both of these templates using the M13 universal primer gave pCW7 sequence information approximately 200bp either side of the origin.

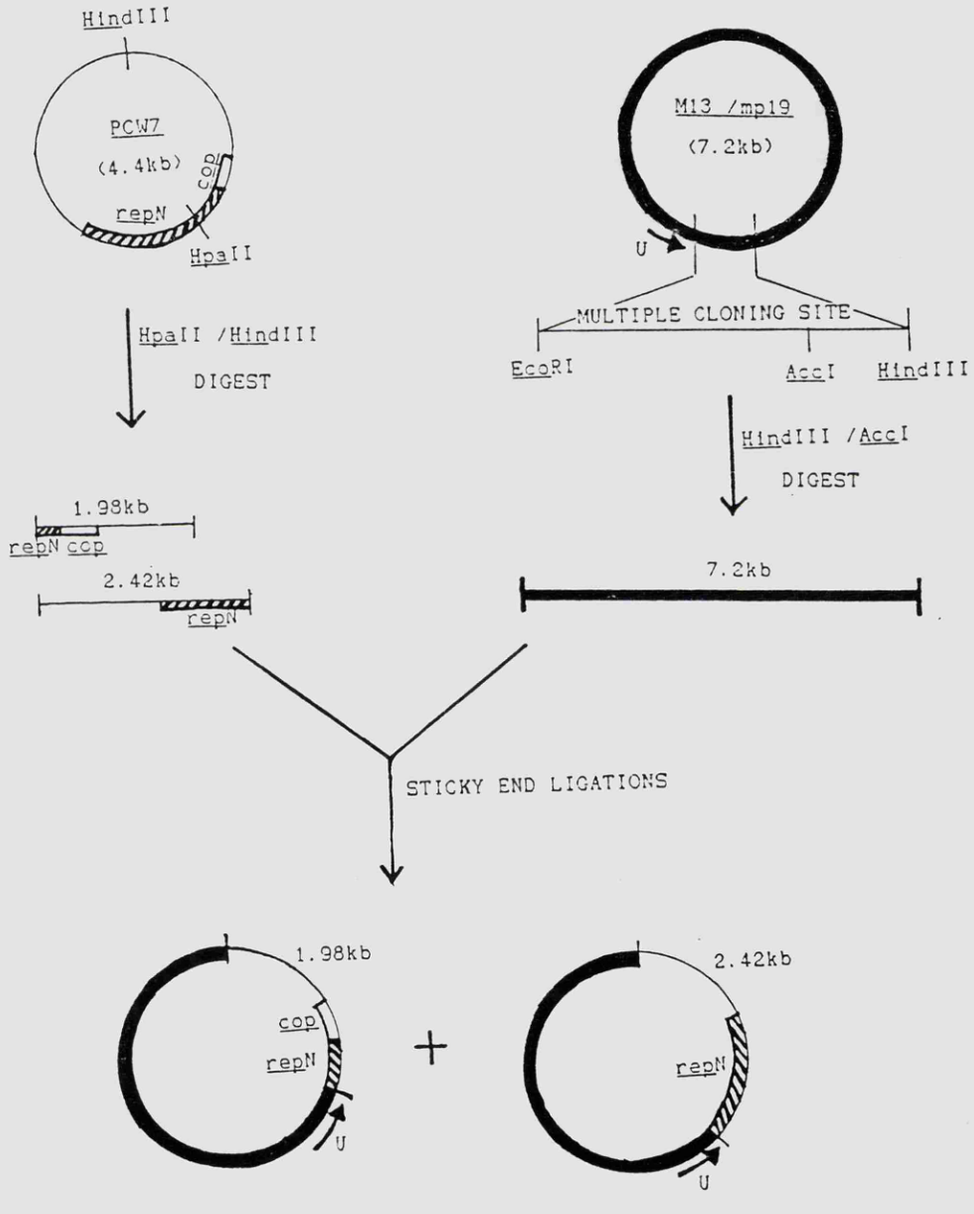
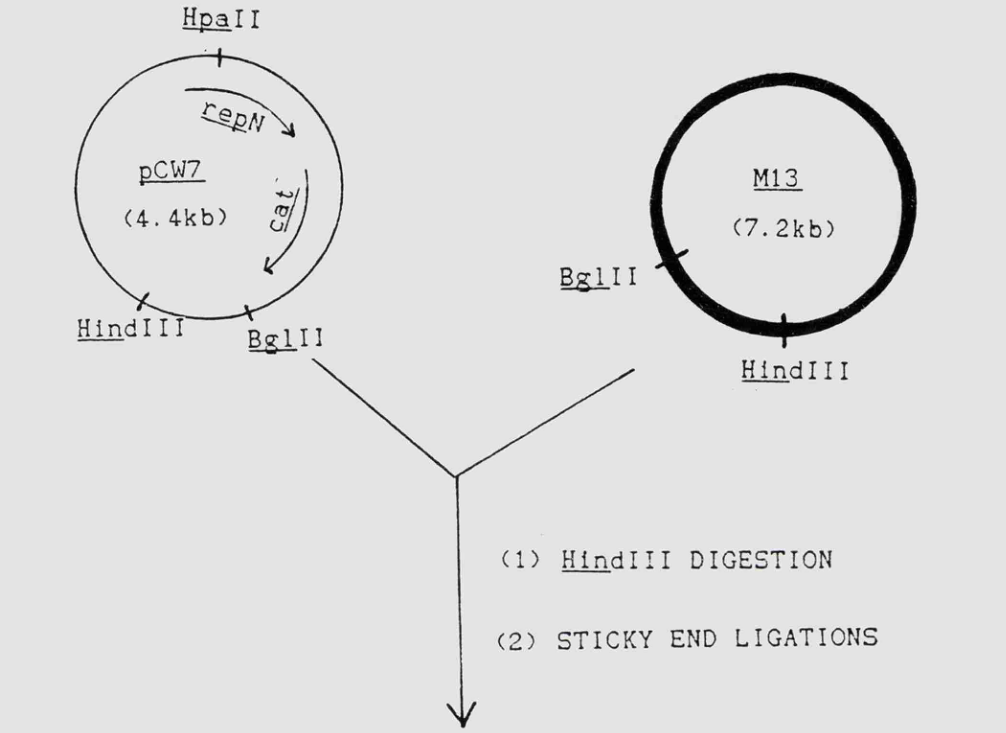
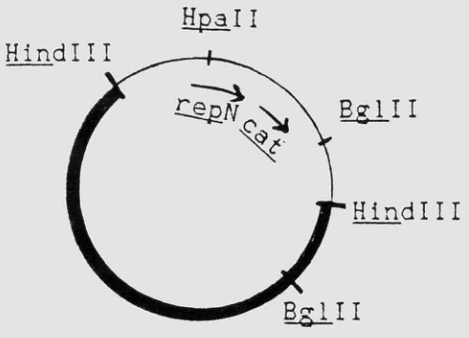

HindIII
PCW7
(4.4kb)
cop
repN
HpaII
HpaII /HindIII
DIGEST
M13 /mp19
(7.2kb)
U
MULTIPLE CLONING SITE
EcoRI
AccI
HindIII
HindIII /AccI
DIGEST
1.98kb
repN cop
2.42kb
repN
7.2kb
STICKY END LIGATIONS
1.98kb
cop
repN
U
+
2.42kb
repN
U

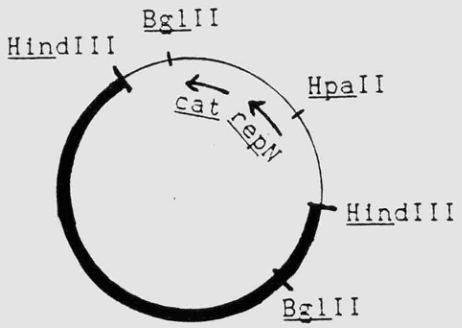
Fig 3.4: Cloning of the whole plasmid pCW7 into M13 /mp19.

The whole of plasmid pCW7 was cloned into M13, in both orientations, via the HindIII site. The orientation of the inserts were checked by BglII digestion of the double-stranded replicative forms. Single-stranded DNA was prepared from each clone to give two universal templates to allow sequencing of the pCW7 repN gene on both strands using synthetic oligonucleotide primers.





ORIENTATION 1



ORIENTATION 2

Fig 3.5: The complete nucleotide sequence of the pCW7 repN gene.

The complete nucleotide sequence of the repN gene, along with some 5' and 3' flanking DNA, is shown on both strands, as determined by dideoxy sequencing methods (Sanger *et al*, 1977). The REP N amino acid sequence is written above the nucleotide sequence.

```
5'--TACCGACATTTTAATTCCGTTGCCGATAAAATTTAATTTCCTTGATATGTACTTATATTATAGATAATTT
3'--ATGGCTGAAAAATTAAGGCTTCGGCTATTTTAAATTAAAGGTTGATATGATGAATATAATATCTATTAAA

    TAATTTTCATTATGTAATAAGTTTTGGTACTAAAAAACACCTGATCTAATTTAGACCAAGTCTTTTGATA
    ATTAAAAGTAATACATTATTCAAAACCATGATTTTTTGTGGACTAGATTAAATCTGGTTCAGAAAACTAT

    GTGTTATATTAATAAACAAATAAAAAGAAGTCGCACGCCCTACCAAAGTTTTGTGAATGCGACTTACCTA
    CACAATATAATTATTTGTTTATTTTTCTTGAGCGTGCGGGATGGAAACAAAACACTTACGCTGAATGGAT

    AATAAAATATAAAGAGTTATGTTTTATTAATTTTTGTATATCTAGATATTAAACGATATAGGTTTTATTC
    TTATTTTATATTTCCTAATACAAAATAATTAAAAACATATAGATCTATAATTTGCTATATCCAAAATAAG

                                                                REP START
                                                                MetSerLysAsn
    TTCAAGATATATATTTGGGTGAGCGACTTCTTAAAGTAAATTAAGGAGTCGATTTTTTATGAGTAAAAAT
    AAGTTCTATATATAAACCCACTCGCTGAAGAATTTCATTTAATTCCTCAGCTAAAAAATACTCATTTTTA

    AsnHisAlaAsnHisSerAsnHisLeuGluAsnHisAspLeuAspAsnPheSerLysThrGlyTyrSerA
    AATCACGCAAATCATTCAAATCATTTAGAAAATCACGATTTAGACAATTTTTCTAAAACCGGCTACTCTA
    TTAGTGCGTTTAGTAAGTTTAGTAAATCTTTTAGTGCTAAATCTGTTAAAAAGATTTTGGCCGATGAGAT

    snSerArgLeuAsnArgHisThrMetTyrThrProGluProLysLeuSerPheAspAlaMetThrIleVa
    ATAGCCGGTTAAACCGACATACTATGTACACCCCGGAACCAAAATTAAGTTTTGACGCTATGACTATTGT
    TATCGGCCAATTTGGCAGTATGATACATGAGGGGCCTTGGTTTTAATTCAAAACTGCGATACTGATAACA

    lGlyAsnLeuAsnLysAsnAsnAlaHisLysLeuSerGluPheMetSerValGluProGlnIleArgLeu
    TGGAAATCTTAATAAAAATAATGCTCACAAACTATCTGAATTTATGAGTGTCGAGCCACAAATTCGACTT
    ACCTTTAGAATTATTTTTATTACGAGTGTTTGATAGACTTAAATACTCACAGCTCGGTGTTTAAGCTGAA

    TrpAspIleLeuGlnThrLysPheLysAlaLysAlaLeuGlnLysLysValTyrIleGluTyrAspLysV
    TGGGATATACTACAAACTAAATTTAAAGCTAAAGCTCTACAGAAAAAGGTTTATATCGAATATGACAAAG
    ACCCTATATGATGTTTGATTTAAATTTCGATTTCGAGATGTCTTTTTCCAAATATAGCTTATACTGTTTC

    alLysAlaArgTyrValGlyArgArgAsnMetArgValGluPheAsnProAsnLysLeuThrHisGluGl
    TAAAAGCCAGATACGTGGGAAGACGTAATATGCGTGTTGAATTTAATCCAAATAAACTTACGCATGAAGA
    ATTTTCGGTCTATGCACCCTTCTGCATTATACGCACAACTTAAATTAGGTTTATTTGAATGCGTACTTCT

    uMetLeuTrpLeuLysGlnAsnIleIleAspTyrMetGluAspAspGlyPheThrArgLeuAspLeuAla
    AATGCTTTGGTTAAAACAAAACATTATCGACTACATGGAAGACGATGGTTTTACAAGATTAGATTTAGCT
    TTACGAAACCAATTTTGTTTTGTAATAGCTGATGTACCTTCTGCTACCAAAATGTTCTAATCTAAATCGA
```

PheAspPheGluTyrAspLeuSerAspTyrTyrAlaMetThrAspLysSerValLysLysThrIlePheT
TTTGATTTTGAATATGATTTAAGTGATTATTATGCAATGACTGATAAATCAGTTAAGAAAACTATTTTTT
AAACTAAAACTTATACTAAATTCACTAATAATACGTTACTGTCTATTTAGACAATTCTTTTGATAAAAAA

yrGlyArgAsnGlyLysProGluThrLysTyrPheGlyValArgAspSerAspArgPheIleArgIleTy
ATGGTCGTAACGGTAAGCCAGAAACGAAATATTTTGGTGTTCGTGACAGTGATAGATTTATTAGAATTTA
TACCAGCATTGCCATTCGGTCTTTGCTTTATAAAACCACAAGCACTGTCACTATCTAAATAATCTTAAAT

rAsnLysLysGlnGluArgLysAspAsnAlaAspIleLysIleMetSerGluHisLeuTrpArgValGlu
TAATAAAAAACAGGAACGCAAAGATAATGCAGATATTAAAATTATGTCTGAACACTTATGGCGTGTAGAA
ATTATTTTTTGTCCTTGCGTTTCTATTACGTCTATAATTTTAATACAGACTTGTGAATACCGCACATCTT

IleGluLeuLysArgAspMetValAspTyrTrpAsnAspCysPheAsnAspLeuHisIleLeuGlnProA
ATTGAATTAAAAAGAGATATGGTTGATTATTGGAACGATTGTTTTAATGATTTACATATATTACAACCAG
TAACTTAATTTTTCTCTATACCAACTAATAACCTTGCTAACAAAATTACTAAATGTATATAATGTTGGTC

spTrpLysThrIleGluArgThrSerAspArgAlaMetValPheMetLeuLeuAsnAspGluGluGluTr
ATTGGAAAACTATCGAACGTACTTCTGATAGAGCAATGGTTTTTATGTTGTTGAATGATGAAGAAGAATG
TAACCTTTTGATAGCTTGCATGAAGACTATCTCGTTACCAAAAATACAACAACTTACTACTTCTTCTTAC

pGlyLysLeuGluArgArgThrLysAsnLysTyrLysLysLeuIleLysGluIleSerLeuIleAspLeu
GGGAAAATTAGAAAGACGTACTAAGAATAAATATAAAAAATTAATTAAAGAAATATCTCTAATTGATTTA
CCCTTTTAATCTTTCTGCATGATTCTTATTTATATTTTTTAATTAATTTCTTTATAGAGATTAACTAAAT

ThrAspLeuMetLysSerThrLeuLysAlaAsnGluLysGlnLeuGlnLysGlnIleAspPheTrpGlnA
ACTGATTTAATGAAATCGACTTTAAAAGCGAACGAAAAACAATTGCAAAAGCAGATTGATTTTTGGCAAC
TGACTAAATTACTTTACGTGAAATTTTCGCTTGCTTTTTGTTAACGTTTTCGTCTAACTAAAAACCGTTG

REP END
rgGluPheArgPheTrpLys
GTGAATTTAGATTTTGGAAGTAAAGTTAACTTTGTTTTATACAAAATCTTACTTTAGTATAATTAACGTT
CACTTAAATCTAAAACCTTCATTTCAATTGAAACAAAATATGTTTTAGAATGAAATCATATTAATTGCAA

TAAGAGGTGACACAATATATGAAAAAATCAGAGGATTATTCCTCCTAAATACAAAGAGATAAGATGAAAT
ATTCTCCACTGTGTTATATACTTTTTTAGTCTCCTAATAAGGATTATTTATGTTTCTCTAATCTACTTTA

CAT START
TTAGGAGGAAATTATATATG--3'
AATCCACCTTTAATATATAC--5'

Fig 3.6: Comparison of REP amino acid sequences.

The pT181 REP C, pS194 REP E, pUB112 REP I, pC223 REP J and pCW7 REP N amino acid sequences are shown, compared against the pC221 REP D amino acid sequence. The sequences are identical except where indicated. The four major regions of divergence are indicated as i, ii, iii and iv. Region ii corresponds to the specificity determinant of the origin. Region iv corresponds to the C-terminal recognition domain of the protein. Tyrosine 188, which has been implicated in the formation of a covalent attachment to the 5' phosphate at the nick site, is indicated by a star. This tyrosine is situated within a run of nine amino acids which are completely conserved in all six sequences.

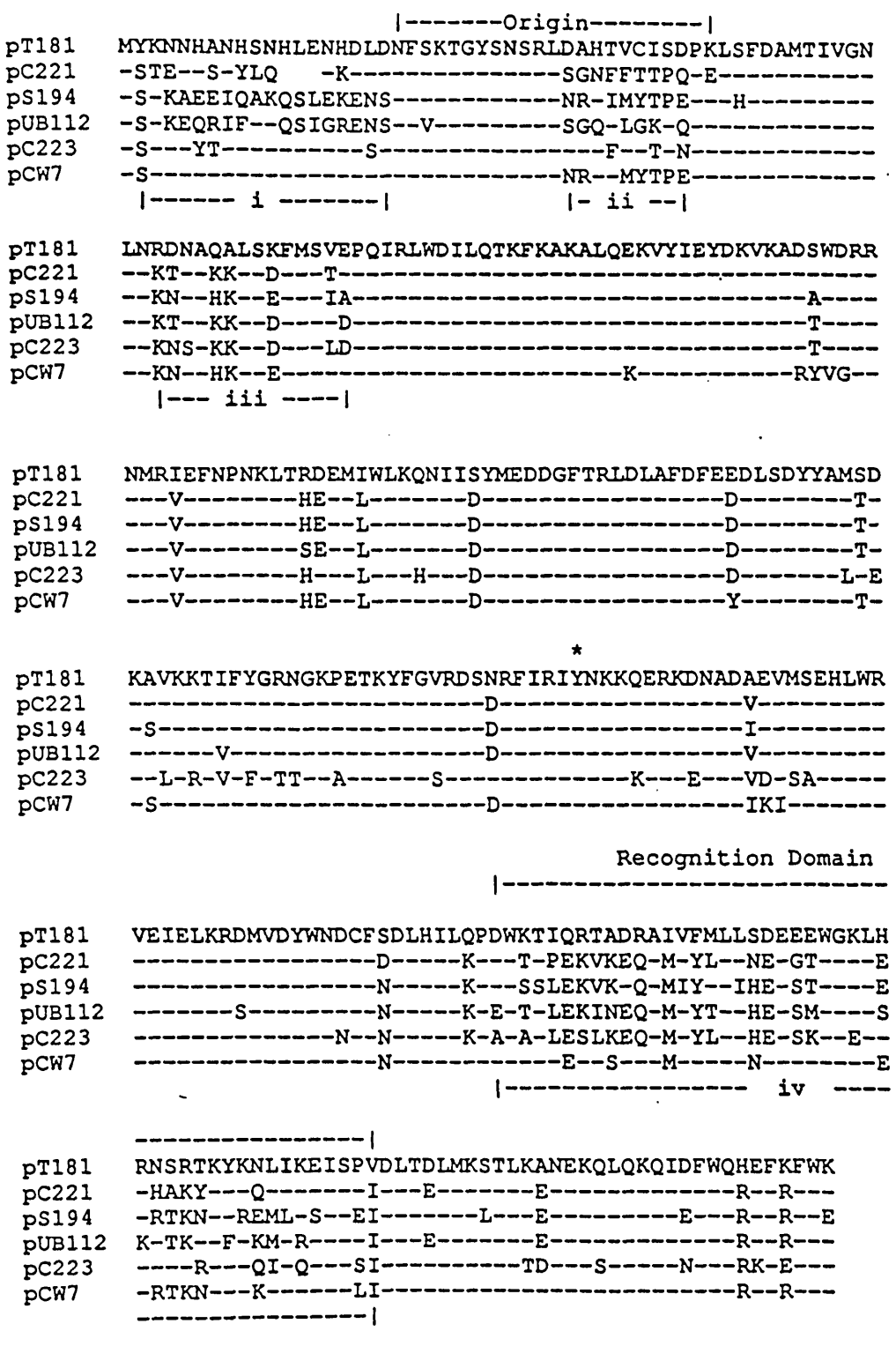
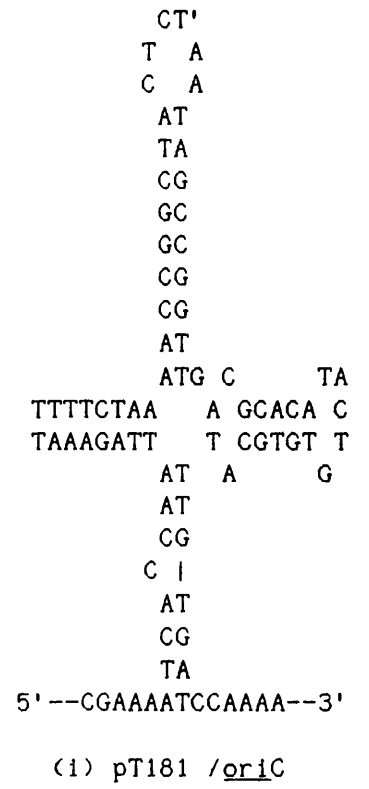
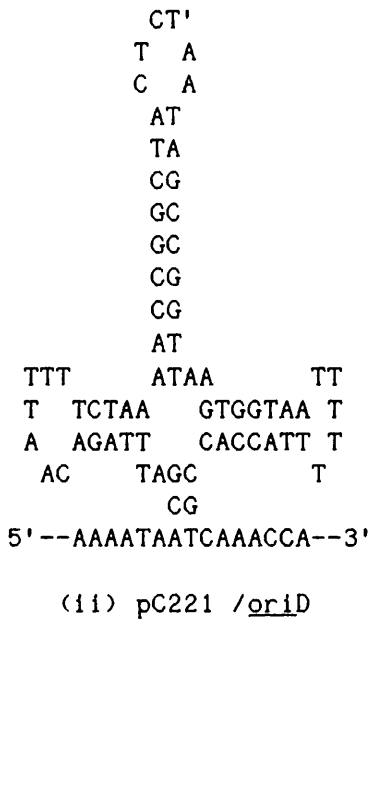

|-------Origin--------|
pT181 MYKNNHANHSNHLENHDLDNFSKTGYSNSRLDAHTVCISDPKLSFDAMTIVGN
pC221 -STE--S-YLQ -K----------SGNFFTTPQ-E----------
pS194 -S-KAEEIQAKQSLEKENS----------NR-IMYTPE---H---------
pUB112 -S-KEQRIF--QSIGRENS--V---------SGQ-LGK-Q----------
pC223 -S---YT----------S----------F--T-N----------
pCW7 -S----------NR--MYTPE----------
|------ i -------| |- ii --|
pT181 LNRDNAQALSKFMSVEPQIRLWDILQTKFKAKALQEKVYIEYDKVKADSWDRR
pC221 --KT--KK--D---T----------
pS194 --KN--HK--E---IA----------A----
pUB112 --KT--KK--D----D----------T----
pC223 --KNS-KK--D---LD----------T----
pCW7 --KN--HK--E----------K----------RYVG--
|--- iii ----|
pT181 NMRIEFNPNKLTRDEMIWLKQNIISYMEDDGFTRLDLAFDFEEDLSDYYAMSD
pC221 ---V--------HE--L-------D----------D--------T-
pS194 ---V--------HE--L-------D----------D--------T-
pUB112 ---V--------SE--L-------D----------D--------T-
pC223 ---V--------H---L---H---D----------D-------L-E
pCW7 ---V--------HE--L-------D----------Y--------T-
*
pT181 KAVKKTIFYGRNGKPETKYFGVRDSNRFIRIYNKKQERKDNADAEVMSEHLWR
pC221 ----------D----------V---------
pS194 -S----------D----------I---------
pUB112 ------V----------D----------V---------
pC223 --L-R-V-F-TT--A------S----------K---E---VD-SA-----
pCW7 -S----------D----------IKI-------
Recognition Domain
|----------
pT181 VEIELKRDMVDYWNDCFSDLHILQPDWKTIQRTADRAIVFMLLSDEEEWGKLH
pC221 ----------D-----K---T-PEKVKEQ-M-YL--NE-GT----E
pS194 ----------N-----K---SSLEKVK-Q-MIY--IHE-ST----E
pUB112 -------S----------N-----K-E-T-LEKINEQ-M-YT--HE-SM----S
pC223 ----------N--N-----K-A-A-LESLKEQ-M-YL--HE-SK--E--
pCW7 ----------N----------E--S---M-----N--------E
|---------- iv ----
----------|
pT181 RNSRTKYKNLIKEISPVDLTDLMKSTLKANEKQLQKQIDFWQHEFKFWK
pC221 -HAKY---Q-------I---E-------E----------R--R---
pS194 -RTKN--REML-S--EI-------L---E---------E---R--R--E
pUB112 K-TK--F-KM-R----I---E-------E----------R--R---
pC223 ----R---QI-Q---SI----------TD---S-----N---RK-E---
pCW7 -RTKN---K------LI----------R--R---
----------|

Fig 3.7: Putative secondary structure at the origins of the pT181 "family" of plasmids.

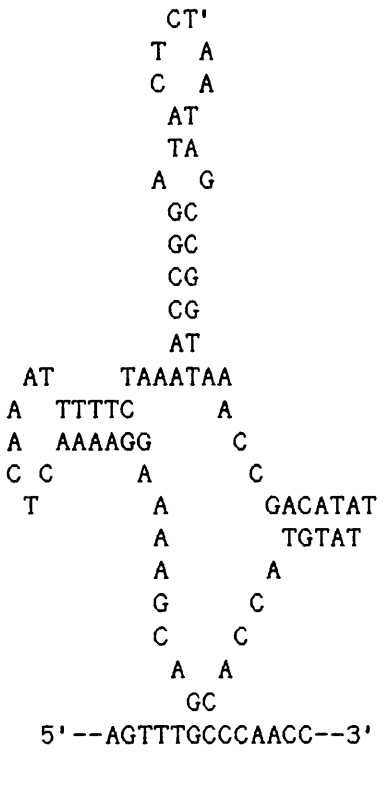
The diagram shows the three inverted complementary repeat sequences (ICR I, II and III) at the ori sequences of plasmids pT181, pC221, pS194, pUB112, pC223 and pCW7. These inverted repeat sequences may form secondary structure in the form of a cloverleaf, as shown. It is not, however, known whether such structures would be stable *in vivo*. Hairpin formation at ICR I, II or III may, though, be important in the replication of these plasmids, especially in the presence of possible stabilising factors, such as negative supercoiling of the DNA or bound proteins.



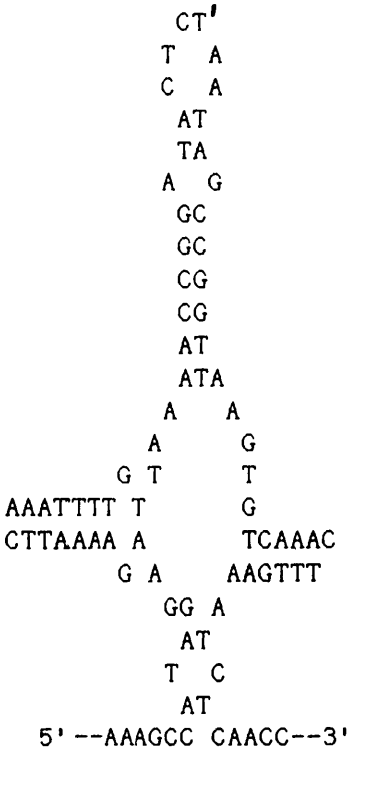
(i) pT181 /oriC



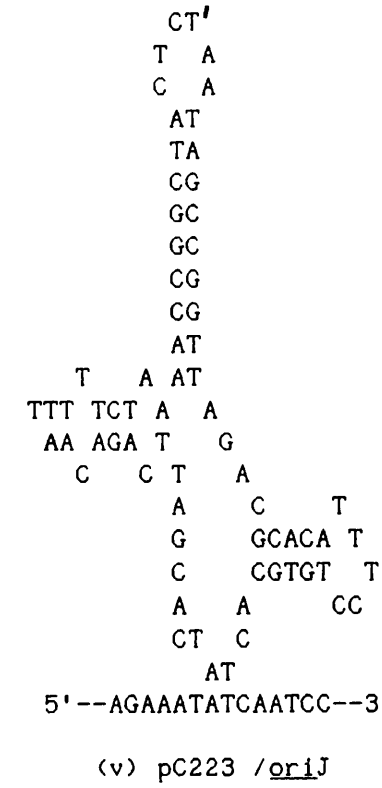
(ii) pC221 /oriD



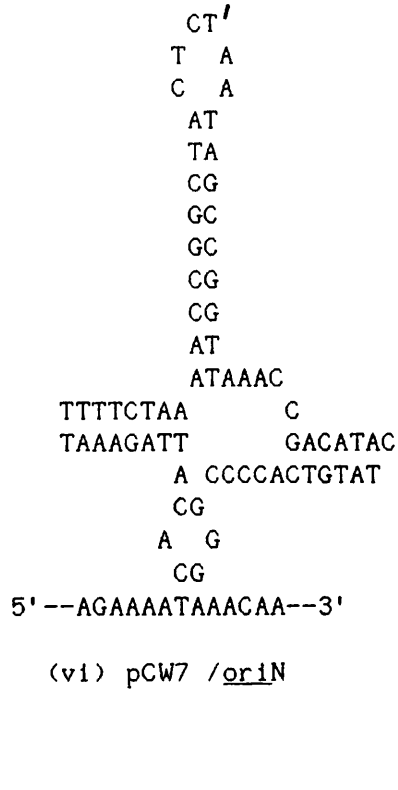
(iii) pS194 /oriE



(iv) pUB112 /oriI



(v) pC223 /oriJ



(vi) pCW7 /oriN

Fig 3.8: The origin of pT181 (oriC) and the footprint of REP C .

The diagram shows the position of binding of REP C on oriC, as determined by various DNA footprinting techniques (Koepsel et al, 1986). ICR II and III are denoted by dotted arrows.

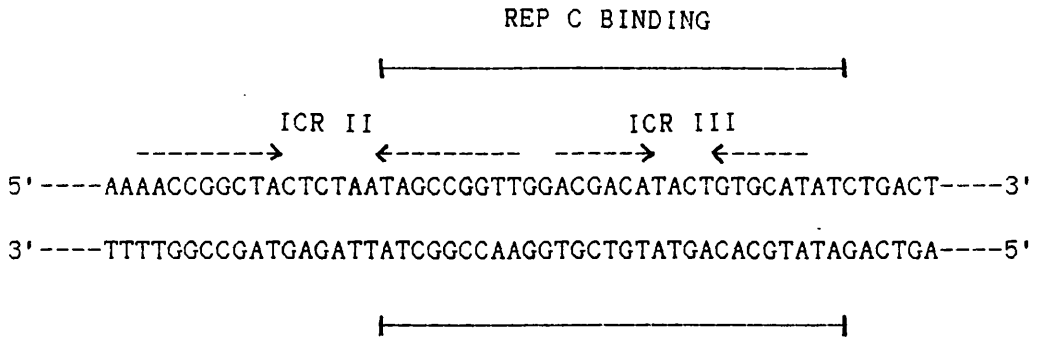
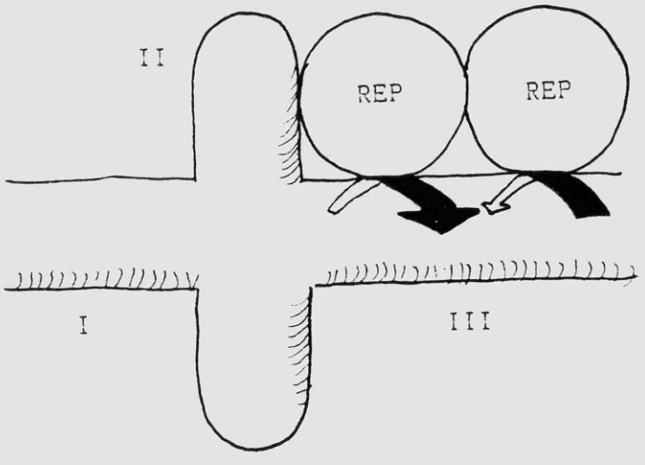

REP C BINDING
ICR II
ICR III
5'----AAAACCGGCTACTCTAATAGCCGGTTGGACGACATACTGTGCATATCTGACT----3'
3'----TTTTGGCCGATGAGATTATCGGCCAAGGTGCTGTATGACACGTATAGACTGA----5'

Fig 3.9: Proposed binding of a REP dimer at ICR III.

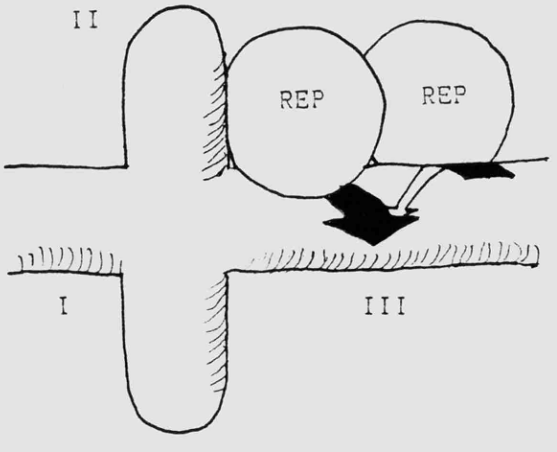
Inverted complementary repeat sequence, ICR III, is proposed as the binding site for REP in the form of a dimer, with the possibility of hairpin formation at the adjacent ICR II. ICR III of pT181, pC221 and pC223 has an inverted repeat separated by a spacer of 4 or 5 bases, whereas ICR III of the other three plasmids has a spacer of only one base. The difference in size of the spacer will change the position of the inverted repeat sequence, ICR III, on the DNA helix, presumably causing displacement of the binding of the REP dimer at these different origin sequences,

COMPARISON OF IOR III REGIONS OF SIX DIFFERENT PLASMIDS.

pT181	GGACGCACATACTGTGTGCATAT
pC221	AAGTGGTAATTTTTTTACCACCC
pC223	AGACGCACATTTCGTGTGCACTT
pCW7	AAACCGACATACTATGTACACCC
pS194	AAACCGACATATTATGTACACCC
pUB112	AAGTGGTCAAACTTTGGGAAAAT



(a) PROPOSED REP BINDING MODEL FOR pT181, pC221 & pC223.

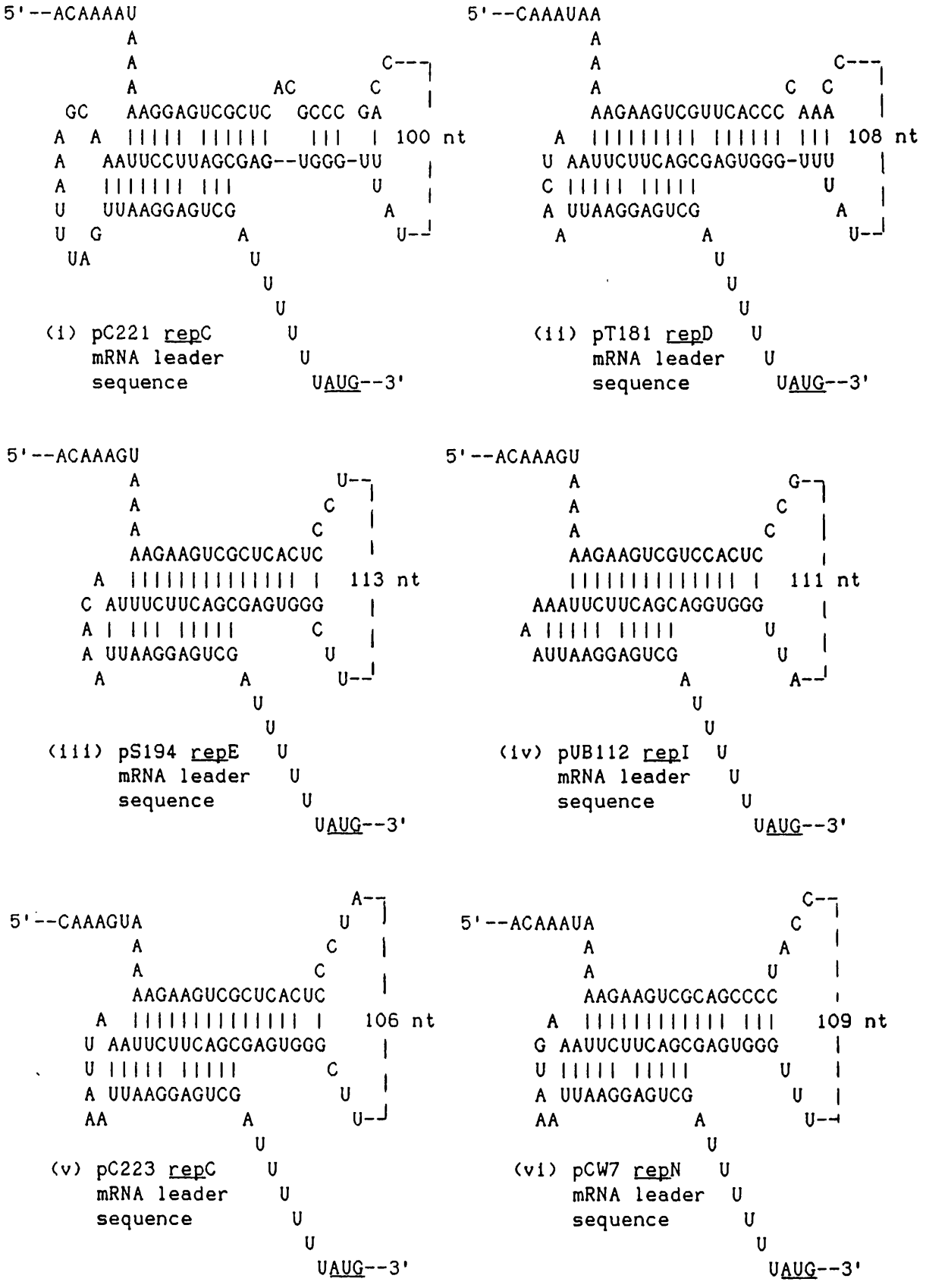


(b) PROPOSED REP BINDING MODEL pCW7, pS194 & pUB112.

Fig 3.10: Putative secondary structure in the rep mRNA leader sequences of the pT181 "family" of plasmids.

The diagram shows proposed secondary structure within the leader sequences of repC, D, E, I, J and N mRNAs. The ribosome binding site (SD) is buried within an inverted complementary repeat sequence, which may form a hairpin structure, thus sequestering the ribosome binding site and inhibiting rep mRNA translation. There is also some evidence for an attenuation control mechanism involving this proposed hairpin, which resembles rho-dependent termination structures seen in *E. coli*.

Binding of the complementary RNA countertranscript to the rep mRNA favours the extrusion of this regulatory hairpin by preventing the formation of a larger, overlapping hairpin, called the preemptor. There are differences within the rep mRNA leader sequences of these plasmids, leading to changes in incompatibility group, as a particular RNA countertranscript will only be complementary to the rep mRNA leader sequence derived from the same plasmid, or from plasmids of the same incompatibilty group.


5'--ACAAAAU
GC AAGGAGUCGCUC GCCC GA
A A 100 nt
A AAUUCCUUAGCGAG--UGGG-UU
U UUAAGGAGUCG
U G
UA
(i) pC221 repC
mRNA leader
sequence
UAUG--3'
5'--CAAAUAA
AAGAAGUCGUUCACCC AAA
A 108 nt
U AAUUCUUCAGCGAGUGGG-UUU
C
A UUAAGGAGUCG
A
(ii) pT181 repD
mRNA leader
sequence
UAUG--3'
5'--ACAAAGU
AAGAAGUCGCUCACUC
A 113 nt
C AUUUCUUCAGCGAGUGGG
A UUAAGGAGUCG
(iii) pS194 repE
mRNA leader
sequence
UAUG--3'
5'--ACAAAGU
AAGAAGUCGUCCACUC
111 nt
AAAUUCUUCAGCAGGUGGG
AUUAAGGAGUCG
(iv) pUB112 repI
mRNA leader
sequence
UAUG--3'
5'--CAAAGUA
AAGAAGUCGCUCACUC
A 106 nt
U AAUUCUUCAGCGAGUGGG
A UUAAGGAGUCG
AA
(v) pC223 repC
mRNA leader
sequence
UAUG--3'
5'--ACAAAUA
AAGAAGUCGCAGCCCC
A 109 nt
G AAUUCUUCAGCGAGUGGG
U
A UUAAGGAGUCG
AA
(vi) pCW7 repN
mRNA leader
sequence
UAUG--3'

CHAPTER 4.

IN VITRO REPLICATION OF pC221 AND RELATED PLASMIDS USING *S. AUREUS* CELL FREE EXTRACTS.

4.1. - Introduction.

4.2. - Results and discussion.

4.2.1. - Preparation of cell free extracts of *S. aureus*.

4.2.2. - *In vitro* replication of pC221.

4.2.3. - The use of different plasmid substrates in the *in vitro* replication system.

4.2.4. - The specificity of REP D appears to increase with substrate plasmids of low superhelical density.

CHAPTER 4.

IN VITRO REPLICATION OF pC221 AND RELATED PLASMIDS USING S. AUREUS CELL FREE EXTRACTS.

4.1: - Introduction.

There have been many examples of *in vitro* plasmid replication using cell free extracts and added nucleotide triphosphates, eg plasmids R1 (Diaz *et al*, 1981; Diaz *et al*, 1982) and Col E1 (Sakikabara *et al*, 1974) from *E. coli* and pT181 (Khan *et al*, 1981) from *S. aureus*. The following chapter describes an *in vitro* study of the initiation of replication of pC221 and related plasmids using cell free extracts prepared from *S. aureus*.

4.2: - Results and discussion.

4.2.1. - Preparation of cell free extracts of *S. aureus*.

The 70% saturation ammonium sulphate precipitate of *S. aureus* cleared whole cell lysates were resuspended in HEPES buffer (pH 8) containing 100mM KCl, 1mM DTT and 10% ethylene glycol. Since the presence of nuclease can give rise to spurious DNA repair replication in the *in vitro* replication system, extracts were prepared using a strain of *S. aureus* which is deficient in nuclease (RN1786).

4.2.2: - *In vitro* replication of pC221.

The cell free extracts were assayed by their ability to incorporate radioactive deoxyribonucleotide triphosphates (dNTP's) into acid precipitable material in the presence of added pC221 DNA (supercoiled), appropriate buffers, nucleotide triphosphates, K^+ and Mg^{++} ions. Assays were carried out at 30°c, typically for 1 hour, and acid precipitable radioactivity was determined by liquid scintillation counting.

Initially, extracts were made from both plasmid-free RN1786 and RN1786 containing pC221. The extracts prepared from cells containing pC221 gave greater incorporation of radioactivity into acid precipitable material (Table 4.1), suggesting that a plasmid encoded product is required for pC221 replication *in vitro*, presumably REP D.

However, the increase in incorporation obtained was quite small (approximately 9 fold). A further increase was observed on addition of REP D (Figure 4.1), which has been purified by overproduction in *E. coli* (Thomas, 1988), to the assay mixtures. Increasing concentrations of REP D gave greater levels of incorporation, showing that REP D is required for pC221 replication *in vitro*.

A time course experiment was carried out over 2½ hours in the presence and absence of added REP D (Figure 4.2). Incorporation of radioactivity into acid precipitable material increased with time and did not appear to level off over the duration of the time course.

The above experiment, however, showed a high level of background incorporation in the absence of REP D, probably due to DNA repair synthesis of nicked DNA present, both plasmid and chromosomal. It was noted that the level of background incorporation in the absence of REP D appeared to a certain extent to be dependent upon the state of the substrate plasmid, as determined by agarose gel electrophoresis. A higher proportion of nicked plasmid DNA generally gave a higher level of background incorporation.

Most of this background incorporation, however, appeared to be complete after 30 minutes. Preincubation of the assay mixtures at 30°c for 30 minutes before addition of the radioactive label was, therefore, effective in cutting down much of this background incorporation (Figure 4.3).

An experiment was carried out to label pC221 with $\alpha[^{32}P]$ dATP using the *in vitro* replication system. The products of the reaction were subjected to restriction endonuclease digestion and agarose gel electrophoresis. Autoradiography of the resulting gel gave an insight into the nature of

labelling of the plasmid (Figure 4.4). The undigested products showed that most of the label had been incorporated into the relaxed forms of the DNA, with only a small amount in the supercoiled form.

A HinfI digest of the labelled products gave several fragments, the largest of which was the 976bp fragment known to contain the pC221 origin (oriD). Therefore, this fragment was expected to be labelled first. The autoradiograph shows that this fragment was predominantly labelled, suggesting that replication is initiating within the region containing oriD.

A HpaII /HindIII double digest of the labelled material gave four fragments. Assuming that replication is initiating from the usual nick site, and, as HpaII cuts very close to, and on either side of the nick site, then the predominantly labelled fragment indicates the direction of replication from the nick site. The results of this experiment suggest that replication of pC221 *in vitro* is unidirectional and in the direction expected for elongation at the free 3' OH group generated at the nick site.

However, replication does not appear to be continuing to give a full round, as only the HinfI fragment containing oriD was appreciably labelled. This may be due to halting of replication, e.g. at some kind of secondary structure which is more favourable *in vitro* than *in vivo*. Alternatively, it may be that one of the substrates or enzymes required for the elongation reaction is limiting. However, since the initiation event itself, rather than elongation, is of principal concern here, the matter of incomplete incorporation was not pursued further.

The *in vitro* replication system may be used to assay for REP D activity, for example, as an aid in the purification of REP D, as an assay for

REP D activity in crude extracts of *E. coli* which over-express the repD gene (Figure 4.5), or to assay fractions collected from a heparin sepharose column (Figure 4.6).

The *in vitro* replication system was also used to investigate the effect of rifampicin, an inhibitor of RNA synthesis, and novobiocin, an inhibitor of DNA gyrase, on pC221 replication (Table 4.2). Rifampicin was seen to have little effect on incorporation, suggesting that rifampicin-sensitive RNA synthesis is not required for the initiation of pC221 replication *in vitro*. This is consistent with a rolling circle replication mechanism for pC221, with initiation being carried out by nicking of oriD by REP D, and the free 3' hydroxyl group generated acting as a primer for replication.

Novobiocin inhibited incorporation by approximately 50%, suggesting that the topological state of the DNA affects the replication of pC221 *in vitro*. As DNA gyrase uses the energy stored in ATP to increase negative supercoiling in the DNA, this suggests that negative supercoiling of the DNA favours replication of pC221 *in vitro*.

4.2.3: - Use of different plasmid substrates in the *in vitro* replication system.

Two other plasmids, pT181 and pCW7, which are related to pC221 were used in the *in vitro* replication system with added REP D. Since a particular REP protein is thought to be specific for its cognate origin *in vivo* (Projan & Novick, in press), pT181 and pCW7 were not expected to give any REP D-dependent incorporation of radioactivity into acid precipitable material *in vitro*. However, when time course experiments were carried out using supercoiled pC221, pT181

and pCW7 in the presence and absence of REP D, all three showed REP D-dependent incorporation of radioactivity into acid precipitable material (Figure 4.7). Therefore, it appears that the specificity which a particular REP protein is thought to show for its cognate origin *in vivo* does not hold *in vitro*.

Labelling of these plasmids with $\alpha[^{32}P]$ dATP was carried out using the *in vitro* replication system with added REP D, followed by restriction endonuclease digestion, agarose gel electrophoresis and autoradiography. This suggested that replication of pT181 and pCW7 , as well as pC221, was initiating within the region containing the origin and proceeding unidirectionally in the expected direction (Figure 4.8).

As a control, a *S. aureus* plasmid called pC194, which has no rep sequences homologous with the pC221 repD gene, was used in the *in vitro* replication system (Figure 4.9). This plasmid showed little significant REP D-dependent incorporation of radioactivity into acid precipitable material over a time course experiment lasting 1 hour.

4.2.4: - The specificity of REP D appears to increase with substrate plasmids of low superhelical density.

In prior experiments, the plasmid DNA used in the *in vitro* replication system had been in the supercoiled state. It is, therefore, possible that the high superhelical density of the DNA may have an effect on the origin which is not normally present *in vivo*. Previous experiments appear to suggest that the superhelical density of the DNA may have an effect on the initiation of replication. Energy stored in negative DNA supercoiling can be used to stabilise

the formation of unusual structures, such as hairpin loops (Lilley, 1980). It is thought that the natural superhelical density of plasmid DNA *in vivo* is low, mainly due to the presence of bound proteins. The latter may stabilise local melting of the DNA or the formation of secondary structures, both of which tend to remove negative supercoils from the DNA.

Therefore, it may give a more accurate picture of what occurs *in vivo* if the substrate plasmid were added to the *in vitro* replication system in the relaxed form. Calf thymus DNA topoisomerase I was used to relax the plasmid DNA. The assays with relaxed plasmid DNA were carried out in the presence of novobiocin, as the extracts were found to contain a very active DNA gyrase (Figure 4.10). Although relaxed pC221 DNA still gave REP D-dependent incorporation, relaxed pT181 did not (Figure 4.11). Therefore it appears that the specificity of REP D for its cognate origin, oriD, *in vitro* is increased relative to oriC when the substrate plasmids are provided in the relaxed form. This would appear to agree with the work carried out *in vivo* (Projan *et al*, in press).

This apparent difference in specificity of REP D between relaxed and supercoiled DNA *in vitro* may be due to a specific structure which is required at the origin for REP to nick. Such a structure could be a hairpin at one or more of the three inverted complementary repeat sequences (ICR I, II and III) of the origin. *In vivo*, or *in vitro* with relaxed plasmid DNA, the initial event is presumably sequence-specific binding over the region of ICR III of the origin. Once REP is bound, it may then stabilise the formation of one or more hairpin structures at the origin. This conformational change in the DNA may then enable REP to nick the DNA within ICR II. *In vitro*, with supercoiled plasmid DNA, such a conformational change might take place in the absence of bound REP, due to

stabilisation by the negative supercoiling of the DNA. This might allow nicking to take place without the need for sequence-specific recognition by REP.

Alternatively, it may be that, when the substrate DNA is in the supercoiled state, the affinity of REP D for the origin is increased to such an extent that any difference in the sequence-specific binding regions is overcome. It has been shown that some DNA binding proteins bind more strongly to their specific targets when the DNA is supercoiled, e.g. the lac repressor protein (Whitson et al, 1987). This is, however, probably due to energy stored in negative supercoiling which can facilitate structural changes in the DNA, such as local melting of the double-stranded helix or formation of hairpin structures or Z DNA, which may be required for binding of the protein.

Table.4.1: Assays with cell free extracts prepared from RN1786 and RN1786/pC221.

Assay mixtures contained 5µl of extract, 0.5µg of supercoiled pC221, appropriate buffers, nucleotide triphosphates, K^+ and Mg^{++} ions in a total volume of 25µl (see section 2.3.5.). Incubation was at 30°c for 1 hour and acid precipitable radioactivity was determined by liquid scintillation counting. The results shown are the average values of duplicate experiments.

EXTRACT SOURCE	DNA ADDED	RADIOACTIVITY INCORPORATED (CPM)
S. aureus RN1786	NONE	403
"	pC221	720
S. aureus RN1786 /pC221	NONE	674
"	pC221	3486

Table 4.2: The effects of rifampicin and novobiocin on pC221 replication *in vitro.*

Typical assays were carried out, in the presence and absence of REP D (50ng), in the presence of rifampicin (50μg/ml) or novobiocin (50μg/ml). The results shown are average values from duplicate experiments.

The presence of rifampicin appeared to have little effect on REP D-dependent incorporation, suggesting that rifampicin-sensitive RNA synthesis is not required for initiation of pC221 replication *in vitro.*

Novobiocin, however, inhibited REP D-dependent incorporation by approximately 50%, suggesting that negative supercoiling of the DNA facilitates the replication of pC221 *in vitro.*

REP D	RIFAMPICIN	NOVOBIOCIN	RADIOACTIVTY INCORPORATED (CPM)
-	-	-	2309
+	-	-	14151
-	+	-	3808
+	+	-	12870
-	-	+	2422
+	-	+	6846

Fig 4.1: REP D dose response curve with pC221 DNA.

The assays were carried out under typical conditions in the presence of purified REP D (0 - 150ng). Incorporation increased in a linear fashion, initially, with increasing REP D concentration. This suggests that REP D is required for the replication of pC221 *in vitro*. The curve levelled off after an addition of 60ng of REP D, showing that some other rate limiting factor comes into play here. This quantity of REP D corresponds to a molar ratio of 2:1 REP D dimers to DNA molecules. This suggests that some of the REP D is inactive and DNA is the limiting factor, or that reinitiation can take place on DNA molecules that are already replicating and some other limiting factor, such as nucleotide triphosphates, comes into play.

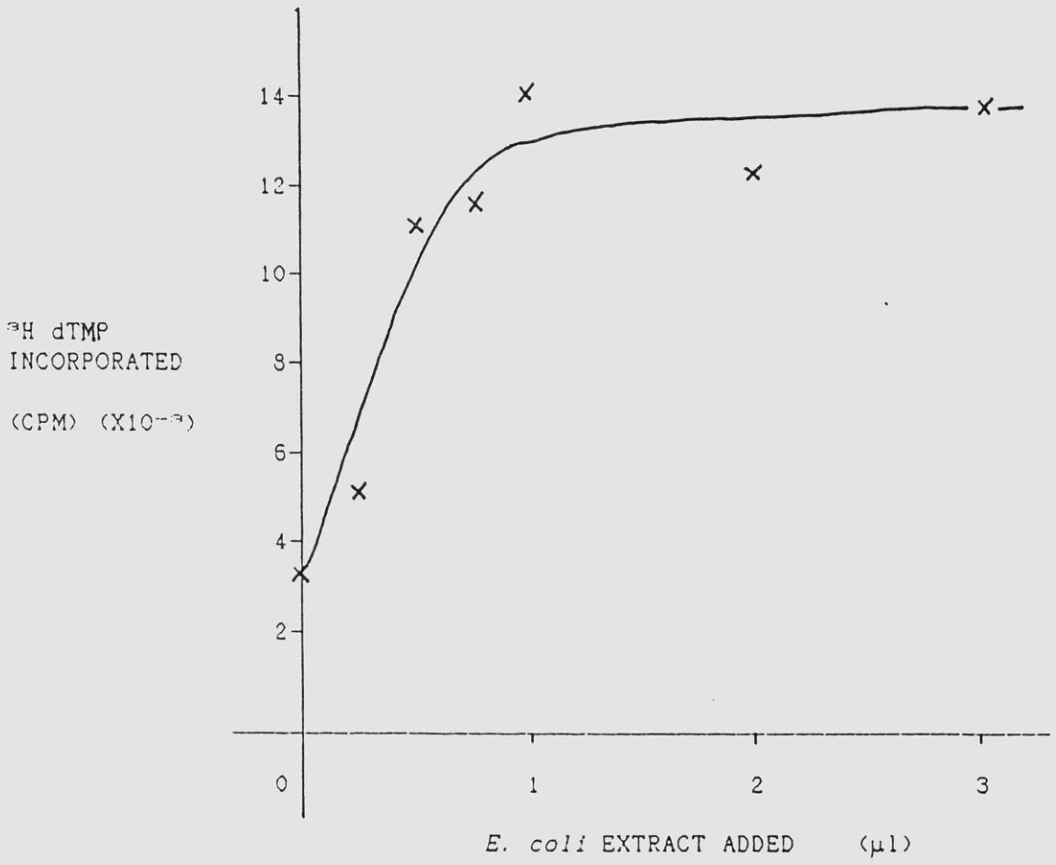

^{3}H dTMP
INCORPORATED
(CPM) (X10^{-3})
14
12
10
8
6
4
2
0
1
2
3
E. coli EXTRACT ADDED (µl)

Fig 4.2: Time course experiment with pC221 in the presence and absence of REP D.

Typical assays were carried out in the presence and absence of REP D (50ng), with incubation times varying between 0 and 160 minutes. Incorporation was increased in the presence of REP D. However, there was some background incorporation in the absence of REP D, probably due to DNA repair mechanisms. This background incorporation appeared to be complete after about 30 minutes.

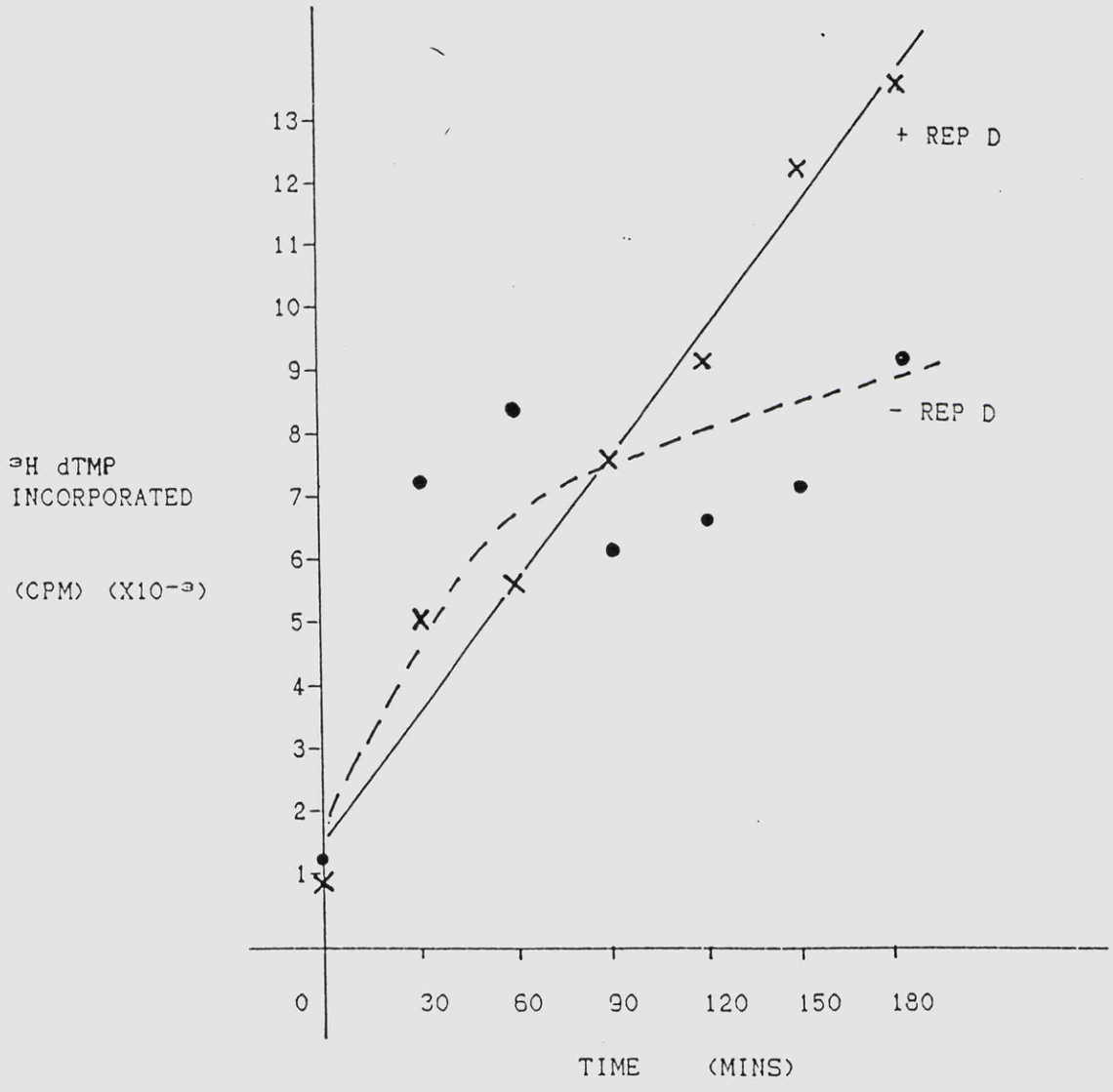

³H dTMP
INCORPORATED
(CPM) (X10⁻³)
13
12
11
10
9
8
7
6
5
4
3
2
1
0
30
60
90
120
150
180
TIME (MINS)
+ REP D
- REP D

Fig 4.3: Time course experiment with preincubation in the absence of radioactive label.

Assays were carried out under typical conditions in the presence and absence of REP D (10ng). Cold assay mixtures were preincubated at 30°c for 30 minutes before addition of the radioactive label and further incubation at 30 °c for the appropriate time (0 - 60 minutes). This cut down the amount of background incorporation appreciably, but still gave increased incorporation in the presence of REP D. Incorporation increased in a linear fashion, initially, tailing off after about 40 minutes. There was an approximately 6 fold increase in incorporated radioactivity over the background level after 60 minutes.

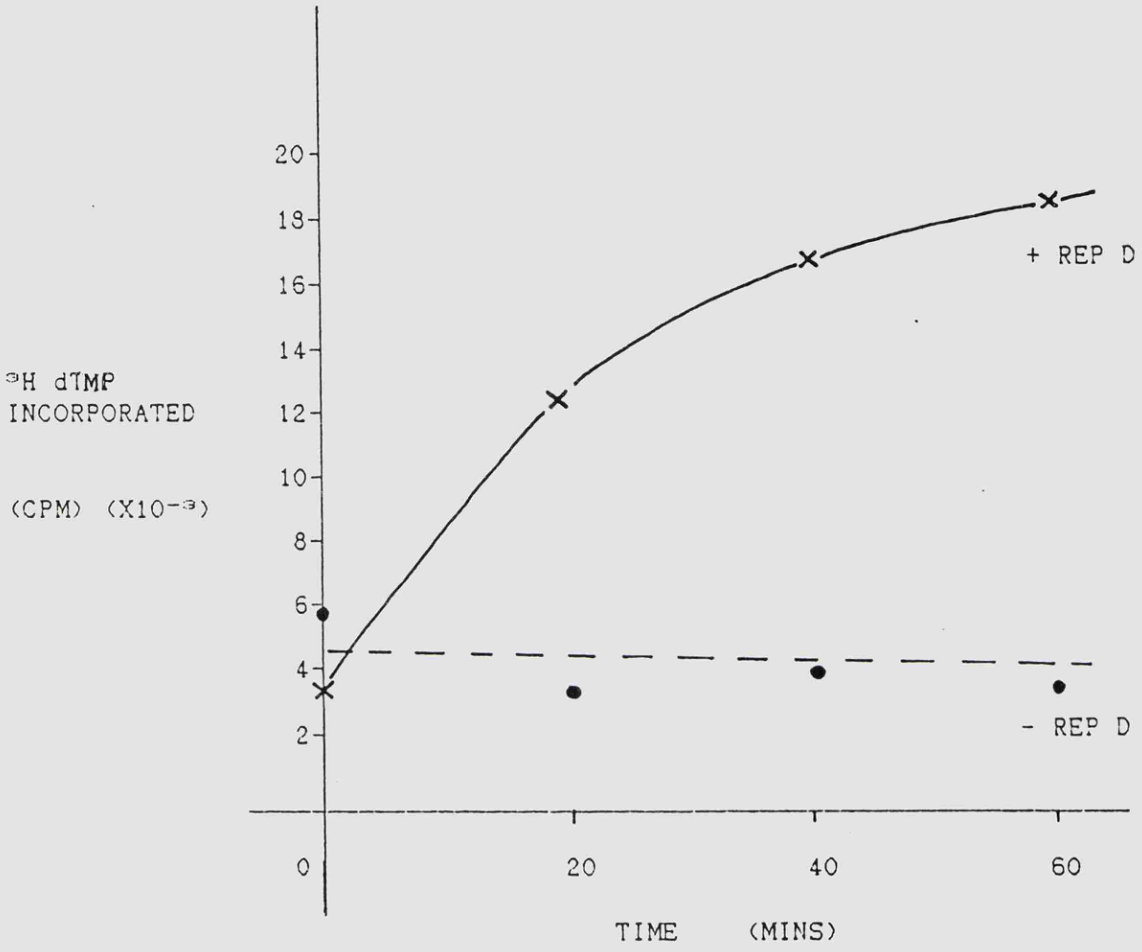
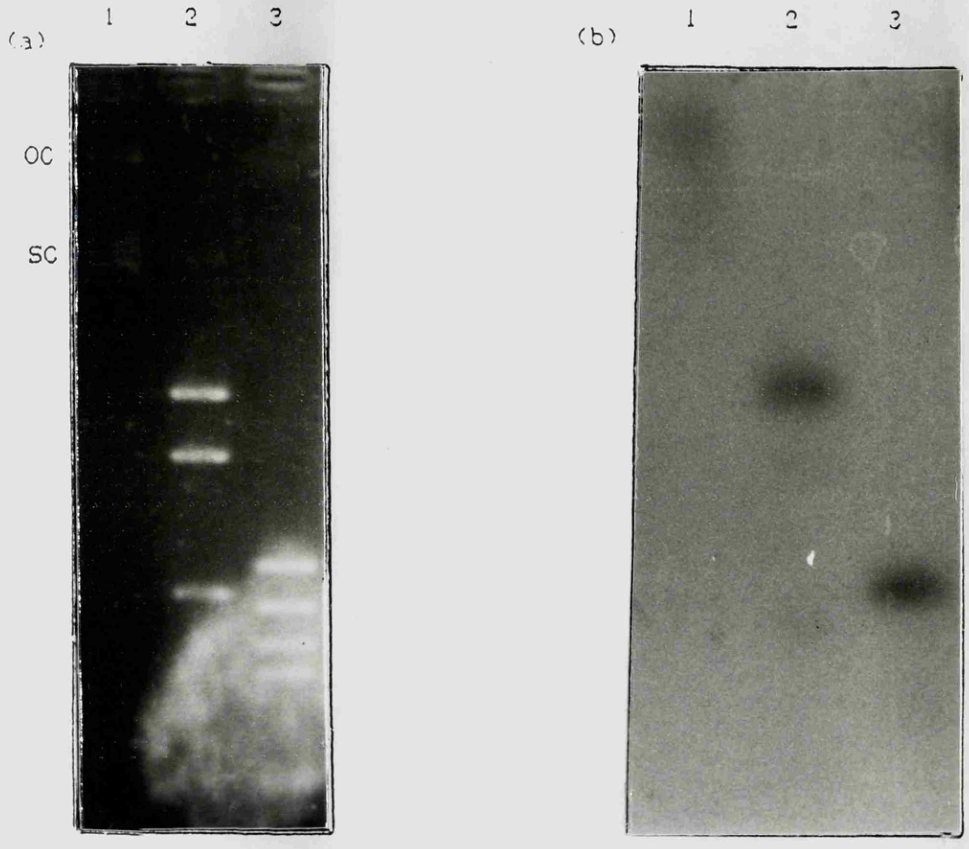

^{3}H dTMP
INCORPORATED
(CPM) (X10^{-3})
20
18
16
14
12
10
8
6
4
2
0
20
40
60
TIME (MINS)
+ REP D
- REP D

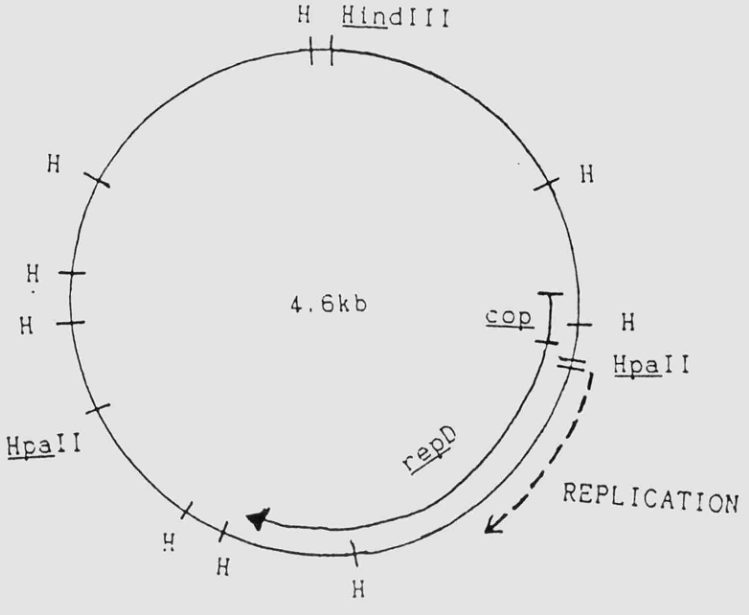
Fig 4.4: Labelling of pC221 with $\alpha[^{32}P]$ dATP using the *in vitro* replication system.

The *in vitro* replication system was used to label pC221 with $\alpha[^{32}P]$ dATP. Assays were carried out under typical conditions in the presence of $\alpha[^{32}P]$ dATP. The products were subjected to restriction endonuclease digestion, agarose gel electrophoresis and autoradiography. The resulting gel (a) and its autoradiograph (b) are shown.

Track 1 - labelled pC221 (undigested), Track 2 - labelled pC221 (HpaII /HindIII digest) and Track 3 - labelled pC221 (HinfI digest). The positions of supercoiled (SC) and relaxed and /or open circular (OC) DNA are indicated.

The pC221 restriction map (c) is shown. The pattern of labelling of the restriction fragments suggests that replication is unidirectional, in the direction expected for elongation at the free 3' OH group generated at the nick site, and is initiated within the region of the plasmid containing the origin.





H = HinfI SITE

(c) pC221 MAP

Fig 4.5: Assay for REP D activity in extracts of *E. coli* which over-express the repD gene.

REP D was purified by cloning of the repD gene into pHD, a temperature-sensitive, inducable, dual-origin vector , followed by overexpression in *E. coli*. The *E. coli* extracts were partially purified by ammonium sulphate factionation (30% saturation). The partially purified extracts were assayed for REP D activity using the *in vitro* replication system.

Varying ammounts of the extract were added to typical assay mixtures, which were then incubated at 30°c for 1 hour. Incorporation was increased on addition of the extract, levelling off at an addition of 1µl of extract.

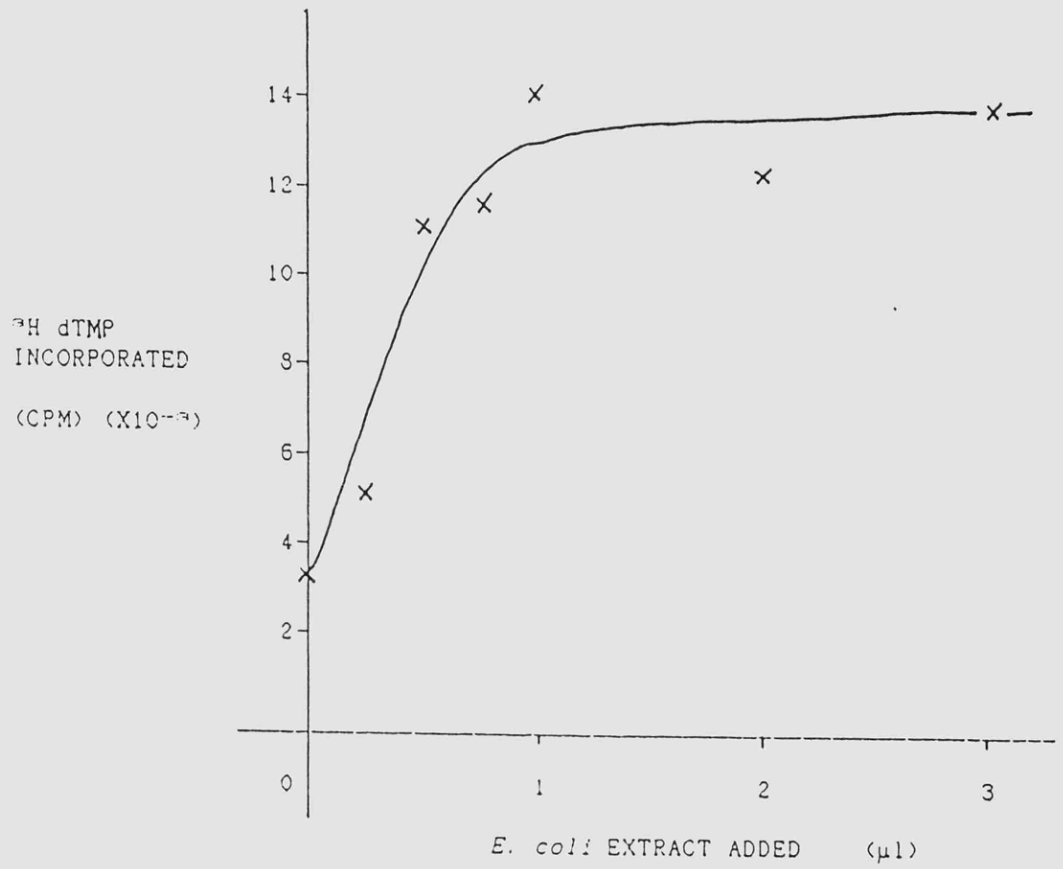

^{3}H dTMP
INCORPORATED
(CPM) (X10^{-3})
14
12
10
8
6
4
2
0
1
2
3
E. coli EXTRACT ADDED (µl)

Fig 4.6: Assay for REP D activity in fractions collected from a heparin sepharose column.

Extracts of *E. coli* in which repD is overexpressed were partially purified by ammonium sulphate fractionation (30% saturation). Extracts were applied to a heparin sepharose column and eluted by increasing salt concentration (KCl).

The A_{280} of the collected fractions was measured and peak fractions were assayed for REP D activity using the *in vitro* replication system. Aliquots (5µl) of each fraction were added to typical assay mixtures, which were incubated at 30°c for 1 hour.

The graph shows part of the elution profile for REP D. The *in vitro* REP D activity (shown by the solid line) corresponds to the A_{280} peak (dotted line), suggesting that the protein peak seen here is REP D.

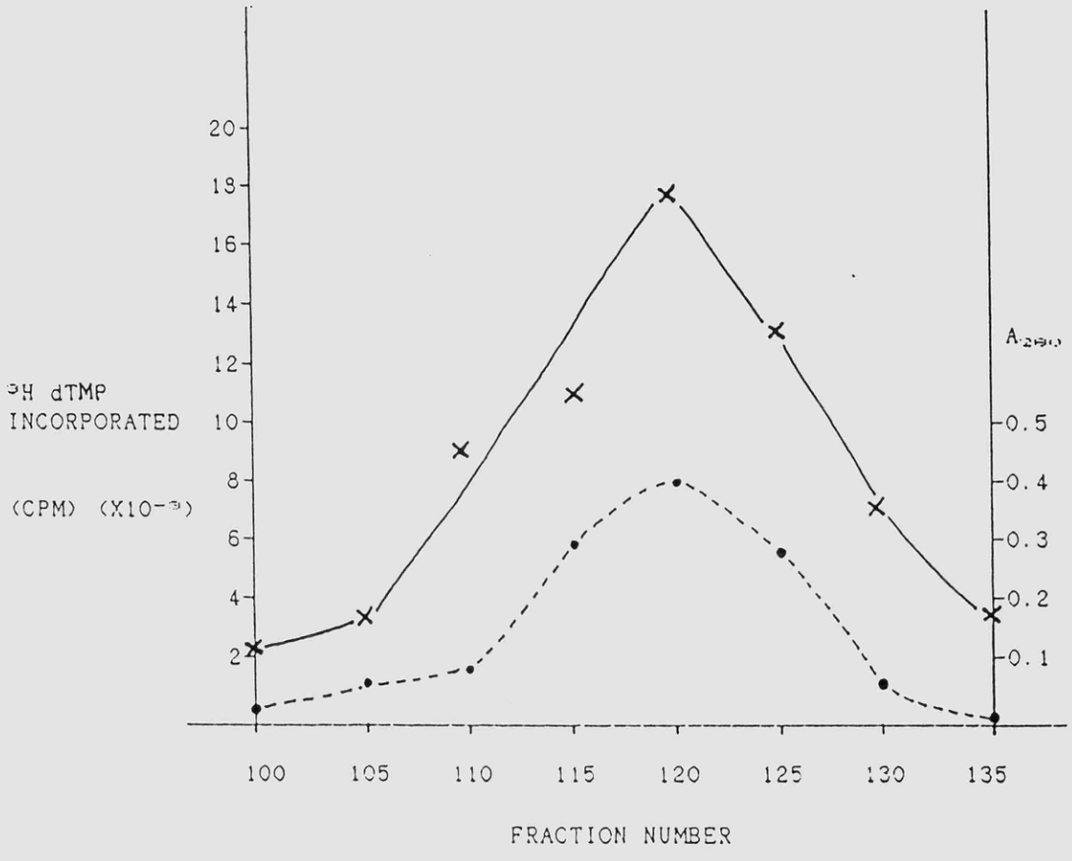

^{3}H dTMP
INCORPORATED
(CPM) (X10^{-3})
20
18
16
14
12
10
8
6
4
2
A_{280}
0.5
0.4
0.3
0.2
0.1
100
105
110
115
120
125
130
135
FRACTION NUMBER

Fig 4.7: Time course experiments with pC221, pT181 and pCW7 in the presence and absence of REP D.

Time course experiments were carried out using the *in vitro* replication system, in the presence and absence of REP D, with pC221, pT181 and pCW7. Assays were carried out under typical conditions. In the absence of REP D, the background incorporation was low in each case. All three showed increased incorporation in the presence of REP D, although incorporation was slightly higher with pC221. These results were unexpected, as experiments carried out *in vivo* have suggested that a particular REP protein is absolutely specific for its cognate origin sequence (Projan *et al*, in press).

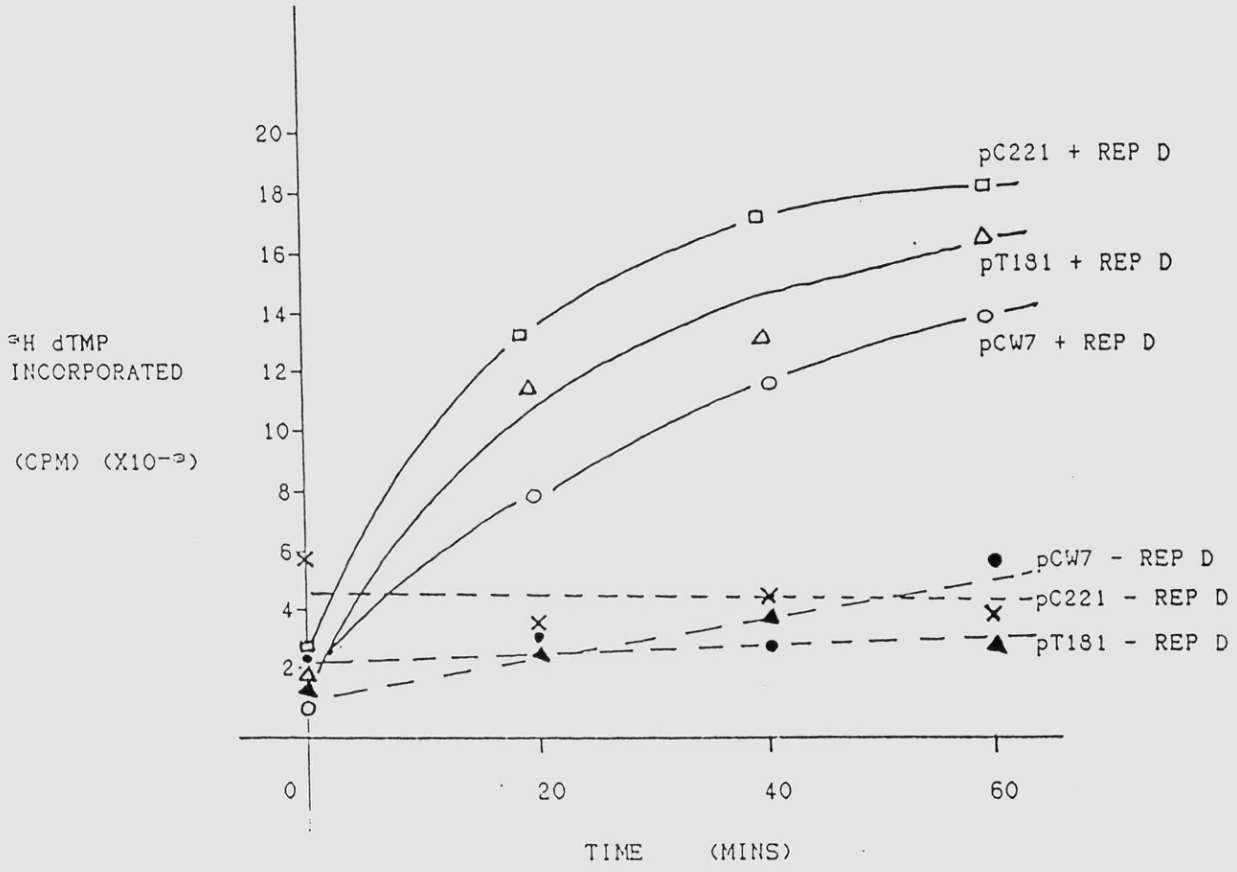

^{3}H dTMP
INCORPORATED
(CPM) (X10^{-3})
20
18
16
14
12
10
8
6
4
2
0
20
40
60
TIME (MINS)
pC221 + REP D
pT181 + REP D
pCW7 + REP D
pCW7 - REP D
pC221 - REP D
pT181 - REP D

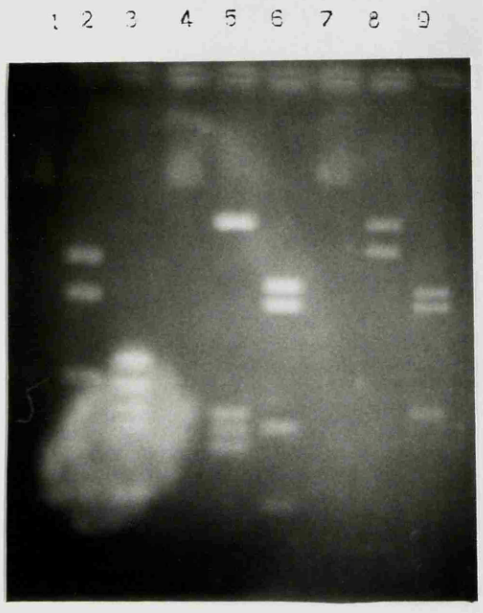
Fig 4.8: Labelling of pC221, pT181 and pCW7 with $\alpha[^{32}P]$ dATP, using the *in vitro* replication system.

The *in vitro* replication system was used to label pC221, pT181 and pCW7 with $\alpha[^{32}P]$ dATP in the presence of REP D protein. The assays were carried out under typical conditions and the products subjected to restriction endonuclease digestion, agarose gel electrophoresis and autoradiography. The resulting gel (a) and autoradiograph (b) are shown.

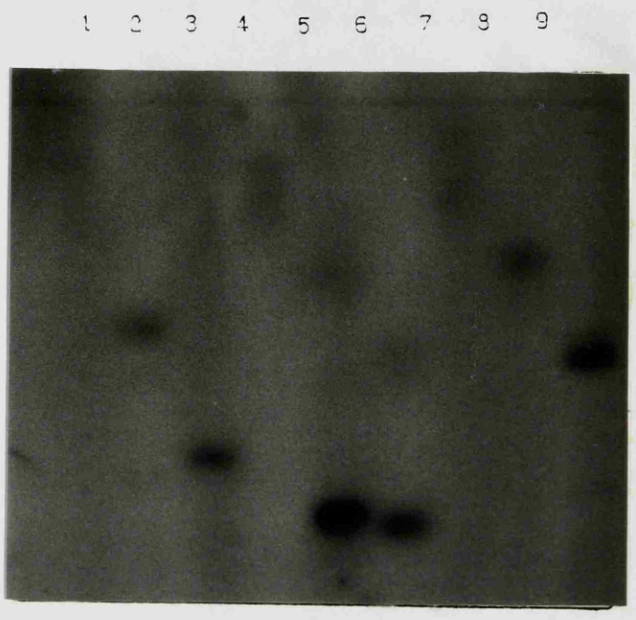
Track 1 - labelled pC221 (undigested), Track 2 - labelled pC221 (HpaII /HindIII digest), Track 3 - labelled pT181 (HinfI digest), Track 4 - labelled pT181 (undigested), Track 5 - labelled pT181 (HpaII /HindIII digest), Track 6 - labelled pT181 (TaqI digest), Track 7 - labelled pCW7 (undigested), Track 8 - labelled pCW7 (HpaII /HindIII digest), Track 9 - labelled pCW7 (HinfI digest).

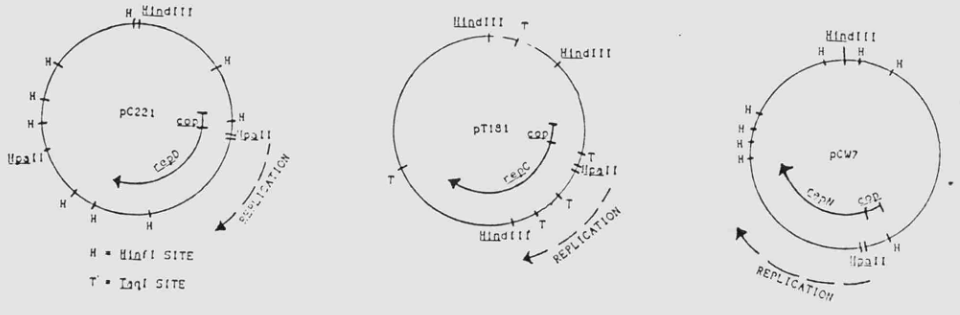
The patterns of labelling of restriction fragments obtained suggested that, in each case, the replication was initiated within the region containing the origin sequence, and was proceeding in the direction expected for elongation at the free 3' OH group generated by nicking of the origin sequence.

(a)



(b)





pC221, pT181 AND pCW7 RESTRICTION MAPS

Fig 4.9: Time course experiments with pC221 and pC194.

Time course experiments were carried out under typical conditions, in the presence and absence of REP D (50ng), with pC221 and pC194 DNA. The results show that, although incorporation with pC221 is increased in the presence of REP D, very little REP D-dependent incorporation is seen with pC194, a staphylococcal plasmid which has no sequences homologous to the pC221 repD gene. The low level of REP D dependent incorporation seen with pC194 may reflect some homology with the pC221 origin within pC194. This would not be entirely surprising as pC194 is thought to replicate via a rolling circle mechanism similar to that for pC221, although there is no obvious homology within the replication regions of these plasmids. Another explanation may be that the presence of a low level of contamination seen in the REP D, which is thought to be associated with non-specific nuclease activity, may degrade the DNA to some extent, leading to DNA repair replication.

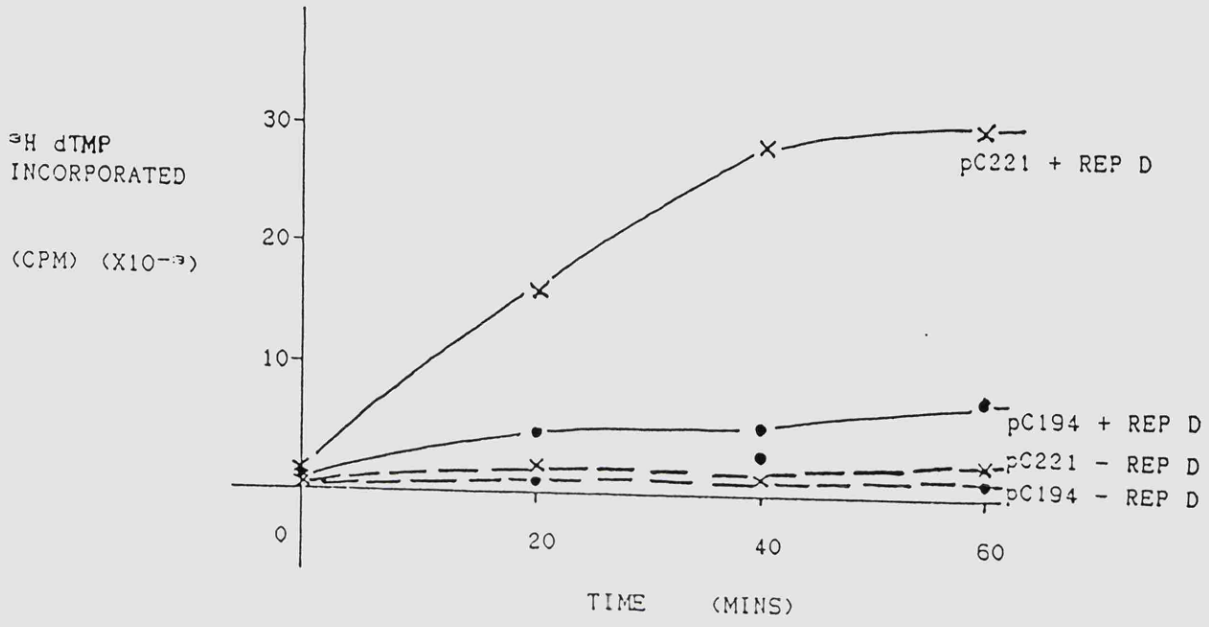
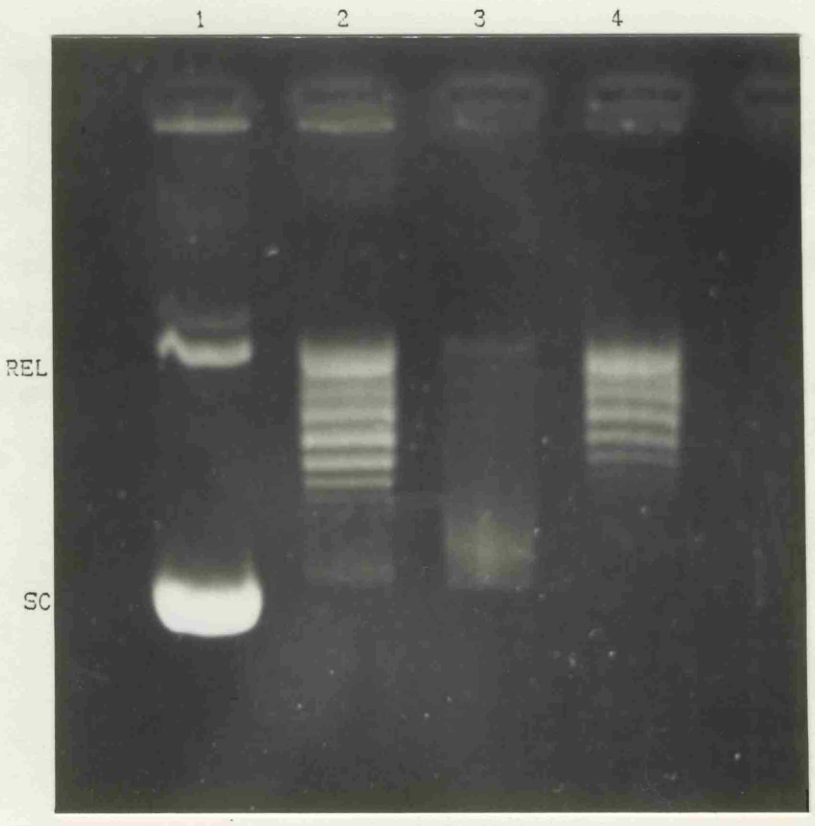

³H dTMP
INCORPORATED
(CPM) (X10⁻³)
30
20
10
0
20
40
60
TIME (MINS)
pC221 + REP D
pC194 + REP D
pC221 - REP D
pC194 - REP D

Fig 4.10: DNA gyrase activity present in *S. aureus* cell free extracts.

Track 1 - supercoiled pC221 DNA (0.5μg), Track 2 - pC221 DNA (0.5μg) treated with calf thymus topoisomerase I (10 units), Track 3 - pC221 DNA (0.5μg) treated with calf thymus topoisomerase I (10 units), followed by treatment with *S. aureus* cell free extract (5μl) and Track 4 - as track 3, in the presence of novobiocin (1.25μg).

Treatment of supercoiled DNA with calf thymus topoisomerase I causes relaxation to a series of topoisomers of low superhelical density. Treatment of this relaxed DNA with the cell free extract increases the superhelical density of the DNA only in the absence of novobiocin, an inhibitor of DNA gyrase.

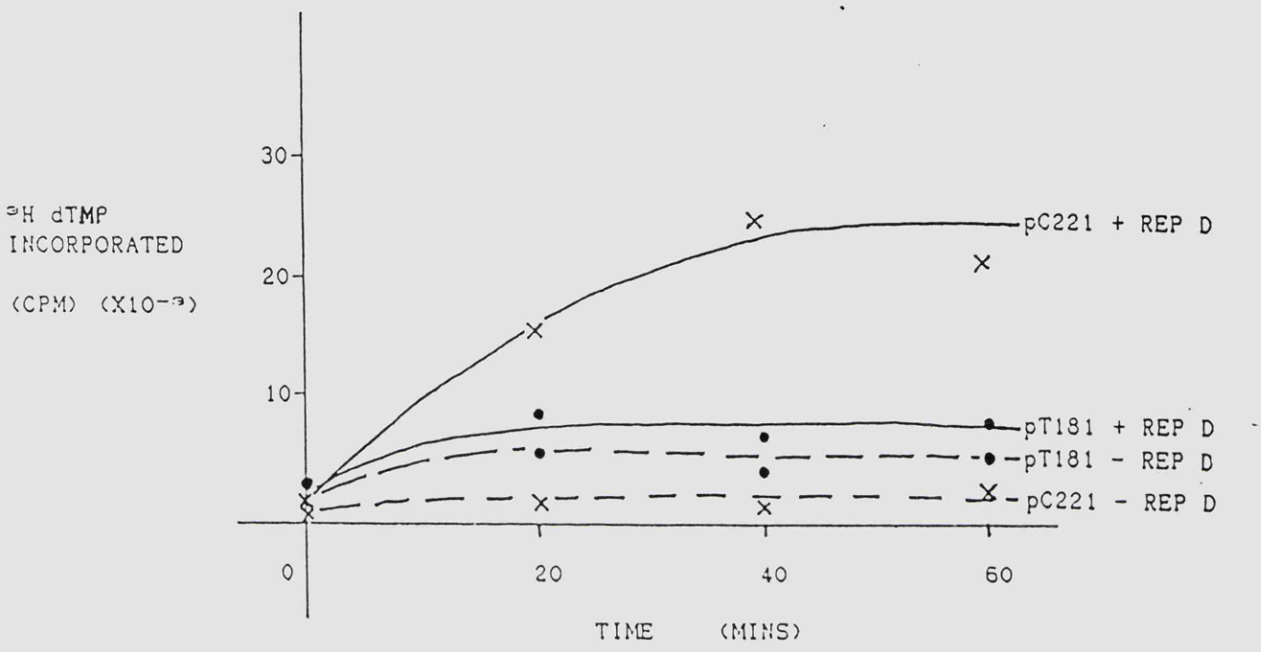


(SC = SUPERCOILED pC221)

(REL = RELAXED pC221)

Fig 4.11: Time course experiment using relaxed pC221 and pT181 DNA in the presence and absence of REP D.

Time course experiments were carried out under typical conditions in the presence and absence of REP D with pC221 and pT181 DNA relaxed with calf thymus topoisomerase I. The assays were carried out in the presence of novobiocin (50μg/ml), as the extracts have been shown to have DNA gyrase activity. Relaxed pC221 DNA still shows REP D-dependent incorporation, but relaxed pT181 DNA does not. This suggests that the specificity of REP D for its cognate origin, oriD, is increased over that for oriC when the superhelical density of the substrate plasmid is low.


³H dTMP
INCORPORATED
(CPM) (X10-³)
30
20
10
0
20
40
60
TIME (MINS)
pC221 + REP D
pT181 + REP D
pT181 - REP D
pC221 - REP D

CHAPTER 5.

FOOTPRINTING OF THE pC221 oriD REGION WITH REP D.

5.1. - Introduction.

5.2. - Results and discussion.

CHAPTER 5.

DNASEI FOOTPRINTING OF THE pC221 ori REGION WITH REP D PROTEIN.

5.1: - Introduction.

DNA footprinting has been carried out in order to demonstrate sequence-specific binding of REP C on a 157bp PvuI /TaqI restriction fragment of pT181 containing oriC (Koepsel *et al*, 1986). This showed that REP C binds specifically to a 32bp region within oriC, immediately downstream of the nick site. Experiments were, therefore, carried out to determine the position of binding of REP D at oriD in order to compare it with REP C. DNase I, a non-specific endonuclease, was used for footprinting studies of REP D.

5.2: - Results and discussion.

DNase I protection experiments were carried out according to the method of Galas & Schmitz (1978). Plasmid DNA (pC221) was digested with HinfI and the largest (976bp) restriction fragment was isolated by agarose gel electrophoresis, followed by electroelution of the fragment from a gel slice. The fragment was then [^{32}P] end-labelled at the 5' or 3' end and cut with a second restriction endonuclease, AluI. A 186bp fragment containing oriD, labelled at one end, was again isolated by gel electrophoresis and electroelution.

In order to determine the exact binding site of REP D on the end-labelled fragment, DNase I protection experiments were carried out. Approximately 10 - 20ng of end-labelled DNA were treated with DNase I (4ng) for 5 minutes at room temperature in the presence and absence of REP D (0.6µg). The samples were run on a 6% polyacrylamide /urea sequencing gel against dideoxy sequencing tracks as size markers (Figure 5.1).

In the absence of REP D, DNase I partial digestion of the end-labelled DNA gave a well distributed ladder of bands, both with the 5' and 3' end-labelled DNA (Figure 5.1. - Tracks [1] and [3]). However, when the DNA was preincubated with a greater than 100:1 ratio of REP D to DNA, the pattern of digestion observed was different.

The DNaseI digestion of 5' end-labelled DNA in the presence of REP D (Figure 5.1. - Track [4]) has a gap in the ladder of bands which shows the position of protection of the DNA from the nuclease by REP D on the upper strand. There are also some strong bands present within this region which, in

some cases , are more intense than the corresponding bands in the digestion pattern in the absence of REP D. These correspond to regions of the DNA which are still accessible to cleavage by DNase I when REP D is bound. In the cases where the band is more intense than in the absence of REP D, i.e. when the cleavage by DNaseI is enhanced, the DNA is somehow made more accessible to DNase I by the binding of REP D.

The very intense band, marked with an asterisk, corresponds to the nicking of the oriD region by REP D. The DNase I footprint of REP D stretches over a region of 38bp immediately downstream of the nick site on the upper strand. There is enhancement of DNase I cleavage at two positions, 20 and 31 bases respectively downstream from the nick site.

The 3' end-labelled DNA also shows a footprint pattern (Figure 5.1. - Track [2]) which shows the position of protection of the DNA by REP D on the lower strand. There is no strong band due to nicking by REP D here, as REP D only cleaves the origin in the top strand. In this case, the DNase I footprint stretches over 34 bases, covering approximately the same region which is protected on the upper strand. There is enhancement of DNase I cleavage at three positions, one 5 bases, one 15 bases and a cluster of three sites at around 25 bases downstream of the nick site.

Therefore, DNase I footprinting data show that REP D binds to a 38bp region of oriD, immediately downstream of the nick site, which encompasses the whole of ICR III, along with the 5' side of ICR II (Figure 5.2). The pattern of DNase I enhanced cleavage obtained appeared to have a periodicity of approximately 10 bases, which is displaced by 5 bases on each strand relative to the other. This might suggest that only one side of the DNA double helix

interacts closely with the protein, with the face of the DNA helix facing away from the protein being accessible to cleavage by DNase I.

Preliminary results using hydroxyl radical footprinting methods (Figure 5.3), which are expected to give a much higher resolution footprint due to the high reactivity and small size of the attacking hydroxyl radical (Tullius, 1987), suggest that close contacts between DNA and protein are not continuous over the entire footprinted region (results not shown). The close contacts appear to show a similar periodic pattern to the enhanced cleavage pattern obtained with DNase I footprinting, giving further evidence that REP D only interacts with one face of the DNA helix.

Bending of the DNA on binding to REP D might explain the enhanced DNase I cleavages observed, as this would cause an opening of the minor groove, giving greater accessibility to DNase I (Figure 5.4). There is evidence for some bending at the pT181 origin (oriC) on binding of REP C (Koepsel & Khan, 1986).

The footprinting data shows no apparent contact between REP D and DNA upstream of the nick site, suggesting that bases here are unimportant for REP D binding. However, it must be remembered that the data obtained here gives information on REP D bound to the DNA after nicking has taken place. Therefore, it may be that other bases are initially involved in REP D binding and a conformational change takes place after nicking, thereby altering the mode of binding.

Fig 5.1: DNase I footprinting of REP D on oriD.

The autoradiograph shows the pattern of protection of oriD by REP D from DNase I. Protection of the lower strand is shown by the footprint on the 3' end-labelled DNA, whereas protection on the upper strand is shown by the footprint on the 5' end labelled DNA.

Track 1 - 3' end-labelled DNA (10ng), DNase I treated (4ng, 5min, room temp), no REP D .

Track 2 - 3' end-labelled DNA (10ng), DNase I treated (4ng, 5min, room temp), plus REP D (0.6μg).

Track 3 - 5' end-labelled DNA (10ng), DNase I treated (4ng, 5min, room temp), no REP D.

Track 4 - 5' end-labelled DNA (10ng), DNase I treated (4ng, 5min, room temp), plus REP D (0.6μg).

The DNA sequence is shown adjacent to the DNase I footprint. Inverted complementary repeat sequence (ICR) III of ori D is denoted by the arrows.

The results suggested that REP D binds to a 38 bp region immediately downstream of the nick site. The pattern of enhanced DNase I cleavage sites observed suggested that only one face of the DNA helix was in close contact with REP D, and that the DNA may be bent on binding of REP D. The very intense band, marked with an asterist, was due to cleavage of the DNA by REP D at the specific nick site.

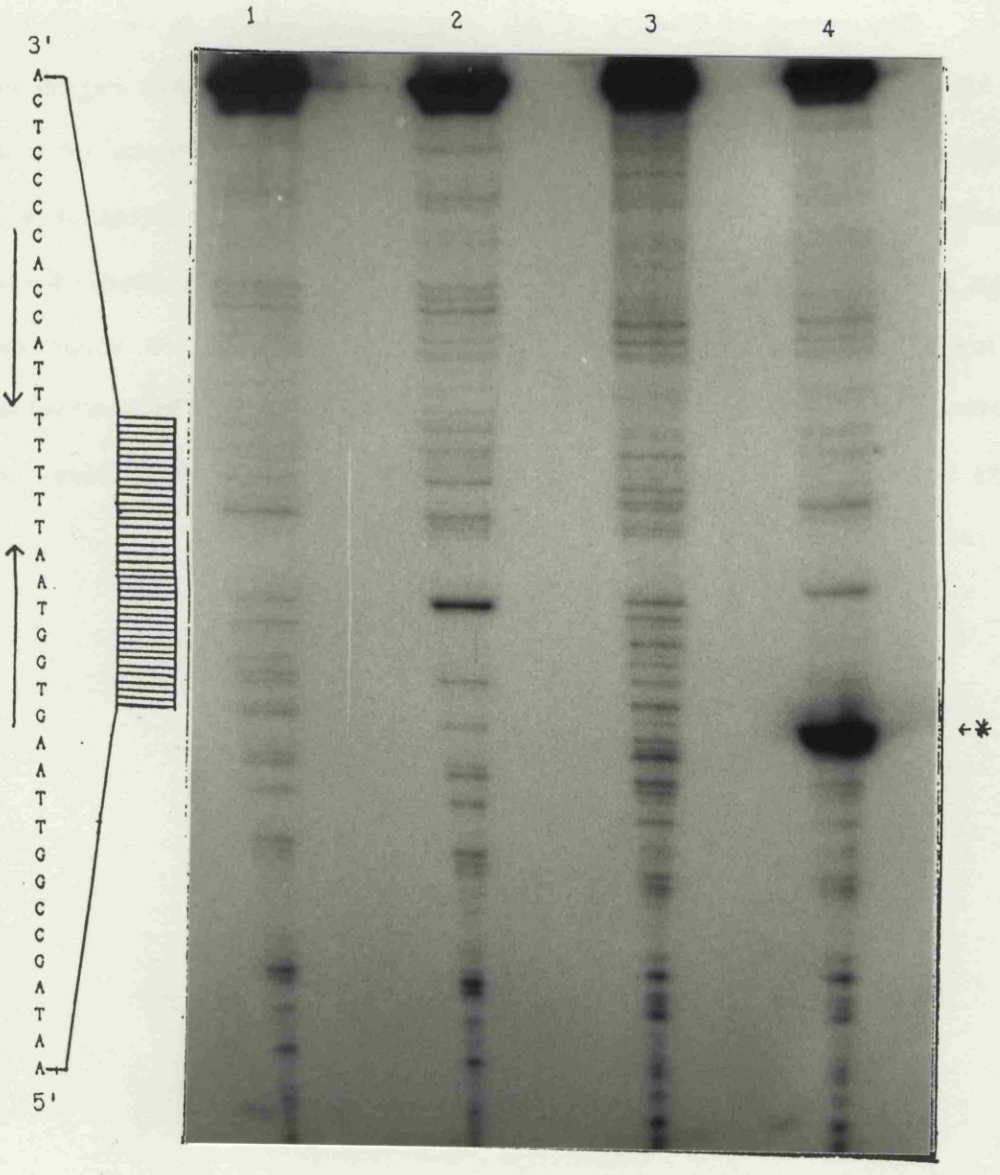

1
2
3
4
3'
A
C
T
C
C
C
C
A
C
C
A
T
T
T
T
T
T
T
T
A
A
T
G
G
T
G
A
A
T
T
G
G
C
C
G
A
T
A
A
5'

Fig 5.2: Diagram of oriD and the DNase I footprint of REP D.

The extent of binding of REP D on either strand is denoted by solid lines. ICR II and III are denoted by dotted arrows. The protected region extends over approximately 38bp on both strands, immediately downstream of the nick site, covering the whole of ICR III and the 5' side of ICR II. Sites of enhanced cleavage by DNase I are shown by solid triangles. Sites of DNase I cleavage which are the same in the presence and absence of REP D are shown by open triangles. There is an approximate periodicity of 10 bases in the pattern of enhanced DNase I cleavage sites observed, which is displaced by 5 bases on each strand relative to the other.

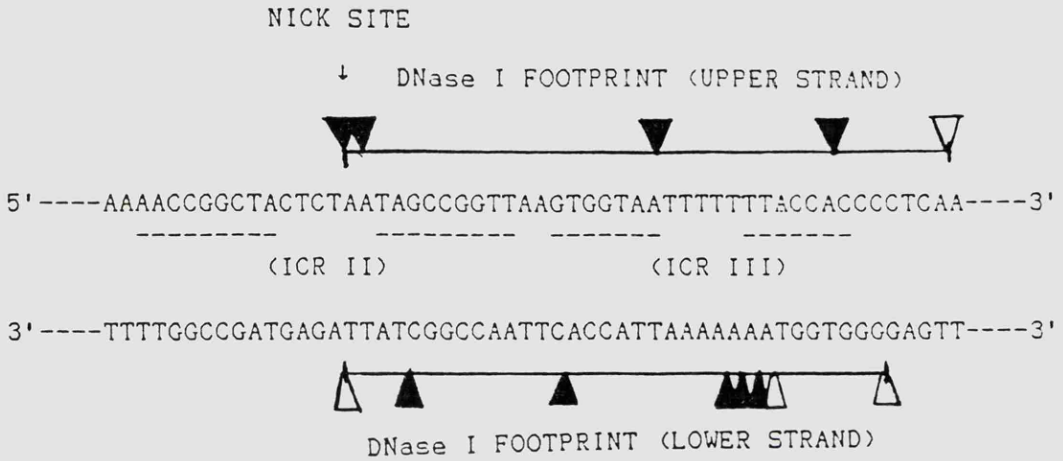

NICK SITE
DNase I FOOTPRINT (UPPER STRAND)
5'----AAAACCGGCTACTCTAATAGCCGGTTAAGTGGTAATTTTTTTACCACCCCTCAA----3'
(ICR II)
(ICR III)
3'----TTTTGGCCGATGAGATTATCGGCCAATTCACCATTAAAAAAATGGTGGGGAGTT----3'
DNase I FOOTPRINT (LOWER STRAND)

Fig 5.3: Hydroxyl radical footprinting (Tullius, 1987).

Footprinting techniques using the hydroxyl radical as the nucleolytic species give a high resolution footprint, with only the regions of the DNA in very close contact with the DNA binding protein being protected. This is due to three properties of the hydroxyl radical, it s small size, it s high reactivity and it s very low preference for the particular nucleotide attacked.

The main features of this method are illustrated. A negatively charged complex of iron (II) EDTA reacts with hydrogen peroxide to produce hydroxyl radicals (·OH). These then make single-stranded nicks in the DNA by extracting a hydrogen atom from a deoxyribose residue in the DNA backbone, which then breaks down by secondary reactions to give chain cleavage. The conditions are adjusted so that each DNA molecule is cleaved, at most, once (as denoted by the heavy arrow).

The site of attack is shown in greater detail, encircled. The products of the reaction are DNA molecules that have at their ends the phosphates that were originally attached to the deoxyribose moiety (the victim of the hydroxyl radical attack).

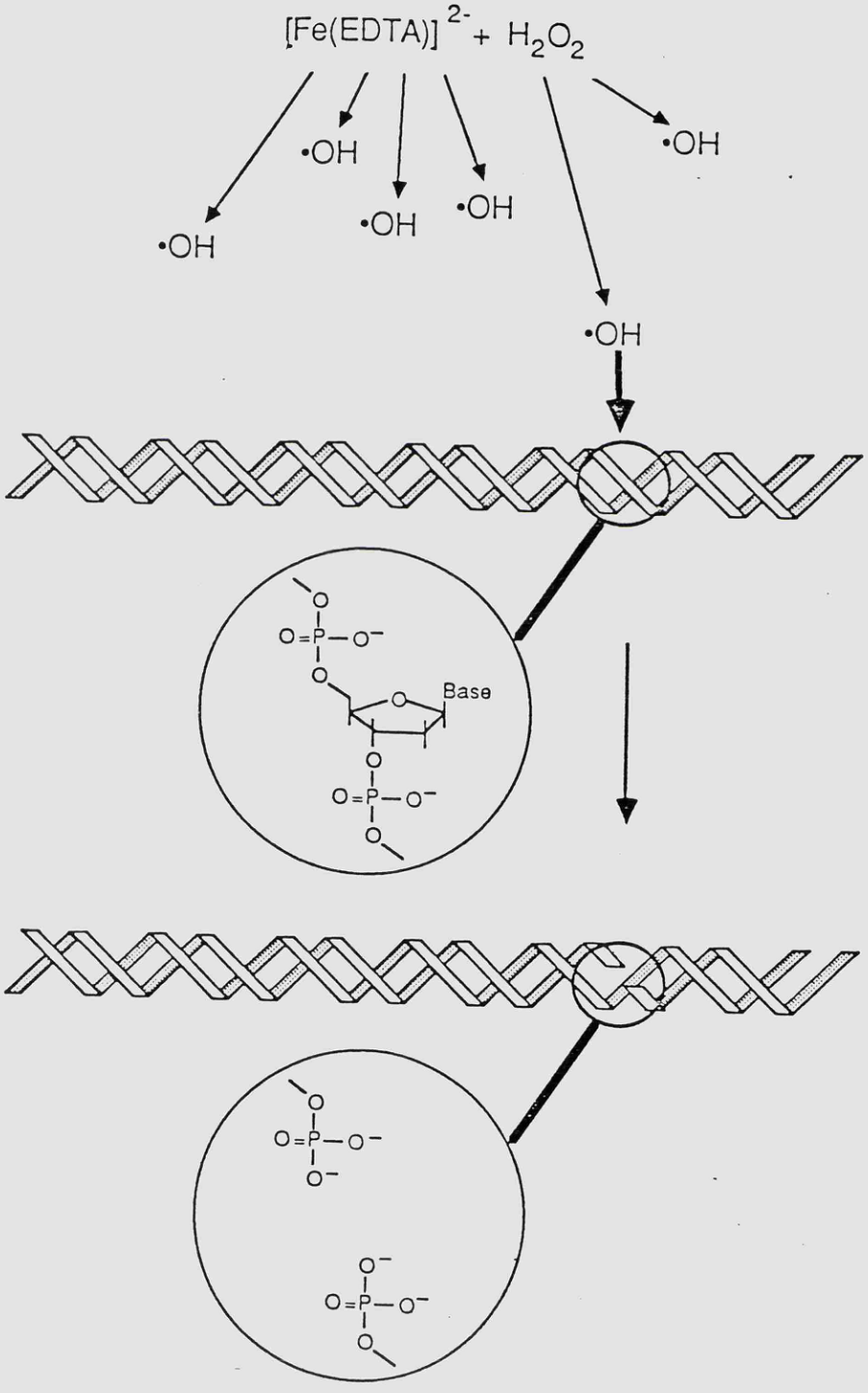
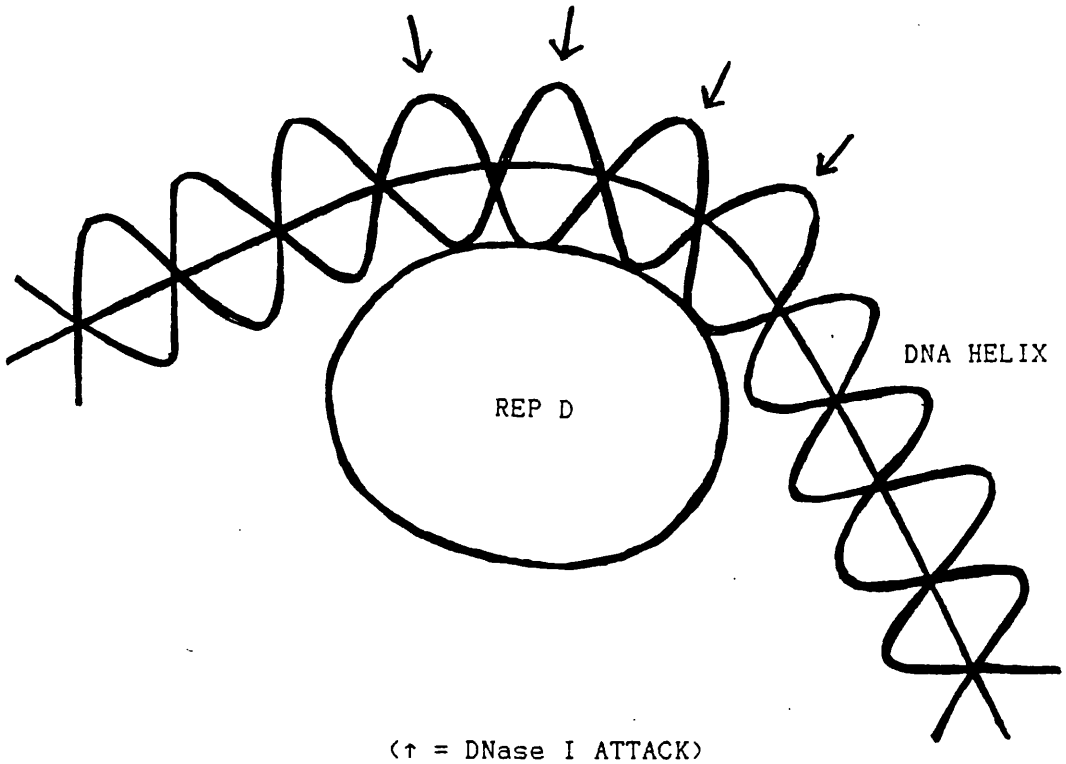

[Fe(EDTA)]2- + H2O2
•OH
•OH
•OH
•OH
•OH
•OH
Base

Fig 5.4: Proposed binding of REP D to the oriD sequence.

The pattern of enhanced cleavage observed in the DNase I footprint of REP D suggests that the protein only makes close contact with one face of the DNA helix of oriD. Bending of the DNA on binding of REP D would open up the minor groove of the DNA on the side of the DNA helix facing away from the protein, making it more accessible to attack by DNase I, denoted by the arrows.



(↑ = DNase I ATTACK)

CHAPTER 6.

In vitro mutagenesis of ICR II of the pC221 origin.

6.1. - Introduction.

6.2. - Results and discussion.

6.2.1. - *In vitro* mutagenesis of pC221 oriD by the method of Kunkel.

6.2.2. - Assay of mutant origin sequences using the REP D *in vitro* nicking /closing activity.

CHAPTER 6.

IN VITRO MUTAGENESIS OF ICR II OF THE pC221 ORIGIN.

6.1: - Introduction.

The pC221 origin (oriD) has three inverted complementary repeat sequences (ICR I, II and III) which may be capable of intra-strand secondary structure formation, e.g. hairpin or cloverleaf structures (see Figure 3.7). However, it is not known whether such structures would be stable *in vivo* or whether they are important for pC221 replication. Work carried out with pC221 *in vitro* (see chapter 4), suggests that, after sequence-specific binding of REP D, a conformational change may be required within oriD before nicking can occur. Such a change appears to be stabilised either by sequence-specific binding of REP D at the origin, or by negative supercoiling of the DNA.

DNase I footprinting data has shown that REP D binding stretches over the 3' half of ICR II, as well as the whole of ICR III. As ICR II is almost completely conserved in all known rep sequences, it may be that the change in the DNA takes place here. Hairpin extrusion at this hairpin would be stabilised by the energy stored in negative supercoiling of the DNA. The proposed hairpin may also be stabilised by protein /DNA contacts with bound REP D.

The above observations raise the question whether mutations within ICR II, ones in particular which affect base pairing within the stem of the proposed hairpin, might affect the ability of REP D to nick at oriD. It was decided to introduce mutations on each side of ICR II. Either mutation would

promote instability of the proposed hairpin stem by virtue of the loss of a hydrogen bonded base pair. This would also determine whether the sequence on the 3' side of ICR II, where REP D appears to bind, is more critical for REP D binding and /or nicking than that on the 5' side of ICR II.

6.2: - Results and discussion.

6.2.1: - *In vitro* mutagenesis of oriD by the method of Kunkel.

A TaqI fragment (278bp) of pC221 containing oriD was cloned into M13 via the AccI site (Figure 6.1). This construct (M13 /oriD) was transformed into *E. coli* JM101. Phage were picked and used to transfect *E. coli* RZ1032 in the presence of deoxyuridine. Single-stranded phage DNA, with deoxyuridine incorporated, was used for *in vitro* mutagenesis (Kunkel, 1985).

Two mutagenic oligonucleotides were designed in order, separately, to change a single base on either side of ICR II (Figure 6.2). The changes were a G to C on the 5' side of ICR II and a C to G on the 3' side of ICR II. Both give disruption of one C-G base pair in the stem of the proposed hairpin. The changes also destroyed each of the HpaII sites present on either side of ICR II.

Mutagenesis was carried out by the method of Kunkel (Figure 6.3). Plaques were picked into phage buffer and single-stranded phage DNA was prepared. Dideoxy sequencing using the single-stranded DNA with the M13 universal primer was carried out to compare the wild type and mutant origin sequences (Figure 6.4).

6.2.2: - Assay of mutant origin sequences using REP D *in vitro* nicking /closing activity.

Double-stranded replicative forms of M13, M13 /oriD wild type and the two mutants, M13 /oriD(5') and M13 /oriD(3') were prepared (5' and 3' denote

mutations in the 5' and 3' sides of ICR II respectively). The supercoiled forms were isolated by CsCl /ethidium bromide density gradient centrifugation. Double-stranded, supercoiled DNA was assayed by its ability to support the *in vitro* nicking /closing activity of REP D (Figure 6.5). The DNA was treated with increasing molar ratios of REP D for 30 minutes at 30°c and the products were analysed by agarose gel electrophoresis in the presence of ethidium bromide.

As expected, M13 DNA was not nicked by REP D. However, M13 containing the wild type cloned origin sequence (M13 /oriD) was nicked by REP D. Nicking was complete at an approximate ratio of 2:1 REP D monomers to DNA molecules, i.e. one REP D dimer per DNA molecule. The same result has been obtained with pC221 DNA (Thomas, 1988). Most of the nicked DNA was religated to give the relaxed, covalently closed circular form.

The extent of nicking obtained with M13 /oriD(5') appeared to be exactly the same as that obtained with wild type M13 /oriD, although this assay may not be sensitive enough to detect very small changes in nicking activity. This result suggested that the mutated G base on the 5' side of ICR II is not important for recognition and nicking by REP D, as might be expected from footprinting data. It also suggests that hairpin formation at ICR II is not important for REP D binding and nicking, as the disruption of a G-C base pair in the 9bp stem of the proposed hairpin would be expected to have some effect on its stability.

A method for estimation of the stability of secondary structure within nucleic acids based on the measured free energies of base paired regions *in vitro* (Tinoco et al, 1973) has been proposed for the prediction of the stability of hairpin structures. The estimated free energy of the putative

hairpin formed in wild type ICR II is -19.6kcal. This increases to -11.4kcal in the mutated hairpin with a single mismatch. Disruption of a C-G base pair would be expected to have a greater effect than disruption of an A-T base pair. Also, work carried out by Lilley (1982) suggests that purine /purine and pyrimidine /pyrimidine mismatches such as these have a much greater destabilising effect than purine /pyrimidine mismatches.

It should be noted that hairpin structures are in competition with perfectly base paired duplexes and even a relatively minor asymmetry, such as a single mismatch, may be sufficient to raise the free energy of the hairpin enough to lower its stability relative to that of the linear duplex conformation. However, such a mismatch may not be sufficient to prevent hairpin formation, especially in the presence of negative DNA supercoiling and bound REP D protein which may have a stabilising effect on the hairpin. The free energy of the mutated hairpin is still negative, albeit much higher than that of the wild type hairpin, suggesting that the mutant hairpin may still form. This result is therefore inconclusive as to whether hairpin extrusion at ICR II is required for nicking of oriD by REP D.

The extent of nicking observed with M13 /oriD(3') was much less than that observed with wild type M13 /oriD and M13 /oriD(5'). Nicking was only complete with ratios of REP D to DNA greater than 16:1 (i.e. 8 REP D dimers per DNA molecule). This reflects an 8 fold decrease in recognition and /or nicking of this mutant origin by REP D compared to the wild type origin. This suggests that the 3' side of ICR II is more important for REP D binding and /or nicking than the 5' side, consistent with the footprinting data.

However, the 5' side of ICR II is almost perfectly conserved in all six known rep sequences, except for one base change in pS194 and pUB112, suggesting that this inverted repeat sequence is important in some aspect of pC221 replication. It may be that the destabilising effect of the mismatched hairpin and the disruption of REP D binding both contribute to the 8 fold decrease in REP D nicking observed. Alternatively, the result may be solely due to disruption of REP D binding, as the base change lies within the REP D binding site, as determined by DNase I footprinting studies.

The pattern of nicking of the M13 /oriD(3') mutant origin also differs from that of the wild type M13 /oriD and M13 /oriD(5') origins in that very little of the nicked product is religated to give the relaxed, covalently closed circular form. It may be that REP D binding to this mutant origin is impaired so that, after nicking, the protein dissociates from its normal binding site before religation can take place. Alternatively, binding may be disrupted so that the REP D /oriD complex is in an altered configuration which allows nicking to occur, albeit at reduced efficiency, but is no longer able to promote religation.

Another interpretation of this observed result is that the DNA of this mutant origin may be nicked and religated, but the relaxed, covalently closed circular form of the DNA is more susceptible to further attack by REP D to give the open circular form than the wild type. As the substrate DNA is already relaxed, the second nicking event would yield open circular DNA, as there is no energy contribution from DNA supercoiling towards the religation reaction. This explanation, however, seems unlikely, as the supercoiled form of

this mutant origin has already been shown to be less susceptible to nicking by REP D than the wild type.

Fig 6.1: Cloning of oriD into M13.

A 278bp TaqI restriction fragment of pC221 containing oriD was cloned into M13 /mp19 via the AccI site. The size of the inserts were checked by restriction endonuclease digestion of the double-stranded replicative forms and dideoxy sequencing using the M13 universal primer on the single-stranded templates.

Single-stranded DNA with deoxyuridine incorporated was prepared to allow *in vitro* mutagenesis of oriD by the method of Kunkel (1985).

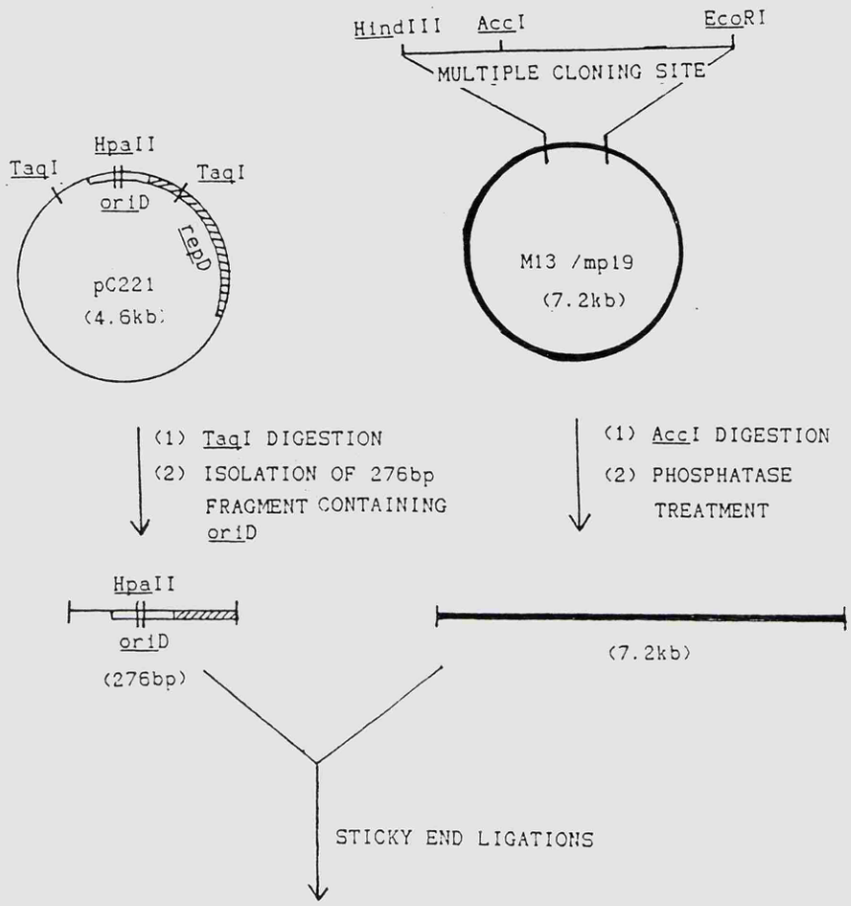

HindIII
AccI
EcoRI
MULTIPLE CLONING SITE
HpaII
TaqI
TaqI
oriD
repD
pC221
(4.6kb)
M13 /mp19
(7.2kb)
(1) TaqI DIGESTION
(2) ISOLATION OF 276bp FRAGMENT CONTAINING oriD
(1) AccI DIGESTION
(2) PHOSPHATASE TREATMENT
HpaII
oriD
(276bp)
(7.2kb)
STICKY END LIGATIONS
HpaII
oriD
OR
HpaII
oriD
M13 /oriD
(7.5kb)

Fig 6.2: Mutations introduced into the 5' and 3' sides of ICR II of oriD.

The sequences of the two mutations introduced into ICR II of oriD (M13 /oriD(5') and M13 /oriD(3'), denoting changes in the 5' and 3' sides of ICR II respectively, are shown compared to the wild type sequence. These mutations disrupt a single C-G base pair in the stem of the proposed hairpin formed at ICR II in each case.

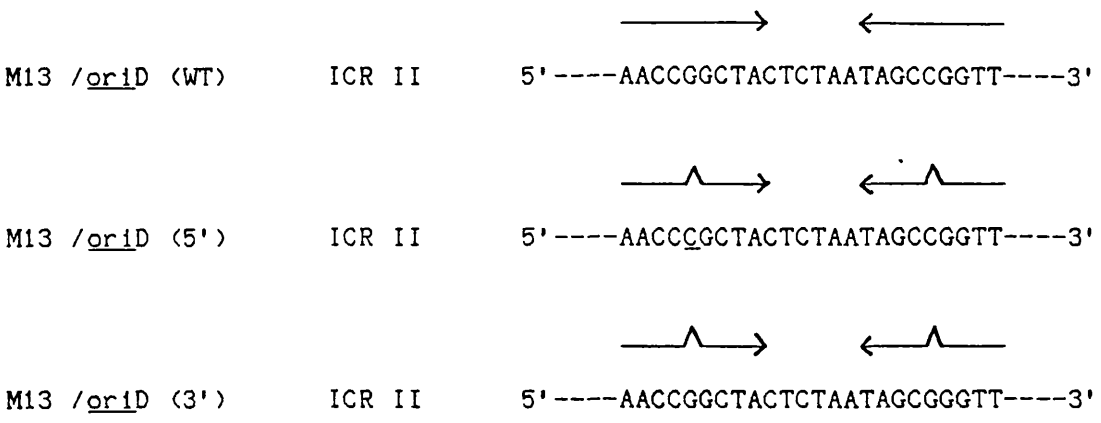

M13 /oriD (WT) ICR II 5'----AACCGGCTACTCTAATAGCCGGTT----3'
M13 /oriD (5') ICR II 5'----AACCCGCTACTCTAATAGCCGGTT----3'
M13 /oriD (3') ICR II 5'----AACCGGCTACTCTAATAGCGGGTT----3'

Fig 6.3: *In vitro* mutagenesis of oriD by the method of Kunkel.

A fragment of pC221 containing oriD was cloned into M13 and single-stranded phage DNA (containing deoxyuridine) was prepared. The appropriate mutagenic oligonucleotide was annealed to the single-stranded template and extension /ligation reactions were carried out in the presence of dNTPs, Klenow fragment of DNA polymerase I and T_4 DNA ligase. Dilutions of the extension /ligation reaction mixtures were transformed into *E. coli* JM101. Plaques were picked and single-stranded phage DNA prepared. Dideoxy sequencing was carried out with these single-stranded templates and the M13 universal primer to check for the required mutations.

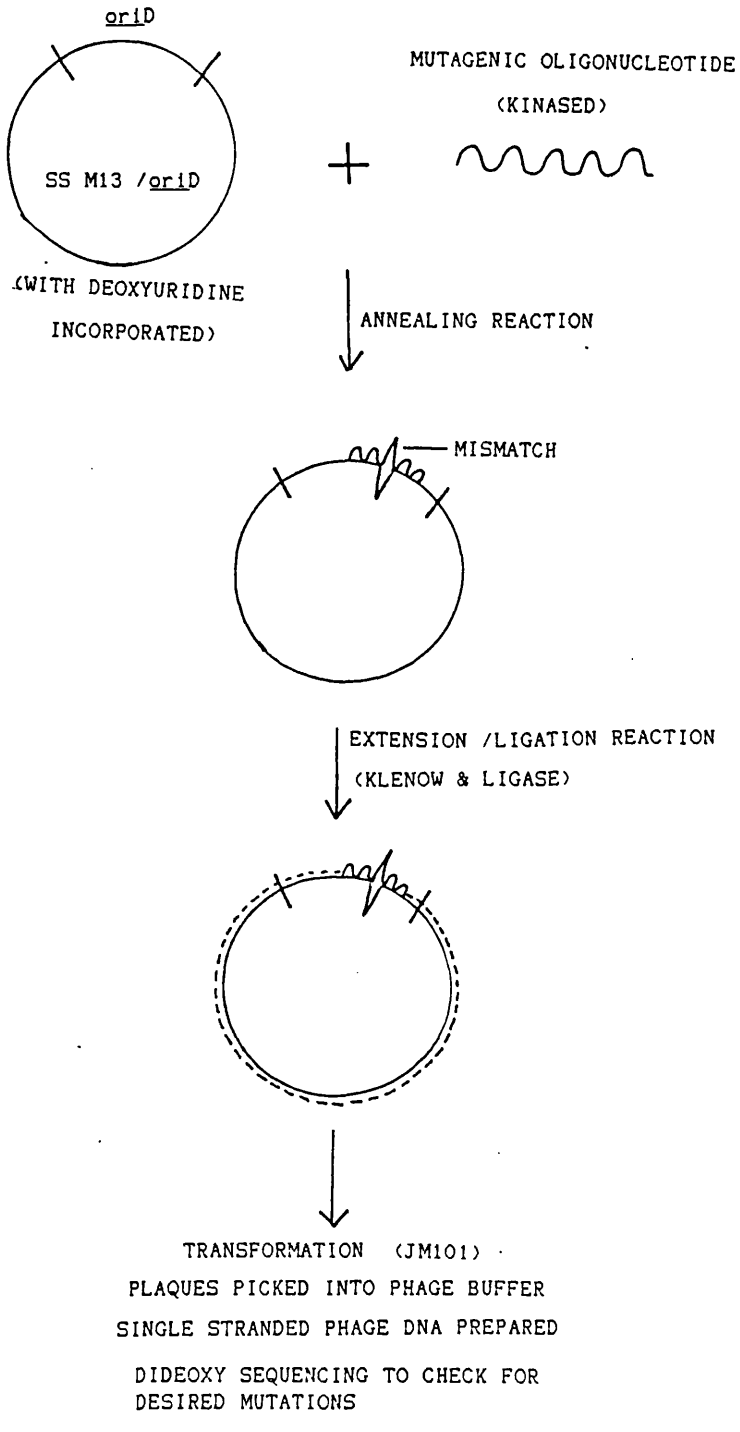

oriD
SS M13 /oriD
(WITH DEOXYURIDINE
INCORPORATED)
MUTAGENIC OLIGONUCLEOTIDE
(KINASED)
ANNEALING REACTION
MISMATCH
EXTENSION /LIGATION REACTION
(KLENOW & LIGASE)
TRANSFORMATION (JM101)
PLAQUES PICKED INTO PHAGE BUFFER
SINGLE STRANDED PHAGE DNA PREPARED
DIDEOXY SEQUENCING TO CHECK FOR
DESIRED MUTATIONS

Fig 6.4: Dideoxy sequencing of wild type and mutant oriD sequences.

The dideoxy sequencing tracks of the wild type and mutant oriD sequences are shown. The sequence of the ICR II region is shown beside the sequencing tracks in each case. The sequencing tracks are in the order ACGT. The sequence of the wild type origin is ambiguous in that it is difficult to tell whether the second HpaII site has lost a G residue. However, treatment of this cloned origin with REP D gives complete relaxation with a REP D dimer /DNA of 1:1, as seen with wild type pC221 DNA (Thomas, 1988). Also, M13 /oriD (5') and (3'), which were derived from the wild type sequence, obviously have this G residue. Therefore, this apparent deletion of a G in the wild type sequence is probably due to false sequencing or to anomalous running of the gel.

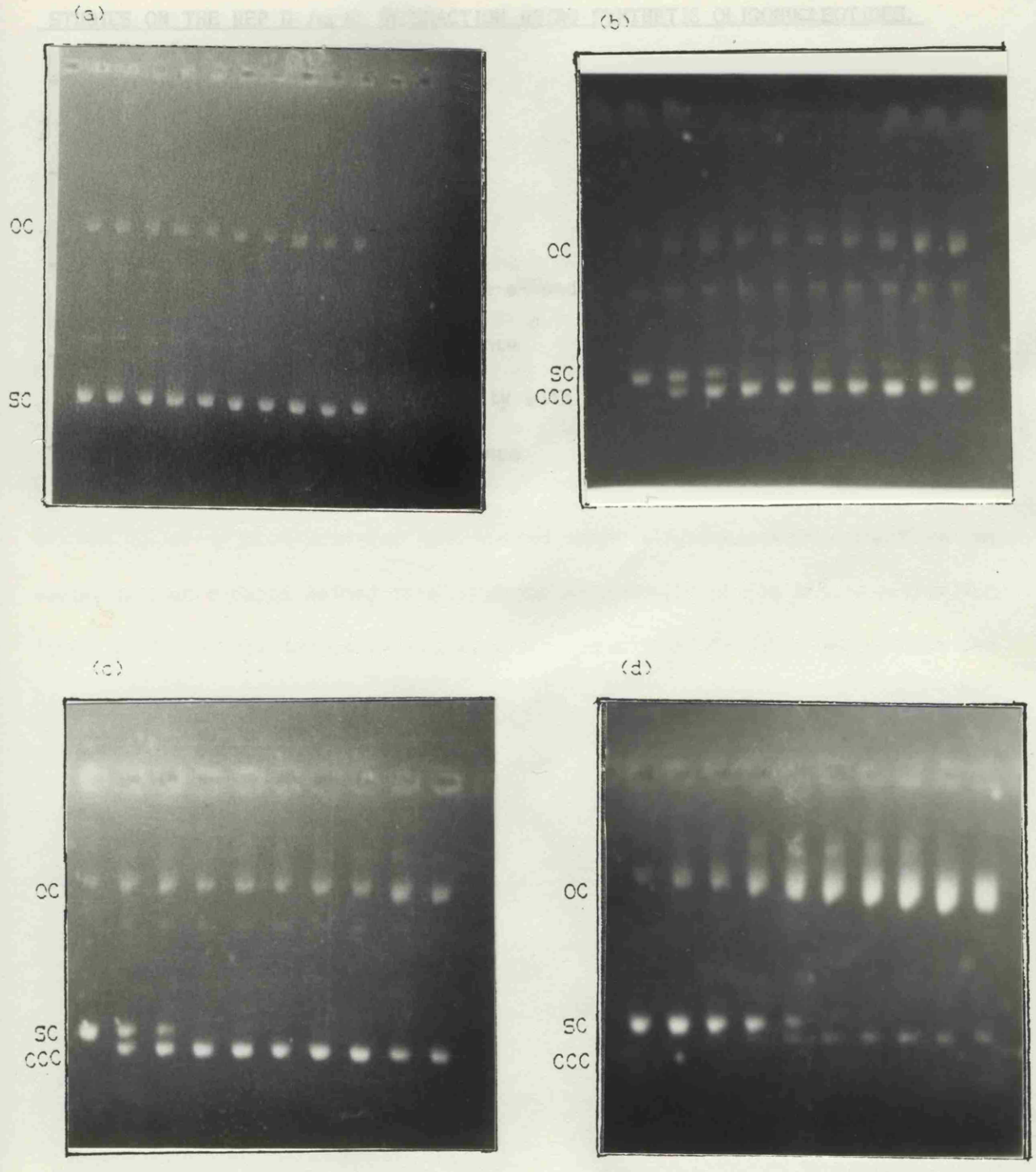
M13 /oriD (WT)

M13 /oriD (5')

M13 /oriD (3')

Fig 6.5: *In vitro* nicking /closing activities of REP D with wild type and mutant oriD sequences cloned into M13.

Double-stranded, supercoiled forms of M13, M13 /oriD, M13 /oriD(5') and M13 /oriD(3') were exposed to increasing molar ratios of REP D at 30°c for 30 minutes. The molar ratios of REP D to DNA used were 0:1, ½:1, 1:1, 2:1, 4:1, 8:1, 16:1, 32:1, 64:1 and 128:1 respectively. Products were analysed by agarose gel electrophoresis in the presence of ethidium bromide. Gel [a] shows that REP D has no detectable nicking /closing activity with M13. Gel [b] shows the nicking /closing activity of REP D with wild type M13 /oriD. Nicking is complete at a ratio of one REP D dimer per DNA molecule. Gel [c] shows the nicking /closing activity of REP D with M13 /oriD(5'), which appears to be exactly the same as with the wild type origin sequence. Gel [d], however, shows that the extent of nicking by REP D with the M13 /oriD(3') origin is approximately 8 fold lower than with the wild type origin. Most of the product here is open circular rather than relaxed, covalently closed circular form.



SC = SUPERCOILED
CCC = COVALENTLY CLOSED CIRCULAR
OC = OPEN CIRCULAR

CHAPTER 7.

STUDIES ON THE REP D /oriD INTERACTION USING SYNTHETIC OLIGONUCLEOTIDES.

7.1. - Introduction.

7.2. - Results and discussion.

7.2.1. - The design of single-stranded oligonucleotides based on the oriD ICR II sequence.

7.2.2. - REP D nicking activity with oligonucleotides based on the oriD ICR II sequence.

CHAPTER 7.

STUDIES ON THE REP D /oriD INTERACTION USING SYNTHETIC OLIGONUCLEOTIDES.

7.1: - Introduction.

Experiments have been carried out using synthetic oligonucleotides based on the oriD sequence to elucidate the REP D amino acid residue to which the DNA becomes covalently attached after the nicking event (Thomas 1988). The results showed that REP D can cleave and become covalently attached to a 24 base single-stranded oligonucleotide corresponding to the upper strand sequence of ICR II of oriD. Therefore, the use of such oligonucleotides might be an easier and more rapid method than *in vitro* mutagenesis of the entire origin for investigation of the effect of sequence changes within ICR II on recognition and nicking by REP D.

7.2: - Results and discussion.

7.2.1: - The design of single-stranded oligonucleotides based on the oriD ICR II sequence.

Oligonucletides were designed based on the sequence of the upper strand of ICR II of oriD. Ten oligonucleotides were synthesized (Figure 7.1). An oligonucleotide corresponding to the "wild type" ICR II sequence (DFB 20) was synthesized to act as a control. Six oligonucleotides were used to investigate the importance of the bases on both sides of the nick site for cleavage by REP D (DFB 21-25). Three more oligonucleotides were used to investigate the effect of changing the sequence on either side of the putative stem structure of the proposed ICR II hairpin, but keeping the sequence of the loop region constant (DFB 27-29).

7.2.2: - REP D nicking activity with oligonucleotides based on the oriD ICR II sequence.

Oligonucleotides were purified by polyacrylamide gel electrophoresis, elution from a gel slice and passage through a Sephadex G-25 spun column. They were then 5' end-labelled with $\gamma[^{32}P]$ ATP using T_4 kinase. Size markers, consisting of a mixture of oligonucleotides betweeen 10 and 20 bases in length, were also end-labelled.

End-labelled ICR II oligonucleotides were treated with a 2:1 molar ratio of REP D to DNA (i.e. one REP D dimer per DNA molecule) for 10 minutes at

30°c. The products of the reaction were subjected to polyacrylamide gel electrophoresis and autoradiography (Figure 7.2).

If REP D cleaves the oligonucleotide at the specific nick site, the product should be a 13 base end-labelled oligonucleotide, with REP D presumably becoming covalently attached to the other (unlabelled) 11 base oligonucleotide. The oligonucleotide corresponding to the "wild type" ICR II sequence (DFB 20) yielded only the expected 13 base end-labelled product on treatment with REP D (Figure 7.2. - Track 3), showing that cleavage was only taking place at the expected site. This suggests that the absence of much of the REP D binding site (i.e. ICR III), as determined by DNA footprinting experiments, has no effect on the position of cleavage of this single-stranded oligonucleotide. However, this does not appear to hold true with double-stranded DNA. The minimum origin sequences determined for the pT181 "family" of plasmids by deletion analysis include ICR II and III (Gennaro et al, in preparation).

The six oligonucleotides with bases on either side of the nick site altered in all possible combinations (DFB 21-26) showed very little detectable cleavage by REP D (Figure 7.3. - Tracks 4-16). This suggests that the bases on both sides of the nick site are important for cleavage by REP D. Some very low level nicking was obtained when the A residue on the 3' side of the nick site was changed, but not when the T residue on the 5' side of the nick site was changed, suggesting that the nucleotide on the 5' side of the nick site is more critical for nicking by REP D than that on the 3' side. This result is unexpected, as footprinting data suggested that the ori sequences on the 3' side of the nick site are generally more important for REP D binding than those on the 5' side. It is not possible from this experiment to discern whether it is

simply the cleavage reaction which is affected, or whether REP D binding is also affected.

Three oligonucleotides were designed to examine the importance of the sequence on either side of the stem of the proposed ICR II hairpin (DFB 27-29). DFB 27 has the sequence of the 3' side of ICR II jumbled so as to give altered sequence but the same base composition. This would totally destroy intra-strand base pairing to form a hairpin structure. DFB 28 has the sequence on the 5' side of ICR II jumbled. DFB 29 has the sequence on both sides of ICR II jumbled to give altered sequence, but to restore complementarity between the two sides of ICR II. Of these three oligonucleotides, only DFB 27 was cleaved by REP D with the same efficiency as the "wild type" oligonucleotide. There was also a minor labelled product corresponding to cleavage of this oligonucleotide two bases upstream of the usual nick site, i.e. cleavage appears to take place, at low efficiency, after the first CT dinucleotide in the putative loop region, as well as after the second CT dinucleotide. DFB 28 was still cleaved, but with much lower efficiency. DFB 29 gave very little detectable cleavage with REP D.

These results suggested that the sequence on the 3' side of ICR II is more critical for recognition and or cleavage by REP D than is the 5' side. This is consistent with the DNase I footprinting data which showed that REP D binds to the 3' side of ICR II, upstream of the nick site, and the whole of ICR III (see chapter 5). It is also compatible with the possibility that hairpin formation is not required for the cleavage of these single-stranded oligonucleotides *in vitro*.

Fig 7.1: Synthetic oligonucleotides designed to investigate the interaction between REP D and ICR II of ori D.

Oligonucleotides DFB 20-29 were designed to investigate the interaction between REP D and ICR II of oriD. DFB 20 corresponds to the "wild type" ICR II sequence. DFB 21-27 have the two bases on either side of the nick site changed to each of the other three bases in turn. DFB 27 and 28 have the bases on the 5' and 3' sides of ICR II, respectively, jumbled in order to destroy intra-strand base pairing to give a hairpin. DFB 29 has the bases on both sides of ICR II jumbled so as to give an altered sequence, but restore the possibility of intra-strand base pairing to give a hairpin. The oligonucleotides DFB 27-29 all have the "wild type" sequence within the proposed loop region.

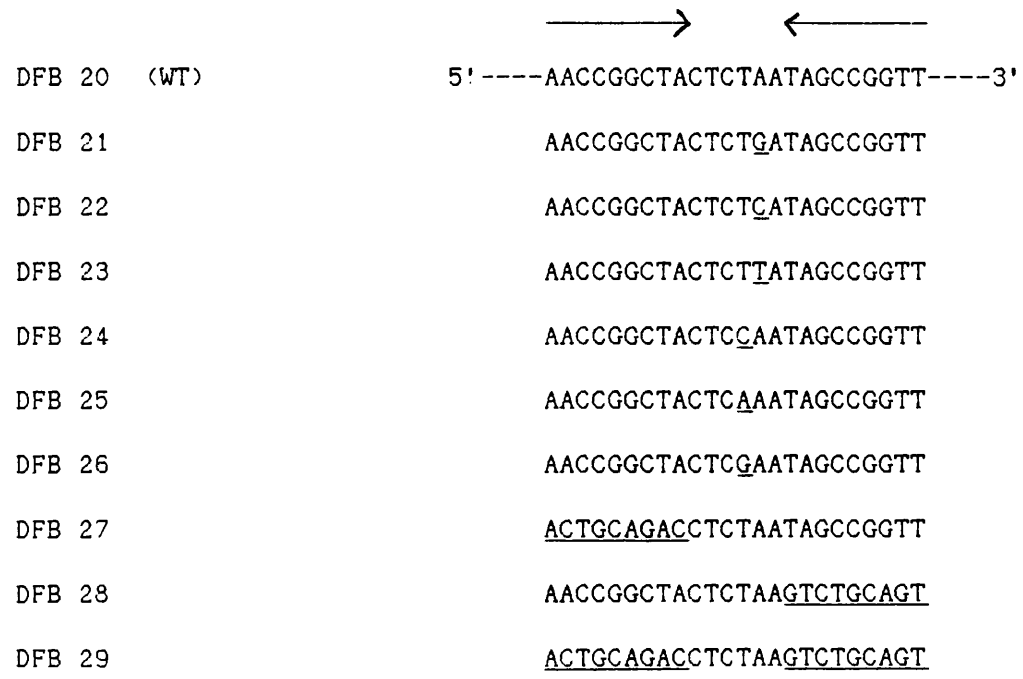
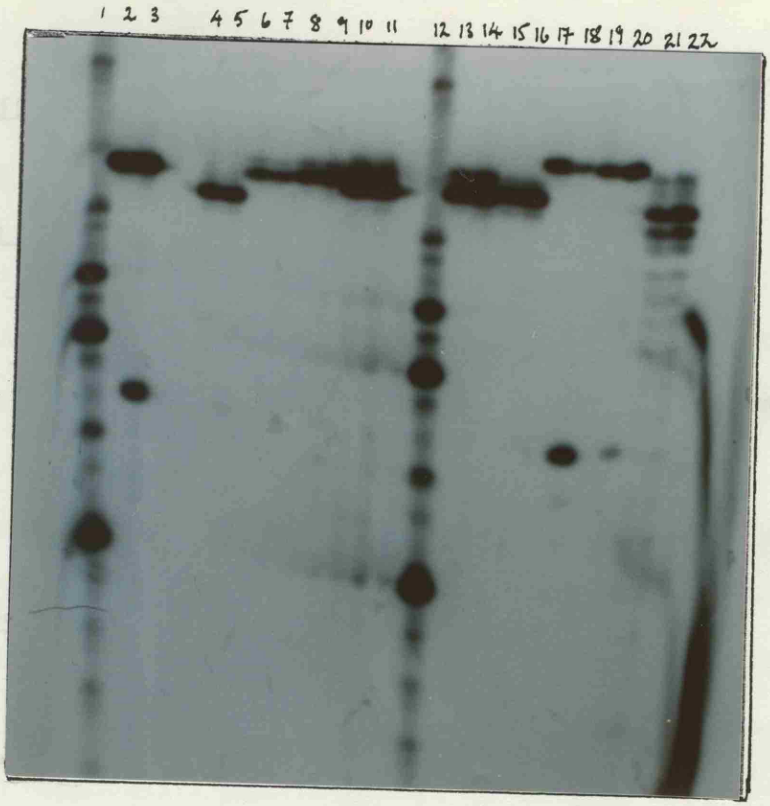

DFB 20 (WT) 5'----AACCGGCTACTCTAATAGCCGGTT----3'
DFB 21 AACCGGCTACTCTGATAGCCGGTT
DFB 22 AACCGGCTACTCTCATAGCCGGTT
DFB 23 AACCGGCTACTCTTATAGCCGGTT
DFB 24 AACCGGCTACTCCAATAGCCGGTT
DFB 25 AACCGGCTACTCAAATAGCCGGTT
DFB 26 AACCGGCTACTCGAATAGCCGGTT
DFB 27 ACTGCAGACCTCTAATAGCCGGTT
DFB 28 AACCGGCTACTCTAAGTCTGCAGT
DFB 29 ACTGCAGACCTCTAAGTCTGCAGT

Fig 7.2: *In vitro* nicking activity of REP D with oligonucleotides based on the oriD ICR II sequence.

The autoradiograph shows the effect of REP D on oligonucleotides based on the oriD ICR II upper strand sequence. 5' End-labelled oligonucleotides (0.01-0.02pmols) were treated with a 2:1 molar ratio of REP D to DNA for 10 minutes at 30°c. The products of the reaction were subjected to polyacrylamide gel electrophoresis and autoradiography. Tracks 1 and 12 show size markers. Track 2 shows the "wild type" ICR II oligonucleotide (DFB 20). Treatment of DFB 20 with REP D yielded the expected 13 base product (Track 3). Oligonucleotides DFB 21 (Tracks 4 & 5), DFB 22 (Tracks 6 & 7), DFB 23 (Tracks 8 & 9), DFB 24 (Tracks 10 & 11), DFB 25 (Tracks 13 & 14) and DFB 26 (Tracks 15 & 16) have the bases on both sides of the nick site altered to each of the other three bases in turn (see figure 7.1). There was very little detectable cleavage of these oligonucleotides by REP D.

DFB 27 (Tracks 17 & 18) and DFB 28 (Tracks 19 & 20) have the 5' and 3' sides of ICR II, respectively, jumbled to destroy putative intra-strand base pairing. DFB 27 was cleaved by REP D with the same efficiency as the "wild type" oligonucleotide. DFB 28 was cleaved with much lower efficiency. DFB 29 (Tracks 21 & 22), which has both sides of ICR II jumbled so as to give altered sequence but restore intra-strand base pairing, was not cleaved by REP D. The results suggested that the 3' side of ICR II is more important for nicking by REP D than the 5' side. It would also appear that the formation of a hairpin at ICR II is not required for cleavage of these oligonucleotides by REP D.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

CHAPTER 8.

DISCUSSION.

8.1. - The specificity of REP D for pC221 and related plasmids *in vitro*.

8.2. - The REP D /oriD interaction *in vitro*.

8.2.1. - oriD may be bent on binding of REP D.

8.2.2. - *In vitro* mutagenesis of the pC221 origin.

8.2.3. - Use of synthetic oligonucleotides to investigate the REP D / oriD interaction.

8.3. - X-ray crystallographic studies of the REP D /oriD complex.

8.4. - Summary.

CHAPTER 8.

DISCUSSION.

8.1: - The specificity of REP D for pC221 and related plasmids *in vitro*.

Initiation of pC221 replication will occur *in vitro* in cell free extracts of *S. aureus* in the presence of REP D, deoxyribonucleotide triphosphates, Mg^{++} ions and monovalent cations (see figure 4.1) . When the substrate plasmid DNA is in the negatively supercoiled form, REP D can also initiate replication *in vitro* with members of the pT181 "family" of plasmids (see figure 4.7), but not with other, unrelated plasmids (see figure 4.9). However, when the plasmids are in the relaxed, covalently closed circular form and the assays are carried out in the presence of novobiocin, the specificity of REP D for pC221, its cognate plasmid, appears to increase relative to that for pT181 (see figure 4.11). Further experiments should, however, be carried out with other relaxed plasmids in order to determine whether REP D is absolutely specific for pC221 *in vitro* with relaxed DNA. Although such results have not been described for other REP proteins, it may be a general property of the rep system. Alternatively, other REP proteins may show greater specificity for their cognate origins *in vitro*.

If REP D is only absolutely specific for pC221 *in vitro* with relaxed plasmid DNA, an explanation for this observation might be that a conformational change takes place within ori upon sequence-specific binding of REP at ICR III, which may then be recognised and nicked by the REP active site. However, were such a conformational change to be favoured or stabilised by negative

supercoiling of the DNA, then the change could take place even in the absence of bound REP. It may be that REP could then recognise and nick the origin without sequence-specific recognition. This would also explain the observed 50% inhibition of pC221 by novobiocin, as supercoiling of the DNA by DNA gyrase might favour the conformational change in the presence of bound REP D, thus facilitating the nicking event.

The possibility of hairpin formation at the three inverted repeat sequences within the origin has already been discussed. Such a hairpin would be stabilised by negative supercoiling of the DNA (Lilley, 1980). Therefore, hairpin formation at any of these inverted repeat sequences might be important for the initiation of replication. However, hairpin formation at ICR I or III would seem unlikely, as these hairpins would have loops of only one base in some of these plasmids, which is thought to be energetically unfavourable (Tinoco *et al*, 1971).

ICR II would seem the most likely to form a hairpin important for initiation of replication. ICR II is the most highly conserved of the three inverted repeat sequences. Nicking takes place within the loop region of the proposed hairpin at ICR II. DNA footprinting data shows that REP binds over the 3' side of ICR II, as well as the whole of ICR III, so that hairpin formation at ICR II might be stabilised by protein /DNA contacts. Finally, the estimated stability of the proposed hairpin at ICR II (Tinoco *et al*, 1973) is greater than the estimated stabilities of the hairpins which might form at the other two inverted repeat sequences (Figure 8.1).

The proposed mechanism for the nicking of oriD by REP D is shown in Figure 8.2. A hairpin structure extruded at ICR II on sequence-specific binding

of REP D may then be recognised and nicked by the REP D active site, within the proposed loop region.

8.2: - The REP D /oriD interaction *in vitro*.

8.2.1: - oriD DNA may be bent on binding of REP D.

DNase I footprinting data with REP D on oriD showed that REP D binds to oriD over ICR III and the 3' side of ICR II. A pattern of enhanced cleavage by DNase I of the DNA bound to REP D was observed. This pattern had a periodicity of approximately 10 bases and was displaced by 5 bases on each strand relative to the other (see figure 5.2). This is consistent with the REP D protein making contact with only one face of the DNA helix. Bending of the DNA around REP D would open up the minor groove of the DNA helix, giving greater accessibility to DNase I cleavage on the side of the DNA facing away from the protein (see Figure 5.4).

It has been reported that pT181 DNA contains a bend at the origin of replication which is enhanced by the binding of REP C (Koepsel *et al*, 1986b) (Figure 8.3). The enhanced bend at oriC on binding of REP C was found to be approximately over the region of ICR III, which would correspond to the proposed induced bend in oriD from DNase I footprinting data with REP D.

It may be that bending of the DNA on binding of REP D is the conformational change required at the origin sequence for nicking to take place. The energy stored in negative supercoiling of the DNA might stabilise such induced bending of the DNA by providing the required energy, thereby increasing the affinity of REP D for the origin sequence. This increased affinity for the origin in the supercoiled state might then account for the apparent loss of specificty of REP D with supercoiled DNA.

DNA bending, either sequence-directed or as a consequence of protein binding, has recently been suggested as a factor in the regulation of gene expression, sequence-specific recognition and DNA packaging, as well as initiation of DNA replication. DNA bending can be detected by altered mobility of the DNA on a polyacrylamide gel (Crothers, 1987). Sequence-directed DNA bending has been found in the bending locus of trypanosome kinetoplast DNA which has tracts of a simple repeat sequence ($CA_{5-6}T$) symmetrically distributed about it with a repeat interval of 10bp (Wu & Crothers, 1984). Bending has also been found to be induced in the promoter of the *E. coli* lac operon on binding of CAP, the catabolite activator protein (Lui-Johnson *et al*, 1986). Sites of sequence-induced DNA curvature have also been found at the origin of replication of bacteriophage λ (Zahn *et al*, 1985) and at the origin region 1 of SV40 (Ryder *et al*, 1986). Bending of the origin of plasmid R6K on binding of the replication initiator protein has also been reported (Mukherjee *et al*, 1985). Thus, it would appear that DNA bending may play an important role in many different aspects of protein /DNA interactions.

It is possible that the role of DNA bending is primarily to create protein /DNA interactions which would not otherwise be sterically possible if the DNA were in the linear form. However, in some cases, there is evidence for DNA bending having a functional role. For example, the bending of the lac promoter by CAP may create a DNA /protein complex which can then be recognised by RNA polymerase (Figure 8.4). Also, significant energy is stored in the induced bend which can be used in transcription, such as local opening of the DNA double helix. DNA bending induced in the origin of plasmid R6K by binding of the replication initiator protein may lead to unwinding of the adjacent AT rich region due to the torsional stress generated.

It must be remembered that the DNase I footprinting pattern observed here corresponds to the REP D /oriD complex after nicking has taken place. Therefore, it may be that other regions of the oriD DNA are involved in the initial interaction. A conformational change might then take place after the nicking reaction has occurred, thereby altering the interaction between protein and DNA. Further footprinting experiments are planned to investigate the interaction between REP D and oriD in the absence of nicking. Mg^{++} ions are required for the REP D nicking reaction. In the absence of Mg^{++} ions, REP D can still bind oriD but can no longer nick it. However, DNase I requires Ca^{++} or Mg^{++} ions in order to cleave DNA. Therefore, as REP D can also nick, but not religate, oriD in the presence of Ca^{++} (Thomas, 1988) another footprinting technique would have to be employed. Also, the absence of Mg^{++} ions may seriously disrupt the binding of REP D to oriD.

The mutant REP D protein, Y188F, which can bind oriD but can no longer cleave it, might also be employed to investigate the REP D /oriD complex in the absence of nicking. However, again it must be assumed that it is only the nicking reaction that is affected and that binding itself is unimpaired. This would seem to be a reasonable assumption for both binding of REP D to oriD in the absence of Mg^{++} and binding of the mutant REP D protein to oriD, as the extent of binding, as determined by gel retardation assays, appears to be unaffected (Thomas, 1988).

8.2.2. - *In vitro* mutagenesis of the pC221 origin.

Since it seemed likely that hairpin formation at ICR II might be important for nicking of the origin by REP, two mutations were introduced into

ICR II to destroy a single G-C base pair in the stem of the proposed hairpin, one in the 3' side of ICR II and the other in the 5' side (see figure 6.2). The mutation in the 5' side of ICR II appeared to have no effect on *in vitro* nicking and closing by REP D (see figure 6.5), suggesting that hairpin formation at ICR II is not absolutely required for REP D recognition and /or cleavage, as the mutation would destabilise the hairpin. However, estimation of the stability of this mutant hairpin (Tinoco *et al*, 1973) showed that its free energy was still negative. Also, the recently published sequences of pS194 and pUB112 (Projan *et al*, in press) show that ICR II of these plasmids has a natural mismatch of 1bp in the stem of the proposed hairpin. In each case, the base change occurs in the 5' side of ICR II, again suggesting that the sequence of the 3' side of ICR II is more critical for recognition and /or cleavage by REP D.

It is clear that further experiments must be carried out in order to probe the 5' side of ICR II in more detail. More mutations might destabilise the proposed hairpin at ICR II even further. If hairpin formation at ICRII were found to be unimportant for initiation of pC221 replication, then the high degree of conservation of this inverted repeat sequence in all known rep sequences could mean that hairpin formation here is involved in some other function, e.g. termination of replication.

The mutation introduced into the 3' side of ICR II inhibits the nicking activity of REP D at the origin by approximately 8 fold. As this mutation lies within the REP D binding site, as determined by DNase I footprinting experiments, the observed effect is presumably due to the interruption of REP D binding. However, it may be that destabilisation of the proposed hairpin at ICR II, along with disruption of REP D binding, could both contribute to the decreased nicking activity. Experiments should be carried out

in order to make a second mutation on the opposite side of the stem of the proposed hairpin in order to restore perfect base pairing. If at least some of the decrease in nicking activity were to be due to destabilisation of the hairpin, then such a mutation might be expected to increase the nicking activity back towards the "wild type" level.

Much of the product obtained with this mutant origin was the relaxed, open circular form of the DNA, rather than the relaxed, covalently closed circular form. The result suggested that a mutation on the 3' side of ICR II may disrupt binding between REP D and the DNA, so that the two tend to dissociate after the nicking event, before religation can occur.

8.2.3: - Use of synthetic oligonucleotides to investigate the REP D /oriD interaction.

REP D has been shown to bind and cleave a single-stranded oligonucleotide corresponding to the upper strand sequence of oriD ICR II (Thomas, 1988). Oligonucleotides (24 bases) based on this sequence were designed to study the REP D /oriD interaction. The "wild type" oligonucleotide was cleaved by REP D at the expected site. However, changing of the bases on both sides of the nick site abolished most of this cleavage, suggesting that the bases on either side of the nick site are important for nicking of these oligonucleotides.

An oligonucleotide with the sequence on the 5' side of ICR II "jumbled", but with the wild type sequence on the 3' side and in the 6 base spacer was cleaved by REP D at the same efficiency as the wild type oligonucleotide. An oligonucleotide with the 3' side of ICR II jumbled was also

cleaved, but with much lower efficiency than the wild type oligonucleotide. This again suggests that the 3' side of ICR II is more important for REP D recognition and /or nicking than the 5' side. However, the 5' side would appear to have some role, as the oligonucleotide with both sides of ICR II jumbled was not cleaved.

These results suggest that hairpin formation at ICR II is not required for cleavage of these oligonucleotides *in vitro*. However, the oligonucleotides used here were single-stranded, as opposed to the normal double-stranded state of oriD. It must also be remembered that these oligonucleotides contain only a part of the REP D binding site, as determined by footprinting experiments.

It is possible that hairpin formation is required with double-stranded oriD DNA to bring the nick site into the single-stranded form within the loop region, to enable nicking to take place. If this were the case, then hairpin formation at ICR II might not be required for nicking of these oligonucleotides, as they are already single-stranded.

Experiments have been carried out with single-stranded pT181 DNA and REP C *in vitro* (Koepsel & Khan, 1987). Although REP C cleaves pT181 specifically at the origin when the DNA is in the double-stranded form, it was found that single-stranded DNA was also cleaved at other secondary sites which show similarities to the primary REP C cleavage site at oriC. This result led to the proposal of a consensus sequence for the nicking of single-stranded DNA by REP C (Figure 8.5). The presence of *E. coli* single-stranded binding protein (ssb) inhibited cleavage at all of the secondary nick sites, but not at the primary nick site. The fact that these secondary cleavage sites show up with single-

stranded DNA, but not with double-stranded DNA suggests that the sequence requirements at the origin are much more stringent with double-stranded DNA.

Based on these findings, oligonucleotides were designed in order to investigate the importance of highly conserved bases within the consensus sequence. Using [^{32}P] end labelled 16 base oligonucleotides, the AA dinucleotide sequence at the nick site was shown to be absolutely required for cleavage of these oligonucleotides, along with a T residue three bases upstream. However, this T residue is only conserved in 76% of the secondary REP C cleavage sites. This suggests that cleavage of small, single-stranded oligonucleotides is different to cleavage of sites within large, single-stranded DNA molecules.

Therefore, although experiments with small, single-stranded oligonucleotides yielded some interesting results, caution must be exercised in the extension of these results to the nicking of the pC221 origin by REP D *in vivo* or *in vitro*. Further experiments may be necessary using longer oligonucleotides containing the whole of the oriD sequence. The complementary strand could also be synthesized to give double-stranded DNA.

8.3: - X-ray crystallographic studies of the REP D /oriD complex.

X-ray crystallographic studies on the REP D protein are in progress. Microcrystals of REP D have already been prepared at Leicester by Dr D.B. Wigley (Figure 8.6).

Tertiary structures have been determined for some other DNA binding proteins bound to their DNA targets, for example, the repressor /operator complex of bacteriophage 434 (Anderson *et al*, 1987), the restriction endonuclease EcoRI bound to its hexanucleotide target (McClaren *et al*, 1986) and the trp repressor /operator complex (Otwinowski *et al*, 1988).

It is thought that the phage 434 repressor recognises its operators by its complementarity to a particular conformation which may be adopted by the target DNA, as well as by direct interaction with base pairs in the major groove.

The crystal structure of EcoRI restriction endonuclease in complex with its hexanucleotide DNA binding site shows that the two identical subunits of the endonuclease each consist of an α/β domain with an extended arm which wraps around the DNA. Sequence-specificity seems to be mediated via 12 hydrogen bonds between α-helical recognition domains and bases within the EcoRI recognition hexanucleotide. The DNA was found to be distorted on binding of EcoRI.

The crystal structure of the trp repressor /operator complex shows that the protein contributes 28 direct hydrogen bonds to the operator, all but 4 being to the non-esterified phosphate oxygens. There are no direct contacts with the bases which might explain specificity. Specificity appears to arise through

a combination of water mediated hydrogen bonds to critical bases in the recognition sequence and a base sequence effect that permits the operator DNA to assume a conformation that forms a large, stable interface with the protein at minimal cost in internal energy.

Therefore, X-ray crystallography of REP D bound to its DNA target should give valuable information on the conformation of the DNA bound to the protein and important protein /DNA interactions, as well as the protein structure itself. As well as X-ray diffraction studies on REP D alone, it is also hoped to study REP D bound to oriD DNA. REP D co-crystallised with a double-stranded DNA fragment containing oriD should give the structure of the covalent complex between the two. Single and double-stranded oligonucleotides corresponding to all or part of oriD could also be used to investigate interactions between REP D and different parts of the origin sequence seperately. The mutant REP D protein (Y188F) could also be co-crystalised with the origin DNA to give the structure of the protein /DNA complex in the absence of the nicking reaction.

Studies such as those described above will, hopefully, enable us to determine whether the oriD DNA is bent on binding of REP D or whether hairpin extrusion occurs at any of the inverted complementary repeat sequences within the origin. It could be that both of these conformational changes take place in the DNA simultaneously. There is evidence for the introduction of bends into linear DNA by the extrusion of cruciform structures (Gough & Lilley, 1985). Therefore, it may be that extrusion of a hairpin within the origin on binding of REP D causes bending of adjacent DNA sequences. Conversely, the binding of REP D

to the DNA may facilitate extrusion of a hairpin structure by stabilization of the bend in the DNA which may be required for the hairpin to form.

8.4: - Summary.

The aims of this project were to study the replication of pC221 *in vitro* and to investigate the binding of REP D to oriD and the probability of secondary structure formation at the origin (see section 1.4).

The findings of this project can be summarised as follows :-

(i) The initiation of replication of pC221 *in vitro* does not require rifampicin-sensetive, DNA-dependent RNA synthesis, consistent with a rolling circle replication mechanism. This does not preclude, however, the involvement of rifampicin-insensitive RNA synthesis.

(ii) The replication of pC221 *in vitro* is dependent to a certain extent on the topology of the DNA. Novobiocin inhibits pC221 replication *in vitro*. Preliminary evidence suggests that the specificity of REP D for its cognate origin sequence, oriD, may be greater than that for oriC when the plasmid DNA is in the relaxed, rather than the supercoiled form.

(iii) DNA footprinting experiments have shown that REP D binds to a 38bp region of oriD, immediately downsteam of the nick site. Enhanced cleavage patterns observed with DNase I footprinting techniques suggest that the DNA may be bent on binding of REP D.

(iv) Experiments using *in vitro* mutagenesis or synthetic oligonucleotides to alter the sequence of oriD suggest that hairpin formation may not be important for the initiation of pC221 replication by REP D *in vitro*, and also, by inference, *in*

vivo.

(v) Plasmid pCW7 appears to belong to the pT181 "family" of plamids. These all have similar genomic organisation and replicate via a common replication mechanism. Sequencing of the pCW7 rep gene has shown that it is homologous to other known rep sequences, but is sufficiently different, especially within the copy control region upstream of the rep gene, to be placed into a new staphylococcal plasmid incompatability group, inc14.

Fig 8.1: Calculated free energies of proposed hairpins at the origin sequences of the pT181 "family" of plasmids.

The free energies of the proposed hairpins at the three inverted complementary repeat sequences at the origns of the pT181 "family" of plasmids (ICR I, II and III) were calculated according to the method of Tinoco *et al* (1973).

For each plasmid, the most stable proposed hairpin was at ICR II. Therefore, a hairpin at this inverted repeat sequence would appear to be most likely to be important in the replication of these plasmids.

In some cases, hairpin formation at ICR I or III appears to be energetically unfavourable due to the Steric hindrance of a loop smaller than three bases. Therefore, it is unlikely that hairpin formation at either of these two inverted repeat sequences is important for the replication of these plasmids.

PLASMID	ΔG OF PROPOSED HAIRPIN (Kcal)		
	ICR I	ICR II	ICR III
pT181	-1.4	-19.6	-4.6
pC221	-1.4	-19.6	-7.6
pS194	+0.2	-16.4	STERICALLY IMPOSSIBLE
pUB112	STERICALLY IMPOSSIBLE	-16.4	STERICALLY IMPOSSIBLE
pC223	STERICALLY IMPOSSIBLE	-19.6	-4.6
pCW7	-1.4	-19.6	STERICALLY IMPOSSIBLE

Fig 8.2: Possible mechanism for recognition and nicking of oriD by REP D.

The initial event is probably sequence-specific recognition by REP D at ICR III. After binding, REP D may then induce a conformational change in the origin DNA, e.g. hairpin extrusion at ICR II. This hairpin may be stabilised by some means, such as DNA /protein interactions with REP D. The proposed hairpin may then be recognised by the REP D active site and nicked at a specific site within the loop region.

When the substrate DNA is in the supercoiled state, such a hairpin would be stabilised by the energy associated with the negative supercoils of the DNA. Therefore, the hairpin may be extruded and nicked without the need for sequence-specific recognition by REP D.

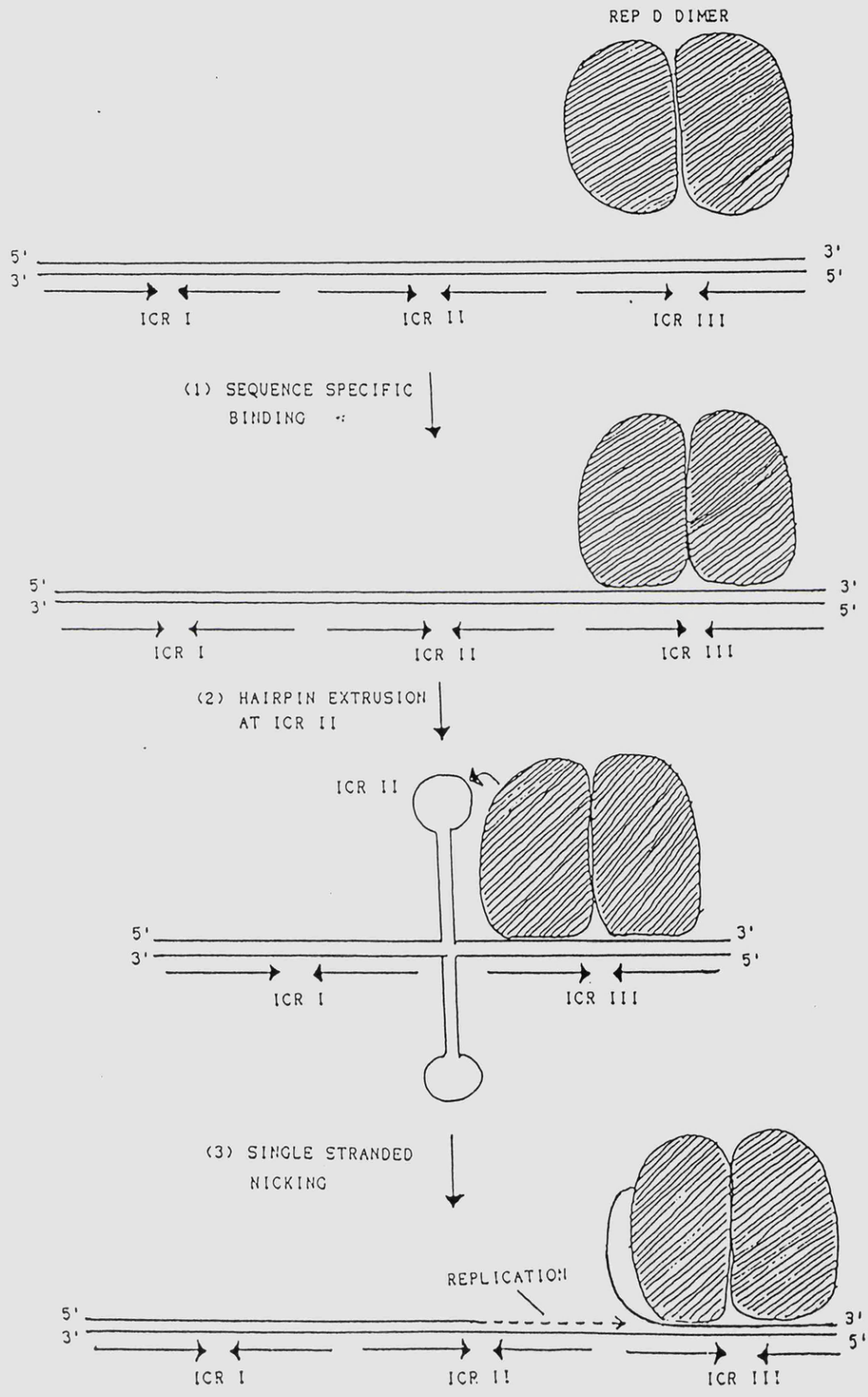

REP D DIMER
5'
3'
3'
5'
ICR I
ICR II
ICR III
(1) SEQUENCE SPECIFIC BINDING
(2) HAIRPIN EXTRUSION AT ICR II
(3) SINGLE STRANDED NICKING
REPLICATION

Fig 8.5: Secondary cleavage sites of REP C with single stranded DNA.

The sequences of 16 major secondary sites cleaved by REP C with single-stranded DNA are shown compared to the primary nick site. The consensus sequence for the nicking of single-stranded DNA by REP C is shown. The most striking feature of these cleavage sites is the presence of an AA dinucleotide immediately 3' of the site of phosphodiester bond cleavage.

REP C SECONDARY NICK SITES :-

5'----TTTTTCT ↓ AAAACGG----3'

ATCATTC AAATCAT

CATTTGG AAAATCA

ATCATGC AAATCAT

TGCGTCC AACCGGA

ATTGTCT AAATCGT

ATTTTCC AAATGAT

TTTTCTT AATCTAA

TAATCTC AATTTCG

TTGGTCT AATCA

AGATTGG AAAACTA

ATCTTGC AACCAGA

GTTTTCC AATCTGG

GATATGT AAATCAC

TTCCAGT AATCCAG

CONSENSUS SEQUENCE :-

5'----TPyNT Py ↓ AAN----3'
G
C

PRIMARY NICK SITE :-

5'----TACTCT ↓ AAT----3'

Fig 8.6. - Micro-crystals of REP D protein.

Micro-crystals of REP D were grown in hanging drops over ammonium sulphate solution (D. B. Wigley, University of Leicester). The crystals are shown at a magnification of 100X.

REFERENCES.

Anderson, J.E., Ptashne, M. & Harrison, S.C. (1987) Nature 326 846-852

Structure of the repressor-operator complex of bacteriophage 434.

Biggin, M.D., Gibson, T.J. & Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80 3963-3965

Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination.

Birnboim, H.C. & Doly, J. (1979) Nucleic Acids Res. 7 1513-1523

A rapid alkaline extraction procedure for screening recombinant plasmid DNA.

Brenner, D.G. & Shaw, W.V. (1985) EMBO Journal 4 561-568

The use of synthetic oligonucleotides with universal templates for rapid DNA sequencing: Results with staphylococcal replicon pC221.

Brady, G., Frey, J., Danbara, H. & Timmis, K.N. (1983) J. Bacteriology 154 429-436

Replication control mutations of plasmid R6-5 and their effects on interactions of the RNA-I control element with its target.

Brougham, M.J., Rowe, T.C. & Holloman, W.K. (1986) Biochemistry 25 7362-7368

Topoisomerase from *Ustilago maydis* forms a covalent complex with single-stranded DNA through a phosphodiester bond to tyrosine.

Carleton, S., Projan, S.J., Highlander, S.K., Moghazeh, S.M. & Novick, R.P. (1984) EMBO Journal 3 2407-2414

Control of pT181 replication II: Mutational analysis.

Catchpole, I., Thomas, C., Davies, A. & Dyke, K.G.H. (1988) J. Gen. Microbiol. 134 697-709

The nucleotide sequence of *Staphylococcus aureus* plasmid pT48 conferring inducible Macrolide-Lincosamide-Streptogramin B resistance and comparison with similar plasmids expressing constitutive resistance.

Crothers, D.M. (1987) Nature 325 464-465

Gel electrophoresis of protein-DNA complexes.

Diaz, R., Nordstrom, K. & Staudenbauer, W.L. (1981) Nature 289 326-328

Plasmid R1 DNA replication dependent on protein synthesis in cell extracts of *Escherichia coli*.

Diaz,R. & Staudenbauer, W.L. (1982) J. Bacteriol. 150 1077-1084

Origin and direction of mini-R1 plasmid DNA replication in cell extracts of *Escherichia coli*.

Eisenberg, S., Scott, J.F. & Kornberg, A. (1976) Proc. Natl. Acad. Sci. USA 73 3151-3155

Enzymatic replication of viral and complementary strands of duplex DNA of phage ØX174 proceeds by separate mechanisms.

Eisenberg, S. & Kornberg, A. (1979) J. Biol. Chem. 254 5328-5332
Purification and characterisation of ØX174 gene A protein.

Eisenberg, S. (1980) J. Virol 35 409-413
Cleavage of ØX174 DNA by gene A protein and formation of a tight protein-DNA complex.

Garner, M.M. & Revzin, A. (1986) Trends in Biochem. Sci. 11 395-396
The use of gel electrophoresis to detect and study nucleic acid-protein interactions.

Galas, D. & Schmitz, A. (1978) Nucl. Acids Res. 5 3157-3170
DNAase footprinting: A simple method for the detection of protein-DNA binding specificity.

Gellert, M. (1981) Ann. Rev. Biochem. 50 879-910
DNA topoisomerases.

Gennaro, M.L. & Novick, R.P. (1986) J. Bacteriol. 168 160-166
cmp, a *cis*-acting plasmid locus that increases interaction between replication origin and initiator protein.

Gennaro, M.L.; Kornblum, J. & Novick, R.P. (1987) J. Bacteriol. 168 2601-2610
A site-specific recombination function in *Stapylococcus aureus* plasmids.

Gough, G.W. & Lilley, D.M.J. (1985) Nature 313 154-156

DNA bending induced by cruciform formation.

Gronostajski, R.M. & Sadowski, P.D. (1985) Mol. Cell. Biol. 5 3274-3279

The FLP recombinase of the *Saccharomyces cerevisiae* 2μM plasmid attaches covalently to DNA via a phosphotyrosyl linkage.

Gruss, A.D., Ross, H.F. & Novick, R.P. (1987) Proc. Natl. Acad. Sci. USA 84 2165-2169

Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids.

Hatfull, G.F. & Grindley, N.D.F. (1986) Proc. Natl. Acad. Sci. USA 83 5429-5433

Analysis of γδ resolvase mutants *in vitro*: Evidence for an interaction between serine-10 of resolvase and site I of *res*.

Horowitz, D.S. & Wang, J.C. (1987) J. Biol. Chem. 262 5339-5334

Mapping the active site tyrosine of *Escherichia coli* DNA gyrase.

Ikeda, J.-E., Yudelevich, A. & Hurwitz, J. (1976) Proc. Natl. Acad. Sci. USA 73 2669-2673

Isolation and characterisation of the protein coded by gene A of bacteriophage ØX174 DNA.

Iordanescu, S. (1975) J. Bacteriol. 124 597-601

Recombinant plasmid obtained from two different, compatible staphylococcal plasmids.

Iordanescu, S. (1979) Plasmid 2 207-215

Incompatability-deficient derivatives of a small, staphylococcal plasmid.

Iordanescu, S., Surdeanu, M., Della Latta, P. & Novick, R.P. (1978) Plasmid 1 468-479

Incompatability and molecular relationships between small staphylococcal plasmids carrying the same resistance marker.

Iordanescu, S. & Surdeanu, M. (1979) Plasmid 2 216-224

Relationships between autonomous and integrated forms of a tetracycline resistance plasmid in *Staphylococcus aureus*.

Iordanescu, S. & Surdeanu, M. (1980) Plasmid 4 256-260

New incompatability groups for *Staphylococcus aureus* plasmids.

Khan, S.A., Carleton, S.C. & Novick, R.P. (1981) Proc. Natl. Acad. Sci. USA 78 4902-4906

Replication of plasmid pT181 DNA *in vitro*: Requirement for a plasmid encoded product.

Khan, S.A., Adler, G.K. & Novick, R.P. (1982) Proc. Natl. Acad. Sci. USA 79 4580-4584

Functional origin of replication of pT181 plasmid DNA is contained within a 168bp segment.

Khan, S.A. & Novick, R.P. (1983) Plasmid 10 251-259

Complete nucleotide sequence of pT181, a tetracycline-resistance plasmid from *Staphylococcus aureus*.

Koepsel, R.R., Murray, R.W., Rosenblum, W.D. & Khan, S.A. (1985a) J. Biol. Chem. 260 8571-8577

Purification of pT181-encoded REP C Protein required for the initiation of plasmid replication.

Koepsel, R.R., Murray, R.W., Rosenblum, W.D. & Khan, S.A. (1985b) Proc. Natl. Acad. Sci. USA 82 6845-6849

The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities.

Koepsel, R.R., Murray, R.W. & Khan, S.A. (1986) Proc. Natl. Acad. Sci. USA 83 5484-5488

Sequence-specific interaction between the replication initiator protein of plasmid pT181 and its origin of replication.

Koepsel, R.R. & Khan, S.A. (1986) Science 233 1316-1318

Static and initiator protein-enhanced bending of DNA at a replication origin.

Koepsel, R.R. & Khan, S.A. (1987) Nucleic acid Res. 15 4085-4097

Cleavage of single-stranded DNA by plasmid pT181-encoded REP C protein.

Kumar, C.C. & Novick, R.P. (1985) Proc. Natl. Acad. Sci. USA 82 638-642

Plasmid pT181 replication is regulated by two countertranscripts.

Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA 82 488-492

Rapid and efficient Site-specific mutagenesis without phenotypic selection.

Lacatena, R.M. & Cesareni, G. (1981) Nature 294 623-626

Base-pairing of RNA I with its complementary sequence in the primer precursor inhibits Col EI replication.

Lacatena, R.M. & Cesarini, G. (1983) J. Mol. Biol. 179 635-650

The interaction between RNA I and the primer precursor in the regulation of Col EI replication.

Lilley, D.M.J. (1980) Proc. Natl. Acad. Sci. USA 77 6468-6472

The inverted repeat as a recognisable structural feature in supercoiled DNA molecules.

Lilley, D.M.J. (1982) Topics in Nucl. Acid Structure - Part 2 (Macmillan) 173-198

Hairpin-loops in supercoiled DNA.

Lindberg, M. (1981) in "Staphylococci and staphylococcal infections", Zentralblatt fur Bakteriologie, Mikrobiologie und Hygiene, Supplemente 10, Jeljaszewicz, J. (Editor), 535-541.

Genetic studies in *Staphylococcus aureus* using protoplasts: Cell fusion and transformation.

Lui-Johnson, H.-N., Gartenberg, M.R. & Crothers, D.M. (1986) Cell 47 995-1005

The DNA binding domain and binding angle of *Escherichia coli* CAP protein.

Lyon, B.R. & Skurray, R. (1987) Microbiol. Rev. 51 88-134

Antimicrobial resistance of *Staphylococcus aureus*: Genetic basis.

Manch-Citron, J.N., Gennaro, M.L., Majumder, S. & Novick, R.P. (1986) Plasmid 16 108-115

REP C is rate-limiting for pT181 plasmid replication.

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) in "Molecular Cloning: A laboratory manual." (Cold Spring Harbour Laboratory Press, NY.)

McClarin, J.A., Frederick, C.A., Wang, B.-C., Greene, P., Boyer, H.W., Grable, J. & Rosenberg, J.M. (1986) Science 234 1526-1541

Structure of the DNA-Eco RI endonuclease recognition complex at 3Å resolution.

Messing, J. (1983) Methods in Enzymology 101 20-78

New M13 vectors for cloning.

Mok, M. & Marians, K.J. (1987) J. Biol. Chem. 262 2304-2309

Formation of rolling-circle molecules during ØX174 complementary strand DNA replication.

Mukherjee, S., Patel, I. & Bastia, D. (1985) Cell 43 189-197

Conformational changes in a replication origin induced by an initiator protein.

Novick, R.P. (1976) J. Bacteriol. 127 1177-1187

Plasmid-protein relaxation complexes in *Staphylococcus aureus*.

Novick, R.P., Adler, G.K., Majumder, S., Khan, S.A., Carleton, S., Rosenblum, W.D. & Iordanescu, S. (1982) Proc. Natl. Acad. Sci. USA 79 4108-4112

Coding sequence for the pT181 repC product: A plasmid-coded protein uniquely required for replication.

Novick, R.P., Projan, S.J., Rosenblum, W. & Edelman, I. (1984a) Mol. Gen. Genet. 195 374-377

Staphylococcal plasmid cointegrates are formed by host- and phage-mediated general *rec* systems that act on short regions of homology.

Novick, R.P., Adler, G.K., Projan, S.J., Carleton, S., Highlander, S.K., Gruss, A., Khan, S.A. & Iordanescu, S. (1984b) EMBO Journal 3 2399-2405

Control of pT181 replication I: The pT181 copy control function acts by inhibiting the synthesis of a replication protein.

Otwinowski, Z., Scheviyz, R.W., Zhang, R.-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. & Sigler, P.B. (1988) Nature 335 321-329

Crystal structure of *trp* repressor /operator complex at atomic resolution.

Pabo, C.O. & Sauer, R.T. (1984) Ann. Rev. Biochem. 53 293-321

Protein-DNA recognition.

Projan, S.J., Kornblum, J., Moghazeh, S.L., Edelman, I., Gennaro, M.L. & Novick, R.P. (1985) Mol. Gen. Genet. 199 452-464

Comparative sequence and functional analysis of pT181 and pC221, cognate plasmid replicons from *Staphylococcus aureus*.

Rosenberg, M. & Court, D. (1979) Ann. Rev. Genet. 13 319-353

Regulatory sequences involved in the promotion and termination of RNA transcription.

Ruby, C. & Novick, R.P. (1975) Proc. Natl. Acad. Sci. USA 72 5031-5035

Plasmid interactions in *Staphylococcus aureus*: Non-additivity of compatible plasmid DNA pools.

Ryder, K., Silver, S., DeLucia, A.L., Fanning, E. & Tegtmeyer, P. (1986) Cell 44 719-725

An altered conformation in origin region I is a determinant for the binding of SV40 large T antigen.

Sakakibara, Y. & Tomizawa, J.I. (1974) Proc. Natl. Acad. Sci. USA 71 802-806

Replication of colicin EI plasmid DNA in cell extracts.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74 5463-5467

DNA sequencing with chain terminating inhibitors.

Sanhueza, S. & Eisenberg, S. (1984) Proc. Natl. Acad. SCi. USA 81 4285-4289
Cleavage of single-stranded DNA by the ØX174 A* protein: The A*-single-stranded DNA covalent linkage.

Scott, J.R, (1984) Microbiological Reviews 48 1-23
Regulation of plasmid replication.

Shaw, W.V. (1983) Crit. Rev. Biochem. 14 1-46
Chloramphenicol acetyltransferase: Enzymology and molecular biology.

Shaw, W.V., Brenner, D.G., LeGrice, S.F.J., Skinner, S.E. & Hawkins, A.R. (1985) FEBS Lett. 179 101-106
Chloramphenicol acetyltransferase gene of staphylococcal plasmid pC221.

Skinner, S.E. (1984) PhD Thesis, Leicester University.
Staphylococcal chloramphenicol-resistance plasmids.

Tautz, D. & Renz, M. (1983) Anal. Biochem. 132 14-19
An optimised Freeze-Squeeze method for the recovery of DNA fragments from agarose gels.

Te Reile, H., Michel, B. & Ehrlich, S.D. (1986) Proc. Natl. Acad. Sci. USA 83 2541-2545
Single-stranded plasmid DNA in *Bacillus subtilis* and *Staphylococcus aureus*.

Thomas, C.D. (1988) PhD Thesis, Leicester University.
Mechanistic studies of REP D, a staphylococcal plasmid replication initiator protein.

Tinoco, I., Uhlenbeck, O.C. & Levine, M.D. (1971) Nature 230 362-367
Estimation of secondary structure in ribonucleic acids.

Tinoco, I., Borer, P.N., Dengler, B., levine, M.D., Uhlenbeck, O., Crothers, D. & Grallia, J. (1973) Nature New Biology 246 40-41
Improved estimation of secondary structure in ribonucleic acids.

Tomizawa, J. & Itoh, T. (1982) Cell 31 575-583
The importance of RNA secondary structure in Col E1 primer formation.

Tse, Y.C., Kirkegaard, K. & Wang, J.C. (1980) J. Biol. Chem. 255 5560-5565
Covalent bonds between proteins and DNA.

Tullius, T.D. (1987) Trends in Biochem. Sci. 12 297-300
Chemical "snapshots" of DNA: Using the hydroxyl radical to study the structure of DNA and DNA-protein complexes.

Wang, J.C. (1985) Ann. Rev. Biochem. 54 665-697
DNA topoisomerases.

Whitson, P.A., Hsieh, W.T-., Wells, R.D. & Matthews, K.S. (1987) J. Biol. Chem. 262 4943-4946

Supercoiling facilitates *lac* operator-repressor-pseudooperator interactions.

Wilson, C.R. & Baldwin, J.N. (1978) J. Bacteriol. 136 402-413

Characterisation and construction of molecular cloning vehicles within *Staphylococcus aureus*.

Wilson, C.R., Skinner, S.E. & Shaw, W.V. (1981) Plasmid 5 245-258

Analysis of two chloramphenicol resistance plasmids from *Staphylococcus aureus*: Insertional activation of Cm resistance, mapping of restriction sites and construction of cloning vehicles.

Wu, H.-M. & Crothers, D.M. (1984) Nature 308 509-513

The locus of sequence-directed and protein induced DNA bending.

Zahn, K. & Blattner, F.R. (1985) Nature 317 451-453

Sequence-induced DNA curvature at the bacteriophage λ origin of replication.

D.F. Balson
Leicester University

PhD Thesis
January 1989

Replication initiation studies for a family of small staphylococcal plasmids.

Abstract.

pC221 belongs to a family of staphylococcal plasmids, including pT181, pS194, pC223, pUB112 and pCW7 . All possess open reading frames with 70-80% homology to the pC221 repD gene. REP D has sequence specific topoisomerase activity at the pC221 origin (oriD) which is thought to be involved in replication initiation.

DNase I footprinting has been carried out, showing that REP D binds to a region of oriD downstream of the nick site. The pattern of DNase I cleavage suggests that REP D contacts one face of the DNA helix, which may be bent around the protein.

Extracts of *S. aureus* support incorporation of radioactive dNTPs into pC221 in the presence of REP D. Labelling with $\alpha[^{32}P]$ dATP shows that replication initiates within the region containing oriD and proceeds in the direction expected for elongation of a 3' OH generated by nicking at oriD. With supercoiled DNA, REP D initiates replication of other members of this plasmid family *in vitro*. However, with relaxed DNA, REP D is specific for oriD, suggesting that a change in the DNA, stabilised by supercoiling of the DNA or by binding of REP D, may be required for nicking.

Of three inverted repeat sequences (ICRI, II & III) at the origin, ICRII has the greatest predicted hairpin stability and is almost totally conserved. Nicking takes place within the loop of this proposed hairpin. Disruption of base pairing within this hairpin has been investigated by mutagenesis of cloned oriD and using oligonucleotides based on the ICRII sequence. These experiments show that the 3' side of ICRII is more important for nicking than the 5' side. This is in agreement with footprinting data which shows that REP D binds the 3' side of ICRII, along with the whole of ICRIII. However, there is no evidence for hairpin formation at ICRII being required for nicking.