Genetics and molecular biology of the *Streptomyces lividans* plasmid pIJ101

A thesis submitted for the degree of Doctor of Philosophy of the
University of London

Sara Zaman
Department of Biology
University College London

1991
Abstract.

The 0.5kb SpeI(53)-BclI(49) fragment containing the plJ101 cop/korB gene was cloned into the E. coli plasmid pUC8 under the control of the lacZ promoter, creating pQR206. pQR206 produces a 10kDa LacZ-Cop/KorB fusion protein in vitro and in vivo in E. coli. Cop/KorB crude protein is able to retard a 0.5kb SpeI(53)-BclI(49) cop/korB fragment, a 0.5kb Sau3A cop/korB fragment, and a 0.65kb SalGI(19)-SstII(16) kilB fragment in band-shift assays; however no retardation was observed with a 0.7kb SpeI(53)-BclI(57) sti fragment. DNA footprinting analysis showed that the Cop/KorB protein protects overlapping 33b sequences on the cop/korB coding and non-coding strands respectively. Furthermore, a hypersensitive site present on each strand divides these 33b regions into two equal 16b regions. The kilB protected region covers 60b and 52b on the coding and non-coding strands respectively. Both these defined kilB and cop/korB operator sites for Cop/KorB binding show strong sequence homology and also correspond with the location of their respective promoters. Competition assays suggest that the Cop/KorB protein may have a higher affinity for the kilB operator than for the cop/korB operator.

A 0.7kb BclI(57)-SpeI(53) sti fragment inserted in its correct orientation into plJ702 (a sti'cop/korB' plJ101 derivative) and into plJ702::pUC8 shuttle vectors prevented them from accumulating ssDNA and structurally rearranging respectively. Constructs which contained sti in the reverse orientation were found to accumulate ssDNA. Thus, sti is only active as the site for second-strand synthesis in its natural orientation with respect to the basic replicon. The sti determinant was further defined to a 0.53kb SstII(55)-SpeI(53) fragment which contains a potential stem-loop structure with a ΔG of -64.00 kcal/mol. Cop/KorB does not inhibit the conversion of ssDNA to the double-stranded form and does not significantly alter copy number.

The minimal replicon was further defined to a 2.0kb Ball(15)-SstII(63) fragment encompassing only the rep ORF and a non-coding region which may either contain the plus origin or the rep promoter. The 1.4kb Norl(12)-SstII(63) rep ORF was cloned into pBGS19- under the control of the lacZ promoter, creating pQR431a. pQR431a produces a 50kDa LacZ-Rep fusion protein in vitro.
Acknowledgements.

I am indebted to Dr John Ward and Dr Hilary Richards for their constant supervision and encouragement throughout the project. I am also indebted to Alex Chapman for proof-reading and formatting the text, and for assisting with the figures.

I would like to thank Dr. Rita Barallon for her generosity and invaluable assistance in the lab during the early years of the project. Thanks also to the rest of the group especially to Carol for organising the numerous parties and Sunil for supplying endless amounts of diet cokes and crunch bars.

A very special thanks to Sami (my best friend, husband, and lab partner) for being so good at all three jobs. Finally, I would like to dedicate this thesis to my father who has supported me throughout my education.
TABLE OF CONTENTS

CHAPTER ONE. Introduction.................................................................9
1.1 Prologue...........................................................................................9
1.2 The biology of *Streptomyces* plasmids..............................................12
  1.2.1 Plasmid maintenance and replication...........................................12
  1.2.2 Plasmid transfer and pock formation..........................................13
  1.2.3 Plasmid-mediated fertility and site-specific integration.................15
1.3 The *Streptomyces* plasmid, pU101..................................................18
  1.3.1 Background..................................................................................18
  1.3.2 Replication and stability determinants.........................................18
  1.3.3 Transfer and spread determinants..............................................22
  1.3.4 Sti (strong incompatibility) and Cop (copy number) determinants....25
  1.3.5 Identification of pU101 promoters and terminators.....................28
1.4 *Streptomyces* plasmids as cloning vectors....................................29
  1.4.1 Construction of cloning vectors based on pU101............................29
  1.4.2 Plasmid structural instability......................................................30
1.5 Aims of the project..........................................................................31

CHAPTER TWO. Materials and Methods............................................33
2.1 Materials.........................................................................................33
2.2 Bacterial strains and plasmids........................................................33
2.3 Culture conditions............................................................................33
2.4 Maintenance of bacterial strains......................................................39
2.5 Preparation and transformation of *E. coli* JM107 competent cells....39
2.6 Small-scale isolation of plasmid DNA from *E. coli*.........................40
2.7 Large-scale isolation of plasmid DNA from *E. coli*.........................40
2.8 Preparation and transformation of *Streptomyces* protoplasts..........41
2.9 Small-scale isolation of *Streptomyces* total DNA............................42
2.10 Small-scale isolation of plasmid DNA from *Streptomyces*...............42
2.11 Large-scale isolation of plasmid DNA from *Streptomyces*..............43
2.12 Restriction enzyme digestions, ligations, and agarose gel electrophoresis.43
2.13 Recovery of DNA from agarose gels..............................................44
2.14 Radioactive-labelling of DNA......................................................45
2.15 Southern blotting..........................................................................46
2.16 Isolation of total mRNA from E. coli/Streptomyces ................................... 47
2.17 Northern blotting ....................................................................................... 47
2.18 High resolution S1 mapping ...................................................................... 48
2.19 Double-stranded DNA sequencing ............................................................. 49
2.20 SDS-polyacrylamide gel electrophoresis (PAGE). ..................................... 50
2.21 In vitro coupled transcription/translation .................................................. 51
2.22 Protein extraction from E. coli ................................................................... 52
2.23 Gel retardation assay ................................................................................ 52
2.24 Competition assay .................................................................................... 52
2.25 DNaseI protection assay (DNA footprinting) ............................................. 53

CHAPTER THREE ........................................................................... 55

3.1 Construction of pQR200 (cop/korB plus sti in pUC8) .................................. 55
3.2 Construction of pQR206 (cop/korB in pUC8) ............................................. 56
3.3 Sequencing of cop/korB (pQR206) ............................................................. 58
3.4 Expression of Cop/KorB protein in vitro ................................................... 61
3.5 Expression of Cop/KorB protein in vivo in E. coli ....................................... 66
3.6 Cloning of cop/korB into other expression vectors ..................................... 72
3.7 Analysis of Cop/KorB binding ................................................................. 77
  3.7.1 Gel retardation assay ....................................................................... 78
  3.7.2 Competition assay ....................................................................... 89
  3.7.3 Identification of cop/korB and kilB recognition sequences ............... 93
3.8 DNaseI protection analysis (DNA footprinting) ........................................... 99
3.9 Summary .................................................................................................... 114

CHAPTER FOUR ............................................................................. 116

4.1 Cloning of cop/korB and sti into pIJ702 ...................................................... 116
4.2 Analysis of ssDNA from plasmids containing sti and cop/korB ............... 119
4.3 Location of the sti function .................................................................... 121
4.5 Expression of Rep protein in vitro and in vivo in E. coli ......................... 131
4.6 Summary .................................................................................................... 135

CHAPTER FIVE. General discussion ......................................................... 137

5.1 Prologue .................................................................................................... 137
5.2 Cloning and expression of the cop/korB gene product ............................ 137
5.3 Cop/KorB binding on cop/korB and kilB genes ....................................... 139
5.4 Model for KorB regulation of korB and kilB genes............................... 144
5.5 Protein structure analysis of Cop/KorB.................................................. 150
5.6 Cop/KorB binding on sti...................................................................... 153
5.7 Definition of the pIJ101 minimal replicon............................................. 158
5.8 Conclusion............................................................................................ 160

REFERENCES....................................................................................... 162

APPENDIX I. Nucleotide sequence of pIJ101............................................ 176
Figure 1.1 A) Diagrammatic representation of the *Streptomyces* life cycle........10
Figure 1.1 B) Scanning electron microscopy of *S. lividans* ..........................10
Figure 1.2 Physical and genetic map of the *S. lividans* plasmid, pIJ101........19
Figure 1.3 Putative active site of pIJ101 Rep protein.................................21
Figure 1.4 Stages of rolling circle replication..............................................27
Figure 2.1 Antibiotic concentrations for selecting resistant strains ..............34
Figure 2.2 Characteristics of bacterial strains used......................................35
Figure 2.3 Characteristics of plasmids used................................................36
Figure 2.4 List of plasmid constructs............................................................37
Figure 3.1 Construction of pQR200 and pQR206.........................................57
Figure 3.2 Sequence of korB gene in pQR206..............................................59
Figure 3.3 *In vitro* coupled transcription-translation of pQR206 resolved on
an 18% SDS-polyacrylamide gel.................................................................63
Figure 3.4 *In vitro* coupled transcription-translation of pQR206 resolved on
an 8M urea 12.5% SDS-polyacrylamide gel..................................................64
Figure 3.5 Northern analysis of pQR206-encoded RNA ...............................67
Figure 3.6 *In vivo* synthesis of cop/korB-encoded protein product from
pQR206 in *E. coli* resolved on an 8M urea 12.5% SDS-polyacrylamide gel.....68
Figure 3.7 *In vivo* synthesis of cop/korB-encoded protein product from
pQR206 in *E. coli* resolved on an 18% SDS-polyacrylamide gel..................69
Figure 3.8 Densitometry scan.......................................................................71
Figure 3.9 Construction of pQR412, pQR413 and pQR414............................73
Figure 3.10 Sequence of pQR414 across the junction of the tac promoter
and korB gene...............................................................................................75
Figure 3.11 *E. coli* total protein extracts containing pQR414 and pKK223.3
resolved on an 18% SDS-polyacrylamide gel..............................................76
Figure 3.12 Construction of pQR427, pQR428 and pQR456............................79
Figure 3.13 Gel retardation assay using crude Cop/KorB protein extract
incubated with kilB and cop/korB (SpeI-cleaved) probes
respectively ......................................................................................................81
Figure 3.14 Construction of pQR452...............................................................83
Figure 3.15 Gel retardation assay using crude Cop/KorB protein extract
incubated with kilB and cop/korB (Sau3A-cleaved) probes
respectively ......................................................................................................84
Figure 3.16 Gel retardation assay using crude Cop/KorB protein extract
incubated with different concentrations of cop/korB (Sau3A-cleaved) probe...85
Figure 3.17 Construction of pQR417, pQR437, and pQR438..........................87
Figure 3.18 Gel retardation assay using crude Cop/KorB protein extract
incubated with sti probe..................................................................................88
Figure 3.19 Competition assay using 2μl of crude Cop/KorB protein extract
preincubated with increasing concentrations of unlabelled kilB
and cop/korB (Sau3A-cleaved) DNA............................................................90
Figure 3.20 Competition assay using 5µl of crude Cop/KorB protein extract preincubated with increasing concentrations of cop/korB (Sau3A-cleaved) DNA.............................. 92
Figure 3.21 Construction of pQR458 and pQR459........................................ 94
Figure 3.22 Gel retardation assay using crude Cop/KorB protein extract incubated with kilB and cop/korB-derived probes............................ 96
Figure 3.23 Construction of pQR457............................................................... 98
Figure 3.24 DNaseI protection analysis of pQR206-derived Cop/KorB protein extract incubated with 0.65kb kilB and 0.5kb cop/korB probes end-labelled with T4 Polynucleotide Kinase .........................101
Figure 3.25 DNaseI protection analysis of pQR206-derived Cop/KorB protein extract incubated with 0.65kb kilB and 0.5kb cop/korB probes filled-in with T7 DNA polymerase .......................... 103
Figure 3.26 Summary of kilB and korB footprint data................................... 105
Figure 3.27 A) Construction of pQR461 .......................................................... 109
Figure 3.27 B) Sequence across the pQR461 SpeI/XbaI junction...................... 109
Figure 3.28 DNaseI protection analysis of pQR461-derived Cop/KorB protein extract incubated with 0.65kb kilB and 0.5kb cop/korB probes end-labelled with T4 Polynucleotide Kinase .........................110
Figure 3.29 DNaseI protection analysis of pQR461-derived Cop/KorB protein extract incubated with 0.65kb kilB and 0.5kb cop/korB probes filled-in with T7 DNA polymerase .......................... 112
Figure 4.1 Construction of pQR410a/b, pQR411a/b, pQR421a/b, and pQR422a/b................................. 118
Figure 4.2 Detection of ssDNA in S. lividans ................................................. 120
Figure 4.3 Construction of pQR441a/b and pQR443a/b................................. 122
Figure 4.4 Construction of pQR434, pQR435, and pQR436............................ 125
Figure 4.5 Construction of pQR445 and pQR446.......................................... 127
Figure 4.6 Construction of pQR453 and pQR455.......................................... 130
Figure 4.7 Sequence of pQR431 across the junction of the lacZ promoter and rep gene.......................................................... 132
Figure 4.8 In vitro coupled transcription-translation of pQR431a/b resolved on a 10% SDS-polyacrylamide gel............................................ 133
Figure 4.9 E. coli total protein extracts containing pQR431a/b and pBGS19- resolved on a 10% SDS-polyacrylamide gel.......................... 134
Figure 5.1 Model for KorB regulation of korB and kilB genes......................... 145
Figure 5.2 The kil-kor regulon of the IncP plasmid RK2................................. 149
Figure 5.3 The kil-kor regulon of the IncN plasmid pKM101.......................... 149
Figure 5.4 A) Putative DNA-binding domains of KorB .................................. 151
Figure 5.4 B) Helical wheel representation of the α-helix turn α-helix region of KorB .......................................................... 151
Figure 5.5 Potential secondary structure of the sti region of pIJ101................. 156
Figure 5.5 Nucleotide sequence comparison of DNA regions containing ssi sites of phages and plasmids........................................... 157
CHAPTER ONE. Introduction.

1.1 Prologue.

*Streptomyces* are Gram-positive, spore-forming soil micro-organisms that exhibit a complex cycle of morphological differentiation (Fig. 1.1A) (Chater 1984). Development involves two spatially separated types of cells: a) the branching, multi-nucleate hyphae of the substrate mycelium (which penetrate the stratum and degrade complex organic material by the secretion of extracellular hydrolytic enzymes), and b) the upwardly protruding hyphae of the aerial mycelium which undergo septation to form long chains of uni-nucleate cells that metamorphose into spores (Fig 1.1B). Aerial mycelium formation is accompanied by partial lysis of the substrate mycelium, which provides a source of nutrients for subsequent development stages (Wildermuth 1970, Wildermuth and Hopwood 1970). The onset of sporulation is probably due to nutritional limitations and coincides with the production of a wide variety of secondary metabolites, including antibiotics and extracellular enzymes (Hopwood *et al.* 1973). Over 60% of known naturally occurring antibiotics come from *Streptomyces* (Berdy 1980).

Renaturation studies on the DNA of *S. coelicolor* A3(2) indicate that the genome size is approximately $10^4$ kilobases (kb), and consists of a single circular chromosome (Benigni *et al.* 1975). *Streptomyces* chromosomal DNA has an unusually high G+C content and is typically about 73% (Enquist and Bradly 1971). Thus, the organization of DNA sequences needed for autonomous replication and for the control of gene expression may be different from those found in other bacteria, such as *Escherichia coli* (Rosenberg and Court 1979; Stormo *et al.* 1982; Hawley and McClure 1983) and *Bacillus subtilis* (Moran *et al.* 1982) which possess a relatively high A+T content.

Plasmids from *Streptomyces* are of major interest, firstly because of their potential use as cloning vectors in genetic manipulation of this industrially important genus of bacteria, and secondly because they may have unique properties
Figure 1.1
A) Diagrammatic representation of the *Streptomyces* life cycle.
(Adapted from Chater and Merrick 1979).

B) Scanning electron microscopy of *S. lividans*.
The panel on the left shows the early stages of aerial hyphae production and the panel on the right shows their eventual differentiation into long chains of spores. (Courtesy of V. Zaman).
which allow them to exist in this structurally and developmentally complex bacteria. Naturally occurring plasmids with a wide range of sizes (4kb to over 200kb) and copy numbers (one to several hundred) have been reported in many *Streptomyces* strains (Hopwood *et al.* 1986a). Three major groups of plasmids have been intensively developed as cloning vectors: the SLP1 family (Bibb *et al.* 1981; Thompson *et al.* 1982), pIJ101 (Kieser *et al.* 1982), and SCP2* (Bibb and Hopwood 1981; Lydiate *et al.* 1985).

SLP1.2, the most stable of the SLP1 family, is a small (14.5kb), low copy number (4-5 copies/chromosome) plasmid with a limited host range. SLP1 was the first plasmid discovered to be naturally integrated into the *S. coelicolor* chromosome. pIJ101 is a small (8.9kb), multi-copy (40-300 copies/chromosome) plasmid with a broad host range. The advantage of using pIJ101 as a cloning vector is that because of its small size and high copy number, it is relatively easy to isolate plasmid DNA for physical analysis of recombinants and it is useful in the amplification of cloned gene products. SCP2* is a large (31.4kb), low copy number (1-5 copies/chromosome) plasmid with a narrow host range, and is the vector of choice for cloning gene products that may be lethal to the cell in high concentrations or for cloning genes required in single copies. The use of *Streptomyces* plasmids as cloning vectors has become possible due to two major advances; a) the development of rapid, effective, and consistent methods for the isolation of plasmid DNA from *Streptomyces* (Kieser 1984) and b) the ability to transform *Streptomyces* protoplasts with plasmid DNA at high frequencies (Bibb *et al.* 1978).

The development of recombinant DNA technology in *Streptomyces* has culminated in the production of novel hybrid antibiotics by introducing antibiotic biosynthetic genes, subcloned on plasmids, into different antibiotic-producing *Streptomyces* strains. For example, the hybrid antibiotics dihydrogranatirhodin (Hopwood *et al.* 1985a) and mederrhodin A and B (Hopwood *et al.* 1985a; Omura *et al.* 1986) were created by introducing *S. coelicolor* A3(2) actinorhodin
biosynthesis genes subcloned on a SCP2*-based vector (Malpartida and Hopwood 1984) into *S. violaceoruber* Tu22 (a natural granaticin producer) and *Streptomyces* sp. AM7161 (a natural medermycin producer) respectively. More recently, the hybrid antibiotic isovaleryl spiramycin was produced by introducing a cloned carbomycin biosynthetic gene from *S. thermotolerans* into *S. ambofaciens* (a natural spiramycin producer) and *S. lividans* (grown in the presence of spiramycin) (Epp et al. 1989).

With the exception of methylenomycin biosynthesis and resistance encoded by the 350kb linear plasmid SCP1 (Kirby and Hopwood 1977; Chater and Bruton 1985; Kinashi et al. 1987; Kinashi and Shimaji-Murayama 1991), the only observed phenotypic properties determined by *Streptomyces* plasmids are concerned with conjugal transfer of chromosomal genes and genetic recombination ("fertility") (Hopwood et al. 1986a). Transfer and fertility functions require only a few gene products in *Streptomyces*, and thus the role of most of the plasmid DNA remains unknown (Hopwood et al. 1986a).

In order for *Streptomyces* plasmids to be maintained in the cytoplasm of growing, dividing and sporulating cells, the plasmids must contain DNA sequences that control their replication, stability, and partition, apart from the transfer and fertility functions mentioned above (Hopwood et al. 1986a). Each of these functions are discussed in general below, after which a more detailed account of pIJ101 is given.

1.2 The biology of *Streptomyces* plasmids.

1.2.1 Plasmid maintenance and replication.

The essential region of a plasmid contains two important features: a) the replication origin, which is a site where replication is initiated and b) regulatory genes, which normally encode for negatively-acting molecules (such as a replication protein and/or a RNA product) that control the initiation of plasmid replication and thereby determine plasmid copy number.
Apart from pIJ101, little is known about the direction of replication of
*Streptomyces* plasmids, whether replication proceeds uni- or bi-directionally, or
whether plasmids possess more than one potential origin of replication. Little also
is known about the basis of control of plasmid copy number in *Streptomyces*.

A partition function may be responsible for the stable inheritance of plasmid
copies to daughter cells (Bibb and Hopwood 1981; Lydiate *et al.* 1985), presumably
by interacting between a specific sequence on the plasmid DNA and a segregation
site within the cell (such as a membrane component). In SCP2*, when this partition
function is deleted, the regular distribution of plasmid copies to daughter cells
collapses, and plasmid-free cells begin to accumulate in the culture even though
replication may be continuing at its normal rate. Partition may be a complex
phenomenon in *Streptomyces* because of their differentiation into mycelia and
spores, and there may be different partition mechanisms operating in these two
growth phases (Hopwood *et al.* 1986a).

The competition for replication initiation and partition sites may be manifested
as incompatibility; that is, the inability of two related plasmids to be stably
maintained in the same cell (Hopwood *et al.* 1986a). Only a few *Streptomyces*
plasmids have been tested for incompatibility, although SLP1, SLP2, SCP1, SCP2,
and pIJ101 are known to be compatible with one another (Hopwood *et al.* 1986a).
Incompatibility has not been observed between derivatives of the high copy number
plasmid pIJ101 (Kieser *et al.* 1982). However, incompatibility exists amongst
SCP1 derivatives and amongst SCP2* derivatives, as is usually the case for low
copy number plasmids (Hopwood *et al.* 1986a).

1.2.2 Plasmid transfer and pock formation.

The property of pock formation, also known as lethal zygosis (Ltz), is a feature
of many conjugative *Streptomyces* plasmids (Hopwood *et al.* 1986a). This
phenotype is useful in the isolation of plasmid-negative strains (Bibb and Hopwood
1981), in the development of a plasmid transformation system (Bibb *et al.* 1978), in
the detection of low frequency interspecific transfer (Bibb and Hopwood 1981), and in the detection of otherwise cryptic plasmids (Bibb et al. 1981).

Pock formation is exhibited in plate cultures when a strain containing a plasmid grows in contact with a plasmid-free strain. Under these conditions, each spore of the plasmid-bearing strain when surrounded by a population of plasmid-free spores, gives rise to a circular zone of retarded growth termed a "pock" (Bibb et al. 1981); the cells within the pock are recipients of the transferred plasmid. Similarly, when a colony or a patch of a plasmid-bearing culture is replica-plated onto a lawn of plasmid-free strain prior to growth, the plasmid-bearing culture becomes surrounded by a narrow zone of inhibition within which the growth of the recipient strain is retarded. This phenomenon was shown to be dependent on plasmid transfer, since transfer-deficient variants of SCP1 failed to exhibit pock formation (Kirby and Hopwood 1977). Furthermore, a plasmid-encoded "spread" function is also needed for the production of normal size pocks. The spread function is distinct from intermycelial transfer and may involve the intramyecelial migration of plasmid copies within the recipient culture after primary transfer from the donor (Bibb and Hopwood 1981).

The inhibition reaction described above was compared to that of lethal zygosis in *E. coli* (Skurray and Reeves 1973), in which F^- cells are killed by simultaneous conjugation with several Hfr cells, although entry of a *Streptomyces* plasmid into a plasmid-free strain probably does not involve the killing of the recipient mycelium. The physiological mechanism responsible for the lethal zygosis reaction is still unknown, but a plausible hypothesis is that this retardation in development may be due to a transient derepression of one or more plasmid functions upon transfer into plasmid-free mycelium (Hopwood et al. 1986a).

Plasmid transfer in *Streptomyces* may involve a simpler mating phenomenon compared to the complex pilus mechanism encoded by self-transmissible plasmids in Gram-negative bacteria (Bradley 1980), since the transfer region in conjugative *Streptomyces* plasmids is relatively smaller than that of *E. coli* plasmids. For
example, the transfer regions of SCP2*, SLP1.2, and pIJ101 have been found to be <15kb (Bibb and Hopwood 1981; Lydiate et al. 1985); <5kb (Bibb et al. 1981, Thompson et al. 1982), and 1.4kb (Kieser et al. 1982) respectively, whereas F-like plasmids have a transfer region of about 35kb (although most of this comprises genes for pilus production; Willets and Skurray 1980). It is interesting to note that Mazodier et al. (1989) have constructed E. coli-Streptomyces shuttle vectors (containing the pBR322 and pIJ101 origins of replication and the RK2 origin of transfer) which are capable of conjugal transfer from E. coli to Streptomyces when RP4 (IncP) transfer functions are supplied in trans.

Most Streptomyces plasmids are self-transmissible at high frequencies and almost 100% of the recipient progeny receive plasmids when donor and recipient strains are mated on solid media (Bibb and Hopwood 1981; Kieser et al. 1982; Hopwood et al. 1986a). Successful interspecific plasmid transfer depends on several factors and may fail because of host restriction, limited host range of the plasmid, strain incompatibility (for example, due to antibiotic or bacteriocin production by one strain), or if the recipient strain already contains an incompatible plasmid (Hopwood et al. 1986a). In the latter case, a strain containing a particular plasmid is resistant (LtzR) to pock formation when exposed to a strain containing the same plasmid (or an Ltz+ derivative of the plasmid), but sensitive (LtzS) to a strain carrying a non-related plasmid (Hopwood et al. 1986a).

1.2.3 Plasmid-mediated fertility and site-specific integration.

Many self-transmissible Streptomyces plasmids have the ability to mobilize chromosomal genes during conjugal matings, thereby leading to the formation of genetic recombinants ("fertility"). This interaction between an autonomous plasmid and the chromosomal DNA has been shown for several plasmids including SLP1 from S. coelicolor A3(2) (Bibb et al. 1981; Omer and Cohen 1984, 1986, 1989; Omer et al. 1988; Lee et al. 1988; Grant et al. 1989), pIJ110 from S. parvulus (Hopwood et al. 1984), pIJ408 from S. glaucescens (Hopwood et al. 1984; Sosio et
pSAM2 from *S. ambofaciens* (Pernodet *et al.* 1984; Simonet *et al.* 1987; Boccard *et al.* 1988, 1989a,b; Kuhstoss *et al.* 1989), pSG1 from *S. griseus* (Cohen *et al.* 1985), pMEA100 from *Nocardia mediterranei* (Moretti *et al.* 1985; Madon *et al.* 1987), pSE101 and pSE211 from *Saccharopolyspora erythraea* (Brown *et al.* 1988a,b, 1990) which are all able to integrate into the host chromosome at specific sites. pSAM2 (11kb), pSG1 (16.6kb), pMEA100 (23.7kb), pSE101 (11.3kb) and pSE211 (18.1kb) exist as low copy number, circular plasmids; whereas SLP1, pIJ110 (13.6kb), and pIJ408 (15kb) are known to occur naturally in the integrated form and only "loop out" to become autonomously replicating plasmids upon transfer.

The SLP1 plasmid attachment site (*attP*) and the *S. coelicolor* A3(2) chromosome attachment site (*attB*) were found to be virtually identical over 112bp (Omer and Cohen 1986), although only 48bp of this *att* sequence is essential for SLP1-mediated recombination (Lee *et al.* 1988). Similarly, the pIJ408 *attP* and *S. glaucescens* *attB* sites are identical over 43bp (Sosio *et al.* 1989) while the pSE211 *attP* and *S. erythraea* *attB* sites are identical over 57bp (Brown *et al.* 1990), and the pMEA100 *attP* and *N. mediterranei* *attB* sites are identical over 47bp (Madon *et al.* 1987). The integration of pSAM2 into the host chromosome is mediated by 58bp and 49bp *att* sequences in *S. ambofaciens* and *S. lividans* respectively (Boccard *et al.* 1989a).

Integration and excision genes, directly involved in plasmid recombination and maintenance, have been identified for pSAM2 and pSE211 adjacent to their *attP* sites (Boccard *et al.* 1989a,b; Omer *et al.* 1988; Brown *et al.* 1990). This strongly suggests that integration is site-specific and occurs via a recombination event between the *attP* site on the circular form of the plasmid and *attB* site of the host. Reiter *et al.* (1989) has found similarities between actinomycete *attB* sites and the termini of tRNA genes. For example, the *attB* sites for SLP1 and pMEA100 were found to have homology to the 3' terminal end of the Tyr-tRNA and Phe-tRNA genes respectively (Reiter *et al.* 1989). Similarly, Brown *et al.* (1990) have found
that the pSE101, pSE211, and pSAM2 \textit{attB} sequences correspond to the Thr-tRNA, Phe-tRNA, and Pro-tRNA genes respectively. Integration of these plasmids at the \textit{attB} site would still leave an intact tRNA (Reiter \textit{et al.} 1989; Brown \textit{et al.} 1990).

As stated, most self-transmissible \textit{Streptomyces} plasmids are able to promote chromosomal recombination when present in one strain and crossed with a plasmid-free strain, and accounts for most of the fertility observed in \textit{Streptomyces} species (Hopwood \textit{et al.} 1986a). These include SCP1 (Vivian 1971), SCP2 (Bibb \textit{et al.} 1977; Bibb and Hopwood 1981), SLP2 (Hopwood \textit{et al.} 1983), and pIJ101 (Kieser \textit{et al.} 1982) in both \textit{S. coelicolor} A3(2) and \textit{S. lividans} 66, and SLP1 (Hopwood \textit{et al.} 1983), pIJ110 (Hopwood \textit{et al.} 1984), and pIJ408 (Hopwood \textit{et al.} 1984) in \textit{S. lividans} 66. Using SLP2- SLP3- strains of \textit{S. lividans} 66 as a common genetic background, the fertility level (expressed as the proportion of recombinants out of the total output of the cross) varies with the plasmid, from $>10^{-3}$ for pIJ101 to $<10^{-5}$ for SCP1 and approximately $10^{-4}$ for SCP2*, SLP1, SLP2, pIJ110, and pIJ408. In the majority of the crosses, the plasmid was transferred to almost 100% of the \textit{potato} recipients. Thus, different levels of fertility are associated with each plasmid (Hopwood \textit{et al.} 1983). It should be noted that since SLP3 and SLP4 are transferrable without promoting chromosomal recombination (Hopwood \textit{et al.} 1983), matings between hyphae do not necessarily result in the transfer of chromosomal genes.

Plasmid-mediated fertility has been used to detect the presence of otherwise cryptic plasmids. However, testing for chromosomal recombination in the original host strain is tedious and a more efficient method is to screen for mobilization of non-self-transmissible plasmids carrying antibiotic resistance markers (Kieser \textit{et al.} 1982).
1.3 The *Streptomyces* plasmid, pIJ101.

1.3.1 Background.

pIJ101 (8.9kb) was isolated from *S. lividans* ISP 5434, together with three smaller plasmid species pIJ102 (4.0kb), pIJ103 (3.9kb), and pIJ104 (4.9kb) which appear to be naturally occurring deletion variants of pIJ101 (Kieser et al. 1982). pIJ101 is a high copy number (40-300 copies/chromosome), broad host range plasmid (Kieser et al. 1982). Derivatives of pIJ101 are known to replicate in other actinomycete genera such as *Micromonospora* (Matsushima and Baltz 1988), *Amycolatopsis* (Matsushima et al. 1987), and *Thermomonospora* (Pidcock et al. 1985). pIJ101 is self-transmissible by conjugation, elicits the lethal zygosis reaction (Ltz+), and promotes chromosomal recombination at high frequencies in *S. lividans* and *S. coelicolor* A3(2) (Kieser et al. 1982). However pIJ102, pIJ103, and pIJ104 are non-conjugative and Ltz−. A restriction map of pIJ101 is shown in Fig. 1.2. More is currently known about the biology of pIJ101 than for any other *Streptomyces* plasmid. Recently, Kendall and Cohen (1988) have determined the complete nucleotide sequence of pIJ101 (Appendix I) and analyzed it for open reading frames (ORFs) with respect to the existing information about the genetic properties of the plasmid. The G+C content of pIJ101 was found to be similar to the chromosomal DNA, that is approximately 73% (Kendall and Cohen 1988).

1.3.2 Replication and stability determinants.

A 2.2kb *Sst*I(16)-*Sst*I(63) region on pIJ101 was found to be essential for maintenance and replication (Fig. 1.2), since the insertion of foreign DNA at four different sites within this segment [(that is, *Bam*HI(1), *Xho*I(3), *Sal*GI(7), and *Bcl*II(10)) destroyed the ability of the plasmid to exist in the host (Kieser et al. 1982). This 2.2kb *Sst*I fragment, when attached to a selectable marker and circularized, defines the limit of the basic replicon as it is capable of autonomous replication (Kieser et al. 1982). This essential region was found to
Figure 1.2
Physical and genetic map of the S. lividans plasmid, plJ101.

The map is based on compiled studies by Kieser et al. (1982); Kendall and Cohen (1987, 1988); and Stein et al. (1989). Restriction enzyme sites are numbered clockwise from the unique BamHI site (site 1). Open reading frames (ORFs) are indicated by open boxes, and arrow-heads show the direction of translation. All the cited genes are explained in the text. The location of promoters and their direction of transcription are indicated by filled boxes and arrows respectively. The position of the site for second-strand synthesis (stf) is boxed.
contain one large ORF (rep; 450 amino acids), a second smaller ORF (orf56), and a noncoding region that may be the origin of replication (Kendall and Cohen 1988). It was not known at the start of the research described in this thesis whether both the Rep and Orf56 proteins were needed for replication, or if the Rep protein alone was sufficient in pIJ101.

Gruss and Ehrlich (1989) have suggested that the pIJ101 origin lies upstream of the rep ORF (that is, between bp 1327 and 1269 in Appendix I), since this region bears structural similarity to the Staphylococcus aureus plasmid pC194 origin of replication which lies 5' to its rep gene (Gros et al. 1987). Furthermore, Gruss and Ehrlich (1989) have suggested that the predicted amino acid motif of the pIJ101 Rep protein active site is similar to those found in several other plasmid-encoded Rep proteins sharing pC194-like consensus replication origins (Fig. 1.3).

The homology between the Rep proteins and plus origins of these plasmids suggests that they possess a similar mechanism of plasmid replication. The staphylococcal plasmids pT181-encoded RepC and pC221-encoded RepD proteins have been shown to possess replication-initiator, sequence-specific endonuclease, and topoisomerase-like activities for their respective origins which are located within the 5' end of the rep coding region (Koepsel et al. 1985; Thomas et al. 1990b). Furthermore, Thomas et al. (1990b) have demonstrated that the site/strand-specific cleavage reaction at the origin is accompanied by the formation of a phosphoryl-tyrosine linkage between RepD and oriD, resulting in the RepD protein being covalently bound to the 5' end of the nicked strand. This is analogous to the action of the φX174 gene A protein which initiates rolling circle replication by introducing a site-specific nick on one strand of the replicative form (RF) DNA and covalently attaching to the 5'-phosphate terminus via a tyrosyl residue, while the 3'-hydroxyl terminus is free to serve as a primer for DNA elongation (van Mansfeld et al. 1986).
Figure 1.3 Putative active site of pIJ101 Rep protein

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Rep active site</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIJ101</td>
<td>ALA GLU TYR ILE ALA LYS THR GLN ASP</td>
<td><em>S. lividans</em> (72% GC)</td>
</tr>
<tr>
<td>ØX174</td>
<td>ALA LYS TYR VAL ASN LYS LYS SER ASP</td>
<td><em>E. coli</em> (51% GC)</td>
</tr>
<tr>
<td>pBAA1</td>
<td>SER LYS TYR PRO VAL LYS ASP THR ASP</td>
<td><em>B. subtilis</em> (42% GC)</td>
</tr>
<tr>
<td>pC194</td>
<td>ALA LYS TYR SER GLY LYS ASP SER ASP</td>
<td><em>S. aureus</em> (34% GC)</td>
</tr>
<tr>
<td>pUB110</td>
<td>ALA LYS TYR PRO VAL LYS ASP THR ASP</td>
<td><em>S. aureus</em> (34% GC)</td>
</tr>
</tbody>
</table>

Enzymatic active site of ØX174 Rep protein compared to Rep proteins from plasmids pIJ101, pBAA1, pC194, and pUB110 (Gruss and Ehrlich 1989). The common Tyr residue thought to be involved in covalent binding to the 5' phosphate of the nicked strand in the origin is highlighted. 1: Kendall and Cohen (1988); 2: Gruss *et al.* (1987); 3: Devine *et al.* (1989).
A 0.4kb PstI(59)-SstII(63) region, situated next to the 2.2kb SstII(16)-SstII(63) essential replicon, may be required for the stable inheritance of the plasmid. Deletions made within this region produced markedly unstable plasmids which were lost from the host at a rate of 13 to 55% after one round of growth on antibiotic-free media (Kieser *et al.* 1982). It is not known whether this region has a partition function similar to that identified in SCP2* (Bibb and Hopwood 1981) which would be involved in promoting stable inheritance of the plasmid.

1.3.3 Transfer and spread determinants.

*In vitro* insertion and deletion studies genetically defined the pIJ101 transfer region to a 1.4kb segment on the plasmid (Kieser *et al.* 1982). Located next to the transfer region is a 1.7kb segment required for the spread of the plasmid within plasmid-free mycelia to produce normal pocks (Ltz+ phenotype) (Kieser *et al.* 1982). Disruption of the transfer (tra) region resulted in no pock production (Ltz- phenotype), since the plasmid is unable to transfer itself from one mycelium to another. Disruption of the spread (spd) region produced small pocks (Ltz+/- phenotype), since the plasmid is unable to migrate within the recipient mycelia although transfer still takes places. Since Ltz+/- pIJ101 derivatives were not induced to spread over the normal distance by the simultaneous presence of Ltz+ plasmids, spreading may involve a *cis*-acting function (Kieser *et al.* 1982). Similarly, transfer may also involve *cis*-acting functions since Ltz- pIJ101 derivatives were not transferred in the presence of Ltz+ plasmids (although Ltz- plasmids were mobilised at low frequencies) (Kieser *et al.* 1982).

Although the pIJ101 essential region was defined to a 2.2kb SstII(16)-SstII(63) segment, the ability of the plasmid to survive was lost when foreign DNA was inserted into a unique BglII(45) site located at a considerable distance from the essential region (Kieser *et al.* 1982). However, functional plasmids were obtained when the BclI fragment containing this BglII(45) site was deleted (Kieser *et al.* 1982). This situation was similar to that observed in the Gram-negative plasmid
RK2 in which the "kill" (kil) functions of the plasmid were lethal to the host in the absence of complementing "kor" (kil override) functions (Figurski et al. 1982). Apart from tra and spd, genetic analysis of pIJ101 identified four more plasmid-encoded loci (kilA, kilB, korA, and korB) thought to be involved in gene transfer and pock formation; the BglII(45) site mentioned above lying within the korA locus (Kendall and Cohen 1987).

The kilA and kilB genes, so called because of their lethal effects, could not be cloned into S. lividans on a minimal pIJ101-based replicon unless suitable kor genes were present in either cis or trans (Kendall and Cohen 1987). Both kil genes are involved in plasmid transfer since insertions into the kilA locus abolished pocking and insertions into the kilB locus produced only very small pocks (Kendall and Cohen 1987). As stated, insertions into the tra and spd loci produced no pocks and small pocks respectively (Kendall and Cohen 1987). Thus, kilA and tra are directly involved in intermycelial transfer whereas kilB and spd are involved in intramycelial spread.

From sequence analysis of pIJ101, only one ORF (termed tra) was found to span the genetically defined kilA and tra loci (Fig. 1.2), and would code for a single 621aa protein of 77kDa in size (Kendall and Cohen 1988). The predicted Tra protein has no significant hydrophobic domains (suggesting that it is not associated with membranes), and has very little homology with other DNA-binding proteins (Kendall and Cohen 1988). However, the putative Tra protein does possess a region which resembles the nucleotide-binding fold domains of ATP-requiring enzymes, suggesting that ATP may be needed for the tra function (Kendall and Cohen 1988). The spd locus was found to contain two ORFs which were called spdA (146aa) and spdB (274aa) respectively (Fig. 1.2) (Kendall and Cohen 1988). The kilB locus contains one ORF (177aa) and is located close to the spd ORFs (Fig. 1.2) (Kendall and Cohen 1988). The predicted SpdB protein contains three strongly hydrophobic domains at the amino-terminal end, suggesting that it may be associated with membranes (Kendall and Cohen 1988). The putative SpdA and
KilB proteins each contain one strongly hydrophobic domain (Kendall and Cohen 1988).

KorA is able to override the lethal effects of both tra (formerly kilA) and kilB, whereas KorB can only override the effects of the kilB gene (Kendall and Cohen 1987). KorB acts by negatively regulating transcription from its own promoter and from the kilB promoter, while KorA represses transcription from the overlapping divergent promoters of the korA and tra genes (Kendall and Cohen 1987; Stein et al. 1989; Stein and Cohen 1990). However, KorA is unable to repress transcription from the kilB promoter, suggesting that korA does not override the lethal effects of kilB at the transcriptional level (Stein et al. 1989).

Although ORFs have been found for the korA (241aa) and korB (80aa) genes (Fig. 1.2), the derived amino acid sequences show no homology to any other known repressor proteins (Kendall and Cohen 1988). However, the predicted amino acid sequence for korA contains an α-helix-turn-α-helix motif similar to the DNA-binding domains of many prokaryotic repressor proteins (Kendall and Cohen 1988). The predicted korB amino acid sequence has two such possible α-helix-turn-α-helix motifs contained within its structure (Kendall and Cohen 1988). Recently, Stein and Cohen (1990) have expressed the 31kDa KorA and 10kDa KorB protein products in E.coli and shown that these proteins are functionally active and able to control transcription from the tra and kilB promoters in vivo, respectively.

A tentative model for plasmid transfer proposes that after the initial fusion of hyphae which is directed by host genes, a cis-acting plasmid function then allows plasmid transfer to take place (Hopwood et al. 1986a). Once transferred, the plasmid then spreads within the hyphae using self-encoded specific functions - one of which may permit the plasmid to cross the transverse septa that it would encounter (Hopwood et al. 1986a). During plasmid transfer in conjugal matings, the absence of kor gene products in the recipient plasmid-free cell may cause the transient expression of newly introduced kil genes (Hopwood et al. 1986a). This would lead to a localized abnormal cell development in the lawn of recipient
plasmid-free cells and produce a pock (Hopwood et al. 1986a). A specific plasmid region promoting fertility has not yet been identified for pIJ101.

1.3.4 Sti (strong incompatibility) and Cop (copy number) determinants

sti has been defined as a non-coding region of DNA which causes strong incompatibility when present in its natural orientation with respect to the basic replicon region of pIJ101 (Deng et al. 1988). For two plasmids to co-exist in the same host, they must both possess sti in the correct orientation (Sti+ phenotype), or both lack sti or possess it in the reverse orientation (Sti- phenotype). Sti+ and Sti- plasmids cannot co-exist in the same cell, in which case the Sti+ plasmid is retained while the Sti- plasmid is lost. Furthermore, when sti is deleted, significant amounts of single-stranded (ss) DNA accumulates in the cell, implicating sti as the site where second-strand synthesis is initiated. As the deletion of sti is not lethal to the plasmid, an alternative lagging-strand initiation site(s) must exist although it would be utilized less efficiently than the primary site, leading to the accumulation of ssDNA in the cell.

The accumulation of ssDNA corresponding to one strand of a plasmid monomer has been discovered as a natural replication intermediate in plasmids isolated from Bacillus subtilis (te Riele et al. 1986a; Devine et al. 1989), Staphylococcus aureus (te Riele et al. 1986a; Gruss et al. 1987; Boe et al. 1989), and Streptococcus pneumoniae (del Solar et al. 1987). It has thus been suggested that pIJ101, like these other Gram-positive plasmids, replicates via a ssDNA intermediate and a rolling circle mechanism of plasmid replication has been proposed for these replicons (Koepsel et al. 1985; te Riele et al. 1986a,b; del Solar et al. 1987; Gros et al. 1987; Gruss et al. 1987; Khan et al. 1988; Devine et al. 1989; Gruss and Ehrlich 1989; Novick 1989; Thomas et al. 1990b).

Rolling circle replication requires three plasmid-encoded elements: a plus origin, the replication protein (Rep), and a minus origin (sti) (Fig. 1.4). Thus, the Rep protein of pIJ101 would initiate transcription by introducing a strand-specific
nick at the plus origin, exposing a 3' hydroxyl group (Fig. 1.4, step 1). The existing plus strand is displaced and the new plus strand is synthesized by 3' hydroxyl extension from the nick using the minus strand as a template (step 2). The displaced plus strand probably retains the Rep protein linked to the 5'-phosphate end. The Rep protein then recognizes a termination sequence that overlaps with the plus origin and introduces a second nick to generate one fully replicated strand and a ssDNA monomer of the displaced strand (step 3). Finally, Rep protein ligates the two ends to yield a new double-stranded plasmid and a free circular single-stranded intermediate (step 4). The single-stranded intermediate is then converted to the double-stranded form when second-strand synthesis is initiated at the sti site present on the ssDNA (step 5). The formation of double-stranded plasmid DNA marks the completion of one cycle of replication in which two daughter plasmids are generated from one parent plasmid (step 6). However, if the single-stranded plus strand in step 4 is not efficiently converted to double-stranded plasmid DNA, replication will be non-productive and result in the accumulation of ssDNA.

Since pT181 replication is sensitive in vivo to the DNA polymerase III inhibitor 5-hydroxylphenyl azouracil (Majumder and Novick 1988), plasmid replication at the plus origin may require a host-dependent multifactor replication complex (consisting of the DNA polymerase III elongation system, the host Rep protein, and SSB proteins) analogous to that used for the RF-RF replication of the single-stranded coliphages (Baas and Jansz 1988). Unlike the φX174-type ssi signals which require the assembly of an n'-dependent primosome complex or the phage G4-type ssi signals which require dnaG primase for the formation of an RNA primer (Baas and Jansz 1988), replication at the minus origin in pC194 and pUB110 may be initiated by RNA polymerase since the addition of rifampin blocks conversion of ssDNA to double-stranded plasmid DNA (Boe et al. 1987). This is similar to the filamentous ssDNA phages, where synthesis of an RNA primer at the minus origin is mediated by RNA polymerase (Baas and Jansz 1988).
Figure 1.4  Stages of rolling circle replication

Step 1
parent ds plasmid

Step 2

Step 3

Step 4

Step 5

Step 6

daughter double-stranded plasmid

ssDNA intermediate

daughter double-stranded plasmid
Cop may be a trans-acting negative regulator that affects the copy number of Sti+ plasmids by inhibiting the initiation of second-strand synthesis, as deletions of the \( \text{cop} \) region increases plasmid copy number to approximately 1000/cell (Deng et al. 1988). Since the \( \text{cop} \) and \( \text{korB} \) genes map at the same position on pIJ101, this suggests that one protein may be controlling both plasmid replication (through interactions with \( \text{sti} \)) and plasmid transfer (through interactions with \( \text{kilB} \)).

### 1.3.5 Identification of pIJ101 promoters and terminators.

A total of five promoters have been localized on pIJ101 (Fig. 1.2; Appendix I). The promoters for the \( \text{kilB} \) gene and the \( \text{tra} \) gene, named pIJ101A and pIJ101B respectively (Buttner and Brown 1985, 1987), have been confirmed to lie upstream of their respective ORFs (Kendall and Cohen 1988; Stein et al. 1989). Transcription from pIJ101B (p\( \text{tra} \)) is thought to extend from \( \text{tra} \) through the clustered \( \text{spdA}, \text{spdB} \), and \( \text{orf66} \) reading frames (Kendall and Cohen 1988). However, the pIJ101C promoter (Deng et al. 1986) was found to lie in the middle of the \( \text{cop/korB} \) ORF and its function is therefore unknown (Kendall and Cohen 1988). Promoters located upstream of the \( \text{korA} \) and \( \text{cop/korB} \) ORF have also been identified (Stein et al. 1989). Interestingly, \( \text{korA} \) and \( \text{tra} \) are transcribed from overlapping divergent promoters; the p\( \text{tra} \) -35 region is located between the p\( \text{korA} \) -10 and -35 regions while the p\( \text{korA} \) -35 region is located between the p\( \text{tra} \) -10 and -35 regions (Appendix I). The initiation of transcription for the \( \text{cop/korB} \) promoter coincides with the translational start codon of the \( \text{cop/korB} \) ORF in \( S. \text{lividans} \) (Stein et al. 1989). Thus, this indicates that the \( \text{cop/korB} \) mRNA may be translated in the absence of a conventional ribosome-binding site, as in the case of the \( \text{Streptomyces} \, \text{neomycin (aph)} \) and \( \text{erythromycin (ermE)} \) genes (Bibb et al. 1985a,b) and the \( \text{E. coli} \, \lambda \, \text{cl} \) and Tn1721 \( \text{tetR} \) genes (Ptashne et al. 1976; Klock and Hillen 1986).

The \( \text{tra}, \text{kilB}, \text{korA} \), and \( \text{korB} \) promoters all have sequence homology to the consensus \( \text{E. coli} \) promoter sequence (Hawley and McClure 1983). Furthermore,
both pkilB (pIJ101A) and ptra (pIJ101B) are able to promote transcription in *E. coli* (Buttner and Brown 1987). Thus, these four promoters may be transcribed in *Streptomyces* by the RNA polymerase-σ^{35} complex which is able to initiate transcription from *E. coli*-type promoters (Jaurin and Cohen 1984; Westpheling et al. 1985; Buttner 1989).

A region of dyad symmetry, which has already been shown to terminate transcription in both *Streptomyces* and *E. coli* (Deng et al. 1987), lies between the convergently transcribed korA and cop/korB genes and may act as a transcriptional terminator for either or both of these genes (Kendall and Cohen 1988). Similar regions of dyad symmetry were found downstream of the rep, kilB, orf56, orf85, and tral/spdA/spdB/orf66 ORFs (Kendall and Cohen 1988).

### 1.4 *Streptomyces* plasmids as cloning vectors.

#### 1.4.1 Construction of cloning vectors based on pIJ101.

Derivatives of pIJ101 have been constructed which contain several drug resistant determinants (Kieser et al. 1982). For example, pIJ350 (4.1kb) was constructed from pIJ102 by the replacement of a BclI fragment with a BclI thiostrepton resistance (tsr) fragment from *S. azureus* (Kieser et al. 1982). pIJ702 was constructed by inserting the tyrosinase (*mel*) gene from *S. antibioticus* IMRU 3720 into the BclI site adjacent to the tsr gene in pIJ350 (Katz et al. 1983); insertions into the *mel* gene disrupts the ability of transformants to convert tyrosine into melanin and white colonies (instead of black colonies) are formed when grown on agar supplemented with tryptone and copper ions.

pIJ702 has been used extensively as a cloning vector in *Streptomyces* (for a review see Tomich 1988; Chater 1990). Reported applications of pIJ702 include the cloning of genes involved in the undecylprodigiosin, clavulanic acid, and tetracenomycin biosynthesis (Feitelson and Hopwood 1983; Bailey et al. 1984; Motamedi and Hutchinson 1987). Several extracellular enzymes have also been cloned using pIJ702, such as the β-lactamase gene from *S. albus* and *S. cacaoi*
(Dehottay et al. 1986; Lenzini et al. 1987), the protease A and B genes from *S. griseus* (Henderson et al. 1987), and the α-amylase gene from *S. hygroscopicus* (Hoshiko et al. 1987). Both the proteases and α-amylase are commercially produced. There have been two interesting reports where eukaryotic genes have been cloned and expressed in *Streptomyces* using pIJ702; namely the bovine growth hormone gene (Gray et al. 1984) and the human interferon α2 gene (Pulido et al. 1986). At Smith Kline and Beecham, the T-cell CD4 surface protein (which is the receptor for HIV binding) has been efficiently expressed in *S. lividans* and active protein is secreted into the culture medium at 100mg/l (personal communication).

1.4.2 Plasmid structural instability.

Plasmid structural instability in *Streptomyces* has been encountered in both naturally occurring and recombinant plasmids, but little is known about the mechanism of rearrangement that occurs in these types of plasmids. As stated, pIJ101 coexists with three smaller plasmid species (that is, pIJ102, pIJ103, and pIJ104) that seem to be naturally occurring deletion variants of pIJ101 (Kieser et al. 1982).

Recombinant plasmids may either be poorly maintained (seggregational instability) or be rearranged (structural instability). Nakano et al. (1984) observed that after successive subcultivations of *S. lavendulae* containing the plasmid pSL1 (3.9kb), a spontaneous tandem duplication of 900bp on the plasmid led to the appearance of a second larger derivative of pSL1, called pSL2 (4.8kb). Furthermore, sequence analysis of pSL2 suggested that the duplication occurred by recombination between short direct repeats of 5bp. A similar event has also been reported by Chen et al. (1987) for the *E. coli-Streptomyces* shuttle vector pIF132 in *S. lividans* 66, where intraplasmid recombination between two direct *mel* repeats on pIF132 resulted in the appearance of a smaller plasmid derivative, pIF138.
Several other examples of structural instability of *E. coli*-Streptomyces hybrid vectors in *Streptomyces* have also been reported (Kieser et al. 1982; Chater et al. 1982; Lee et al. 1986; Pigac et al. 1988). Generally, the hybrid vectors were found to be stable in *E. coli* but underwent rearrangement when transformed into *Streptomyces*. Lee et al. (1986) showed that while the hepatitis B viral surface antigen (HBsAg) was stably maintained in the *Streptomyces* plasmid pIJ702, insertion of the *E. coli* pUC12 plasmid either alone or in conjunction with the HBsAg gene induced deletions when transformed into *S. lividans* 1326. The deletions were found to occur within the *mel* gene of pIJ702, resulting in the loss of part of the *mel* sequence and the HBsAg insert.

Pigac et al. (1988) reported that pZG1, a hybrid plasmid consisting of the *E. coli* plasmid pBR322 and the *S. lividans* plasmid pIJ350, was structurally unstable when introduced into *S. lividans* 1326 and *S. rimosus* R6 and resulted in derivatives of pZG1 which lacked parts of the pBR322 sequence. Both Lee et al. (1986) and Pigac et al. (1988) also observed that the size of the deletions differed depending on the age of the transformant colony chosen for analysis. Furthermore, Pigac et al. (1988) suggested that the structural instability of the pBR322:pIJ350 hybrid plasmid in *Streptomyces* was due to the presence of ssDNA. Such ssDNAs are likely to be highly recombinogenic and may lead to illegitimate recombination resulting in plasmid rearrangements (Pigac et al. 1988). Inefficient conversion of single-stranded replication intermediates has also been reported to underlie plasmid structural instability in *Staphylococcus aureus* (Gruss et al. 1987) and *Streptococcus pneumoniae* (del Solar et al. 1987).

1.5 Aims of the project.

The main aim of the project was to investigate the biological function of the *cop/korB* gene from the *S. lividans* plasmid pIJ101. This protein is potentially interesting to study because it has been assigned two distinct functions: firstly, as a negative regulator of the *kilB* gene (Kendall and Cohen 1987) and secondly, as a
repressor of plasmid copy number through the *sti* determinant (Deng et al. 1988). The examination of the regulatory interactions of *cop/korB* may elucidate mechanisms involved in plasmid transfer and copy number control in pIJ101. The long term objective is to understand more about the biology of pIJ101 in order to gain information which could lead to the construction of more efficient cloning vectors.
CHAPTER TWO. Materials and Methods.

2.1 Materials.

All chemical reagents used in this thesis were obtained from BDH, Fisons, or Sigma, and were of analytical grade or the equivalent. Bacteriological reagents were purchased from Difco or Oxoid.

Ampicillin and kanamycin were purchased from Sigma. Thiostrepton was a kind gift from SJ Lucania (ER Squibb and Sons Inc, NJ, USA). Table 2.1 shows the stock and working concentrations of antibiotics used in liquid and solid media.

Restriction enzymes were either made by Dr Linda Wallace or purchased from Anglian Biotechnology. L-[\textsuperscript{35}S]methionine, [\alpha-\textsuperscript{35}S]dATP, [\alpha-\textsuperscript{32}P]dNTP, and [\gamma-\textsuperscript{32}P]ATP were purchased from DuPont.

2.2 Bacterial strains and plasmids.

The characteristics of \textit{E. coli} JM107 and \textit{S. lividans} TK24 are shown in Table 2.2. Plasmids used in this thesis are shown in Table 2.3. pUC8, pUC18/19, pBluescript II (KS+), pKK223.3, and pBGS19- were used for cloning in \textit{E. coli} JM107. pIJ702 was used for cloning in \textit{S. lividans} TK24. A list of the plasmids constructed during the course of this research is shown in Table 2.4 (according to figure order).

2.3 Culture conditions.

\textit{E. coli} JM107 was grown on nutrient agar plates or nutrient broth overnight at 37°C. These media were supplemented with the appropriate antibiotic (Table 2.1), IPTG (40ug/ml) and X-gal (80ug/ml) when necessary.

\textit{S. lividans} TK24 was grown on R2YE agar (Hopwood \textit{et al.} 1985b) or malt extract yeast extract agar (2.4% w/v Difco malt extract, 0.5% w/v Oxoid yeast extract, 2% w/v Bactoagar) at 30°C until sporulation (typically one week). For liquid cultivation, \textit{S. lividans} TK24 was grown in YEME (Hopwood \textit{et al.} 1985b).
Figure 2.1
Antibiotic concentrations for selecting resistant strains.

<table>
<thead>
<tr>
<th>Antibiotic stock solution</th>
<th>Final concentration (ug/ml)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E. coli solid</td>
<td>E. coli liquid</td>
<td>S. lividans solid</td>
<td>S. lividans liquid</td>
<td></td>
</tr>
<tr>
<td>Ampicillin (50)</td>
<td>500</td>
<td>50</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Kanamycin (25)</td>
<td>25</td>
<td>2.5</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Thiostrepton (50 in DMSO)</td>
<td>-</td>
<td>-</td>
<td>50</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* conc. in soft agar overlays.
**Figure 2.2**

**Characteristics of bacterial strains used.**

<table>
<thead>
<tr>
<th>Bacterial strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> JM107</td>
<td>Δ(lac-proAB) endA1 gyrA96 thi hsdR17 supE44 relA1 λ- [F' traD36 proAB+ lacI9ZΔM15]</td>
<td>Yanisch-Perron <em>et al.</em> (1985)</td>
</tr>
<tr>
<td><em>S. lividans</em> TK24</td>
<td>str-6 SLP2- SLP3-</td>
<td>Hopwood <em>et al.</em> (1983)</td>
</tr>
</tbody>
</table>

**Genotype abbreviations:**

- **endA** : mutation in DNA endonuclease gene allowing improved isolation of plasmid DNA.
- **gyrA** : mutation in DNA gyrase subunit A gene.
- **hsdR** : mutation in restriction gene permitting introduction of foreign DNA into the host.
- **lacI9** : mutation in repressor protein of *lac* operon resulting in a high binding-affinity repressor protein.
- **lacZΔM15** : deletion in host-encoded β-galactosidase gene which is complemented by plasmid-encoded α-segment to produce enzymatically active protein.
- **proAB** : deletion of proline gene.
- **relA** : phenotype permitting RNA synthesis in the absence of protein synthesis.
- **str** : streptomycin resistance.
- **supE** : suppressor of amber (UAG) mutations.
- **thi** : mutation in thiamine gene.
- **traD** : mutation in transfer gene inactivates conjugal transfer of F' plasmid.
### Figure 2.3

**Characteristics of plasmids used.**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC8 (2.7kb)</td>
<td>Ap&lt;sup&gt;r&lt;/sup&gt;, LacZ&lt;sup&gt;'&lt;/sup&gt;</td>
<td>Vieira and Messing 1982</td>
</tr>
<tr>
<td>pUC18/19 (2.7kb)</td>
<td>Ap&lt;sup&gt;r&lt;/sup&gt;, LacZ&lt;sup&gt;'&lt;/sup&gt;</td>
<td>Yanisch-Perron et al. 1985</td>
</tr>
<tr>
<td>pBluescript II (3.0kb)</td>
<td>Ap&lt;sup&gt;r&lt;/sup&gt;, LacZ&lt;sup&gt;'&lt;/sup&gt;</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pKK223.3 (4.6kb)</td>
<td>Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>pBGS19- (4.4kb)</td>
<td>Km&lt;sup&gt;r&lt;/sup&gt;, LacZ&lt;sup&gt;'&lt;/sup&gt;</td>
<td>Spratt et al. 1986</td>
</tr>
<tr>
<td>pIJ101 (8.9kb)</td>
<td>-</td>
<td>Kieser et al. 1982</td>
</tr>
<tr>
<td>pIJ303 (10.8kb)</td>
<td>Tsr&lt;sup&gt;r&lt;/sup&gt;,</td>
<td>Kieser et al. 1982</td>
</tr>
<tr>
<td>pIJ702 (5.8kb)</td>
<td>Tsr&lt;sup&gt;r&lt;/sup&gt;, Mel&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Katz et al. 1983</td>
</tr>
</tbody>
</table>

**Abbreviations:**
- Ap<sup>r</sup>: ampicillin resistance
- Km<sup>r</sup>: kanamycin resistance
- Tsr<sup>r</sup>: thiostrepton resistance
- Mel: melanin production
### List of plasmid constructs. (cont’d overleaf)

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQR200</td>
<td>1.2kb BclII-49-BclII(57) cop/korB, sti, orf85/79 fragment in pUC8 BamHI site</td>
</tr>
<tr>
<td>pQR206</td>
<td>0.5kb BclII-49-SpeI(53) cop/korB fragment in pUC8 BamHI/EcoRI sites</td>
</tr>
<tr>
<td>pQR412</td>
<td>0.5kb BclII(49)-SpeI(53) cop/korB fragment in pUC8 Smal site (in both</td>
</tr>
<tr>
<td>pQR413</td>
<td>orientations)</td>
</tr>
<tr>
<td>pQR414</td>
<td>0.5kb BclII(49)-SpeI(53) cop/korB fragment in pKK223.3</td>
</tr>
<tr>
<td>pQR427</td>
<td>2.8kb BclII(10)-BclII(29) spdB, orf66, kilB, orf56 fragment in pUC19</td>
</tr>
<tr>
<td>pQR428</td>
<td>BamHI site</td>
</tr>
<tr>
<td>pQR456</td>
<td>0.8kb Ball(15)-SalGI(19) kilB fragment in pUC19 Smal/SalGI sites</td>
</tr>
<tr>
<td>pQR452</td>
<td>0.5kb Sau3A cop/korB fragment in pUC19 BamHI site</td>
</tr>
<tr>
<td>pQR417</td>
<td>0.7kb SpeI(53)-BclII(57) sti fragment in pUC19 XbaI/BamHI sites</td>
</tr>
<tr>
<td>pQR437</td>
<td>0.6kb SpeI(53)-SstI(56) sti fragment in pUC19 XbaI/SalI sites</td>
</tr>
<tr>
<td>pQR438</td>
<td>0.53kb SpeI(53)-SstI(55) sti fragment in pUC19</td>
</tr>
<tr>
<td>pQR458</td>
<td>0.5kb BclII(49)-FspI(52) cop/korB fragment in pUC18 BamHI/Smal sites</td>
</tr>
<tr>
<td>pQR459</td>
<td>0.45kb XhoII cop/korB fragment in pUC18 BamHI site (in both orientations)</td>
</tr>
<tr>
<td>pQR460</td>
<td>pQR457</td>
</tr>
<tr>
<td>pQR457</td>
<td>0.73kb Ball(15)-FspI(18) kilB fragment in pUC19 Smal site</td>
</tr>
<tr>
<td>pQR461</td>
<td>1.1kb BglII(45)-SpeI(53) cop/korB fragment in pUC19 BamHI/XbaI sites</td>
</tr>
<tr>
<td>pQR410a</td>
<td>pQR200 in pJJ702 PstI site</td>
</tr>
<tr>
<td>pQR410b</td>
<td>sti&lt;sup&gt;Rev&lt;/sup&gt;cop&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>pQR411a</td>
<td>1.2kb BclII(49)-BclII(57) fragment in pJJ702 PstI site</td>
</tr>
<tr>
<td>pQR411b</td>
<td>sti&lt;sup&gt;Rev&lt;/sup&gt;cop&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>pQR421a</td>
<td>pQR417 in pJJ702 PstI site</td>
</tr>
<tr>
<td></td>
<td>sti+cop&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>sti+cop&lt;sup&gt;-&lt;/sup&gt;</td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>pQR421b</td>
<td>0.7kb SpeI(53)-Bcll(57) fragment in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR422a</td>
<td>pQR417 in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR422b</td>
<td>0.7kb SpeI(53)-Bcll(57) fragment in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR441a</td>
<td>pQR437 in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR441b</td>
<td>0.6kb SpeI(53)-SstI(56) fragment in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR442a</td>
<td>pQR437 in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR443a</td>
<td>pQR438 in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR443b</td>
<td>0.53kb SpeI(53)-SstI(55) fragment in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR444a</td>
<td>pQR438 in pIJ702 PstI site</td>
</tr>
<tr>
<td>pQR420</td>
<td>2.2kb SstI(16)-SstI(63) rep, orf56 fragment in pBGS19-</td>
</tr>
<tr>
<td>pQR430</td>
<td>2.0kb Bcll(15)-SstI(63) rep fragment in pBGS19-</td>
</tr>
<tr>
<td>pQR431a/b</td>
<td>1.4kb NotI(12)-SstI(63) rep fragment in pBGS19-</td>
</tr>
<tr>
<td>pQR433</td>
<td>1.0kb Bcll tsr fragment in pUC19 SacI site</td>
</tr>
<tr>
<td>pQR434</td>
<td>pQR420 + 1.0kb tsr fragment from pQR433</td>
</tr>
<tr>
<td>pQR435</td>
<td>pQR430 + 1.0kb tsr fragment from pQR433</td>
</tr>
<tr>
<td>pQR436</td>
<td>pQR431a + 1.0kb tsr fragment from pQR433</td>
</tr>
<tr>
<td>pQR445</td>
<td>1.7kb Apal(14)-SstI(63) rep fragment in pBluescript</td>
</tr>
<tr>
<td>pQR446</td>
<td>1.7kb Apal(14)-SstI(63) rep fragment and 1.0kb tsr fragment in pBGS19-</td>
</tr>
<tr>
<td>pQR453</td>
<td>pQR438 (Fig. 3.17) in pBGS19- HindIII site</td>
</tr>
<tr>
<td>pQR455</td>
<td>pQR435 (Fig. 4.4) + 0.53kb sti fragment from pQR453</td>
</tr>
</tbody>
</table>
for 2-3 days or in malt extract peptone (MEP) media (2% w/v glycerol, 1% w/v Oxoid malt extract, 1% w/v Difco Bacto-peptone, 0.174% w/v K$_2$HPO$_4$) overnight at 30°C with good aeration. These media were supplemented with thiostrepton when required (Table 2.1).

*(CA Wrigley-Jones, personal communication)*

2.4 Maintenance of bacterial strains.

*E. coli* JM107 was maintained on M9 minimal media plates supplemented with thiamine (10ug/ml) and 1% (w/v) glucose as a carbon source. *E. coli* JM107 harbouring plasmids were streaked onto nutrient agar slopes and kept at room temperature, or were stored as cell suspensions in 20% (v/v) glycerol at -70°C. *Streptomyces* strains were maintained as spore suspensions in 20% (v/v) glycerol at -70°C.

2.5 Preparation and transformation of *E. coli* JM107 competent cells.

This procedure is adapted from Morrison (1979) and Hanahan (1983). A single *E. coli* colony taken from a minimal media plate was used to inoculate 5ml nutrient broth and grown overnight at 37°C with shaking. 1ml of this overnight culture was transferred to 200ml nutrient broth supplemented with 20mM MgCl$_2$, and incubated for 1-2hrs at 37°C with shaking until an OD$_{550}$=0.2 was reached. The cells were then harvested by centrifugation (8000rpm, 10min), the pellet resuspended in 100ml ice-cold 75mM CaCl$_2$, 15% (v/v) glycerol and respun (8000rpm, 10min). The pellet was resuspended in 20ml ice-cold 75mM CaCl$_2$, 15% (v/v) glycerol, dispensed in 0.5ml aliquots, and stored at -70°C until needed.

For transformation, one aliquot (0.5ml) of *E. coli* competent cells was thawed, mixed with plasmid DNA in up to 50ul TE buffer (10mM Tris/HCl pH7.5, 1mM EDTA), and left on ice for 45min. The cells were then incubated at 37°C for 10min, transferred to 5ml nutrient broth and grown for 2hrs at 37°C with shaking. 100ul of this transformation mixture was plated onto the appropriate selective media and left
at 37°C overnight. Typically, transformation frequencies of between $10^6$ to $10^8$ transformants/ug of plasmid DNA were obtained.

2.6 Small-scale isolation of plasmid DNA from *E. coli*.

Small-scale isolation of plasmid DNA was performed essentially as described by Birnboim and Doly (1979). 1.5ml of an *E. coli* overnight culture was placed in a 1.5ml Eppendorf microfuge tube and spun in a MSE microfuge (13000rpm, 10min). The supernatant was discarded and the pellet resuspended in 100ul Birnboim buffer (0.9% w/v glucose, 10mM EDTA, 25mM Tris/HCl pH7.5). 200ul 0.2M NaOH, 1% (w/v) SDS was added and the mixture vortexed. 150ul 4M Na acetate pH6.0 was then added, the mixture vortexed and spun (13000rpm, 10min). The supernatant was transferred to a clean Eppendorf tube and extracted once with an equal volume of phenol/chloroform (500g phenol, 200ml distilled water, 30ml 1M Tris/HCl pH7.4, 15ml 2M NaOH, 60ml 0.2M EDTA, 500ml chloroform), and nucleic acids in the aqueous phase precipitated with 1/10 volume 4M Na acetate pH6 and two volumes absolute alcohol at -20°C for 20min. The plasmid DNA was collected by centrifugation (13000rpm, 10min), the pellet rinsed with 70% (v/v) alcohol, and dried in a 50°C oven. The dried pellet was finally resuspended in 100ul of TE buffer and stored at -20°C until needed.

2.7 Large-scale isolation of plasmid DNA from *E. coli*.

Large-scale isolation of plasmid DNA was performed using a modification of the method described by Birnboim and Doly (1979). 500ml of an overnight *E. coli* culture was spun down in a Sorvall centrifuge (8000rpm, 10min). The supernatant was discarded and the pellet resuspended in 20ml Birnboim buffer. 50ml 0.2M NaOH, 1% (w/v) SDS was added and the mixture thoroughly shaken. 40ml 4M Na acetate pH6 was added, the mixture shaken, and placed on ice for 20min. The mixture was then spun (8000rpm, 30min), and the supernatant mixed with 50ml 50% (w/v) PEG 6000 and left on ice for 2hrs. The preparation was spun (5000rpm,
20min), the pellet resuspended in 5ml TE buffer, extracted once with phenol/chloroform, and nucleic acids in the aqueous phase precipitated using two volumes of absolute alcohol and 1/10 volume 4M Na acetate pH 6. The precipitate was collected by centrifugation (9000rpm, 10min). The pellet was resuspended in 5ml TE buffer, mixed with 5.5g caesium chloride (CsCl) and 200ul 10mg/ml ethidium bromide (EtBr), and placed in a heat-sealable ultracentrifuge tube. The sample was spun in a Beckman L7 ultracentrifuge at 50000rpm for 20hrs and relaxed at 40000rpm for 30min before removing the gradients.

The plasmid DNA band was recovered using a syringe and extracted 3X with a CsCl-saturated solution of isopropanol (10g CsCl, 10ml distilled water, 100ml isopropanol); each time an equal volume of CsCl/isopropanol solution was mixed with the DNA, left on the bench for 1min and the top layer discarded. After removal of the EtBr, the remaining volume was made up to 5ml with TE buffer and the DNA precipitated on ice for 20min by adding two volumes of ethanol and 1/10 volume of 4M Na acetate pH 6. The precipitate was centrifuged (9000rpm, 10min), the pellet rinsed with 70% (v/v) alcohol and respun (9000rpm, 10min). The pellet was dried, resuspended in 1ml TE buffer, and stored at -20°C until required. Quantitation of DNA was done by measuring absorbance at OD260; an OD260 of 1 is equivalent to 50ug/ml for double-stranded DNA.

2.8 Preparation and transformation of Streptomyces protoplasts.

Preparation and transformation of S. lividans TK24 protoplasts was done as described by Bibb et al. (1978) and Thompson et al. (1982). S. lividans was grown in 25ml YEME (supplemented with 0.5% w/v glycine) for 36-40hrs at 30°C with shaking. The mycelia were harvested (3000rpm, 10min), washed with 15ml 10.3% (w/v) sucrose, and respun (3000rpm, 10min). The pellet was resuspended in 4ml protoplast (P) buffer (10.3% w/v sucrose, 0.025% w/v K2SO4, 0.2% w/v MgCl2, 0.005% w/v KH2PO4, 0.368% w/v CaCl2, 0.573% w/v TES buffer pH 7.2) with
1mg/ml lysozyme, and incubated at 30°C for 1hr. 5ml P buffer was then added, the mixture filtered through cotton wool, and the protoplasts sedimented gently (1500rpm, 5min). The protoplasts were resuspended in 5ml P buffer, dispensed in 1ml aliquots, and stored at -70°C until required.

For transformation with plasmid DNA, 3ml P buffer was added to an aliquot of thawed protoplasts and spun (1500rpm, 5min). The protoplasts were resuspended in the remaining drop of P buffer. Plasmid DNA (in up to 20ul of TE buffer) was added, after which 0.5ml T buffer (25% w/v PEG 1000 in P buffer) was added. Next, 5ml P buffer was added and the mixture spun (1500rpm, 5min). The pellet was resuspended in 0.5ml P buffer and 100ul plated onto five R2YE plates. After incubation for 20hrs at 30°C, the plates were overlayed with 3ml soft agar containing 500ug/ml thiostrepton. Resistant colonies were restreaked onto malt extract yeast extract plates containing 50ug/ml thiostrepton.

2.9 Small-scale isolation of *Streptomyces* total DNA.

Total DNA was isolated from *Streptomyces* as described by Fisher (see Hopwood *et al.* 1985b). Approximately 50mg of mycelia were resuspended in 500ul *Streptomyces* lysozyme solution (0.3M sucrose, 50mM glucose, 25mM Tris/HCl pH8, 25mM EDTA, 50ug/ml RNase, 2mg/ml lysozyme), and incubated for 30min at 37°C. 250ul 2% (w/v) SDS was added, the mixture vortexed, and extracted several times with phenol/chloroform. The DNA was precipitated by adding 1/10 volume 3M Na acetate pH4.8 and equal volumes of isopropanol, and incubating at -20°C for 20min. The DNA was spun (13000rpm, 10min) and the pellet resuspended in 100ul TE buffer. Typically 0.5 to 5ug of DNA was obtained per sample.

2.10 Small-scale isolation of plasmid DNA from *Streptomyces*.

This procedure is based on the alkaline lysis method developed by Kieser (1984). *S. lividans* grown overnight at 30°C in 20ml MEP was harvested (3000rpm,
10min), the pellet resuspended in 5ml *Streptomyces* lysozyme solution, and incubated at 37°C for 1hr. 3ml 0.3M NaOH, 2% (w/v) SDS was added and the mixture extracted once with 3ml acid phenol/chloroform (50% w/v phenol, 10% v/v distilled water, 50% v/v chloroform). The aqueous phase was precipitated with 1/10 volume 3M K acetate pH4.8 and equal volumes isopropanol for 5min at room temperature, and spun (3000rpm, 10min). The pellet was dissolved in 2ml TE buffer, after which 4ml 0.2M NaOH, 1% (w/v) SDS was added and the mixture left at -20°C for 5min. 3ml 3M K acetate pH4.8 was then added and the mixture left at -20°C for a further 20min. After spinning (3000rpm, 10min), the supernatant was extracted with 3ml acid phenol/chloroform. The aqueous phase was precipitated with equal volumes of isopropanol for 5min at room temperature, and spun (3000rpm, 10min). The DNA pellet was resuspended in 500ul TE buffer and extracted once with 250ul neutral phenol/chloroform. The aqueous phase was precipitated with 1/10 volume 3M K acetate pH4.8 and equal volumes of isopropanol at -20°C for 20min, and spun (13000rpm, 10min). The pellet was washed with 70% (v/v) alcohol and dried in a 50°C oven. Finally, the dried pellet was resuspended in 100ul TE buffer. Typical yields of plasmid DNA obtained were between 2 to 10ug per sample.

2.11 Large-scale isolation of plasmid DNA from *Streptomyces*.

Large-scale isolation of plasmid DNA from *Streptomyces* was done essentially as described in Section 2.7, except that mycelia from a 500ml 2 day YEME culture were incubated with Birnboim buffer plus 2mg/ml lysozyme at 37°C for 1hr before proceeding with the next steps.

2.12 Restriction enzyme digestions, ligations, and agarose gel electrophoresis.

All restriction enzyme digestions and ligations were performed as described by Maniatis *et al.* (1982). 0.2 to 1ug of plasmid DNA (in 18ul of TE buffer) was
usually mixed with 1ul (5 units) enzyme and 2ul 10x restriction buffer (500mM Tris/HCl pH7.5, 50mM MgCl₂, 0 to 150mM NaCl depending on the restriction enzyme used). If the DNA had to be dephosphorylated, 5 units of calf-intestinal alkaline phosphatase was also added to the reaction. After 2-3hrs incubation at 37°C, 5ul stop mix (0.1M EDTA, 40% w/v sucrose, 0.15mg/ml bromophenol blue) was added and the sample loaded onto a 1% (w/v) agarose gel. Gels were made in 1x Tris-borate electrophoresis buffer (90mM Tris/base, 90mM boric acid, 10mM EDTA, 0.5ug/ml EtBr) and electrophoresed in the same buffer at 150V for 2-3hrs. The DNA bands were visualized on an ultra-violet transilluminator and photographed.

Ligation reactions typically contained 0.5-5ug DNA in 20-50ul TE buffer, 2ul (20 units) to 5ul (50 units) T4 ligase, and 1/10 volume 10x ligase cocktail (660mM Tris/HCl pH7.5, 100mM MgCl₂, 100mM DTT, 1mM ATP). Ligation reactions were incubated overnight at 4°C before being used to transform E. coli (Section 2.5) or Streptomyces (Section 2.8).

2.13 Recovery of DNA from agarose gels.

DNA was recovered from agarose gels by electrophoresis onto a dialysis membrane as described by Girvitz et al. (1980). Two incisions were made, one below and the other above the DNA band of interest. A piece of Whatman 3MM paper backed by pre-boiled dialysis membrane was inserted into each cut and the agarose gel electrophoresed at 300V for 10min. The bottom paper and dialysis membrane were removed from the gel and washed 3x with 100ul elution buffer (0.2M NaCl, 50mM Tris/HCl pH7.6, 1mM EDTA, 0.1% w/v SDS). The pooled washes were extracted once with phenol/chloroform, the aqueous phase precipitated with two volumes of ethanol and 1/10 volume 4M Na acetate pH6 on ice for 20min, and spun (13000rpm, 10min). The DNA pellet was washed with 70% (v/v) alcohol, dried, and resuspended in a small volume of TE buffer. Quantitation of DNA was done as described in Section 2.7.
2.14 Radioactive-labelling of DNA.

The Amersham nick-translation kit was used to obtain [α-32P]dCTP labelled DNA probes with a high specific activity (>1x10^8 cpm/ug DNA) for use in Southern and Northern blots. Essentially, 1-2ug plasmid DNA was added to 20ul nucleotide buffer solution (100um dATP, 100um dGTP, 100um dTTP in Tris/HCl pH7.8), 10ul [α-32P]dCTP (specific activity 3000Ci/mmole), 10ul enzyme solution (5 units DNA polymerase I, 100pg DNase I), and the volume brought up to 100ul with water. The reaction was incubated at 15°C for 2hrs and then terminated by adding 2ul 0.5M EDTA. Labelled DNA was separated from unincorporated [α-32P]dCTP by centrifugation through a G-50 sephadex column prepared in a 1ml syringe. An aliquot (1-2ul) of purified labelled probe was measured in a scintillation counter to assess the level of radioactivity.

Probes used in SI mapping and DNaseI experiments were labelled specifically at the 5' protruding end with [γ-32P]ATP using T4 polynucleotide kinase (PNK). The reaction mixture contained 1ug dephosphorylated DNA, 10ul 10x kinase buffer (0.5M Tris/HCl pH7.6, 0.1M MgCl2, 50mM DTT, 1mM spermidine, 1mM EDTA), 150uCi [γ-32P]ATP (specific activity 3000 Ci/mmole), 10-20 units T4 PNK, and the volume made up to 50ul with distilled water. This mixture was incubated at 37°C for 30min, after which the reaction was stopped by heating at 65°C for 10min. A secondary digest was then carried out and the DNA run on a 1% (w/v) agarose gel. The required end-labelled fragment was isolated by slicing it out from the gel and spinning it through glass wool (6000rpm, 10min). The eluent was collected and the DNA precipitated with two volumes of absolute alcohol and 1/10 volume 4M Na acetate pH6 for 20min on ice. The DNA was spun (13000rpm, 10min) and the pellet was resuspended in 50-100ul TE buffer.

Probes used in DNase I footprinting experiments were also labelled by filling in a 5' overhang end with [α-32P]dNTP using T7 DNA polymerase. Essentially, 1ug DNA was mixed with 10 units T7 DNA polymerase, 10ul 10x reaction buffer
(100mM Tris/HCl pH7.5, 100mM MgCl₂, 100mM DTT), and 10uCi [α-³²P]dNTP. This mixture was incubated at 37°C for 1hr, after which the reaction was stopped by heating at 65°C for 10min. The remaining steps of the procedure have been described above.

2.15 Southern blotting.

Southern blotting of DNA was done essentially as described by Southern (1975). After electrophoresis of DNA samples was complete, the agarose gel was photographed alongside a ruler so that the distance migrated of any given DNA band could be calculated. The gel was denatured by soaking it in several volumes of 1.5M NaCl, 0.5M NaOH for 1hr at room temperature with constant shaking. The gel was then neutralized by soaking it in several volumes of 1M Tris/HCl pH8.0, 1.5M NaCl for 1hr at room temperature with constant shaking. The gel was inverted onto a Whatman 3MM paper (supported by a glass plate) whose sides were immersed in 10x SSC (for 20x SSC: 175.3g NaCl, 88.2g Na citrate in 1 l water). A sheet of Biorad Zeta-probe nylon membrane, soaked in water, was placed on top of the gel. Two pieces of Whatman 3MM paper, soaked in water, were placed on top of the nylon membrane. A stack of tissues (5-8cm high) was placed on top of the 3MM paper. Finally, a glass plate and 500g weight were placed on top of the tissue. After overnight transfer, the nylon membrane was recovered, dried briefly at room temperature, and baked for 2hrs at 80°C under vacuum. For detecting single-stranded DNA, agarose gels were blotted directly onto Biorad Zeta-probe nylon membranes without any prior pretreatment of the gels.

The baked nylon membrane was placed in a heat-sealable plastic bag containing prehybridisation solution (50% v/v formamide, 4x SSC, 1% w/v SDS, 5x Denhardt's solution, 500ug/ml denatured salmon sperm DNA) to a volume of 150ul/cm² of filter. The bag was then placed in a 55°C oven for 1hr with shaking. The bag was cut open and denatured [α-³²P]dCTP-labelled probe (that is, probe boiled for 5min
and rapidly chilled on ice) was added. The bag was resealed and incubated at 55°C overnight with shaking.

The membrane was removed from the bag, rinsed briefly with 2x SSC, and then washed successively in each of the following solutions for 15min at room temperature: 2x SSC, 0.1% (w/v) SDS; 0.5x SSC, 0.1% (w/v) SDS; and finally 0.1x SSC, 0.1% (w/v) SDS. The filter was then wrapped in Saran wrap and exposed to Fuji RX X-ray film.

2.16 Isolation of total RNA from *E. coli*\textit{Streptomyces}.

This procedure is based on that of Kirby \textit{et al.} (1967) and Glisin \textit{et al.} (1974). A 100ml culture of *E. coli* grown in nutrient broth or *Streptomyces* grown in YEME, was harvested by centrifugation (8000rpm, 10min) or vacuum filtration on Whatman #1 discs respectively. The *E. coli* cells or *Streptomyces* mycelia were resuspended in 5ml modified Kirby mixture (1% w/v sodium triisopropyl naphthalene sulphonate, 6% w/v sodium 4-amino salicylate, 6% v/v phenol/chloroform, 50mM Tris/HCl pH8.3) and 14g 4.5-5.5mm glass balls in the case of *Streptomyces*. The mixture was vortexed vigorously for 2min. 5ml phenol/chloroform was then added and the mixture re-vortexed. The homogenate was transferred into a fresh tube and spun (3000rpm, 10min). The aqueous phase was layered onto a 5.7M CsCl, 0.1M EDTA cushion in an ultracentrifuge tube and spun (35000rpm, 12hrs). The supernatant was carefully discarded, the RNA pellet resuspended in sterile water, and stored at -70°C in 5ul aliquots. Quantitation of RNA was done by measuring absorbance at OD260; an OD260 of 1 is equivalent to 40ug/ml for RNA.

2.17 Northern blotting.

Northern blotting of RNA samples was done essentially as described by Fourney \textit{et al.} (1988). 5ul (10-20ug) RNA was added to 25ul RNA sample buffer [75% v/v formamide, 24% v/v 37% formaldehyde, 10% v/v glycerol, 0.8% w/v bromophenol
blue, 15% 10x MOPS buffer (0.2M MOPS, 50mM Na acetate, 10mM EDTA pH7.0)], incubated at 65°C for 15min and placed on ice. 1ul 10mg/ml EtBr was then added and the sample loaded onto a 1% (w/v) agarose gel (made in 5.1ml 37% formaldehyde, 10ml 10x MOPS buffer, 87ml water). The gel was electrophoresed in 1x MOPS buffer at 14V/cm.

When electrophoresis was completed, the gel was soaked in several volumes of 10x SSC for 30min and then blotted onto Biorad Zeta-probe nylon membrane in the presence of 10x SSC (as described in Section 2.15). The membrane was baked at 80°C for 2hrs and incubated with prehybridisation solution (5x SSC, 7% w/v SDS, 10x Denhardt's solution, 20mM phosphate buffer pH7.0) to a volume of 150ul/cm² at 55°C for 1hr. The bag was cut open, the prehybridisation solution replaced with hybridisation solution (prehybridisation solution plus 10% (w/v) dextran sulphate and denatured [α-32P]dCTP labelled probe), and incubated at 55°C overnight with shaking. The membrane was then washed and exposed to Fuji RX X-ray film as described in Section 2.15.

2.18 High resolution S1 mapping.

This was done essentially as described by Berk and Sharp (1977). An RNA sample and the test DNA were co-precipitated with isopropanol at -20°C (usually overnight). The test DNA was labelled specifically at the 5' position using T4 PNK. The pellet was washed with 70% (v/v) ethanol and dried. The pellet was resuspended in 20ul hybridisation solution (40mM PIPES pH6.4, 400mM NaCl, 1mM EDTA, 80% v/v formamide), and placed in an 85°C water bath for 10min, after which the temperature was allowed to equilibrate to 62-63°C over 30-60min and the sample incubated at this temperature for a further 3-4hr. The sample was then mixed with 300ul chilled S1 digestion solution (1000 units S1 nuclease, 0.28M NaCl, 30mM Na acetate pH4.4, 4.5mM zinc acetate, 20ug/ml partially cleaved denatured non-homologous DNA), and incubated at 37°C for 45min. The reaction was stopped by adding 75ul S1 termination solution (2.5M ammonium acetate,
0.05M EDTA). 10ul carrier tRNA was then added and the mixture precipitated with 400ul isopropanol at room temperature for 5min. It was spun for 5min and the pellet was washed with 70% (v/v) ethanol and dried. The pellet was resuspended in 5ul denaturing dye solution (80% v/v formamide, 10mM NaOH, 1mM EDTA, 0.1% w/v xylene cyanol, 0.1% w/v bromophenol blue), boiled for 3min at 90°C and quick-chilled on ice. The S1 protected fragments were electrophoresed on an 8% polyacrylamide-urea sequencing gel (Section 2.19). The gel was then dried under vacuum and exposed to Fuji RX X-ray film.

2.19 Double-stranded DNA sequencing.

DNA sequencing was done using a Pharmacia sequencing kit which employs the dideoxy chain termination method developed by Sanger et al. (1977). Essentially, 5ul (0.2ug/ul) double-stranded DNA template was mixed with 3ul distilled water and 2ul 2M NaOH, and incubated at room temperature for 10min. 7ul distilled water, 3ul 3M Na acetate pH4.5 and 60ul absolute alcohol were then added, the mixture incubated at -20°C for 20min, and the precipitated DNA collected by centrifugation (13000rpm, 10min). The pellet was washed with 70% (w/v) ethanol, dried, and resuspended in 6.5ul distilled water. 1.5ul sequencing buffer (100mM Tris/HCl pH7.5, 100mM MgCl₂, 100mM DTT) and 2ul primer were added, and the mixture incubated at 37°C for 20min, and then at room temperature for at least 10min. 3ul Klenow fragment and 2ul (20uCi) [α-35S]dATP were then added to produce an enzyme/label/primer/template (ELPT) mixture.

Four Sarstedt tubes were labelled A,T,G,C, and 3ul A mix (ddATP in solution with dATP, dTTP, dGTP, dCTP) was added to A, 3ul T mix (ddTTP in solution with dATP, dTTP, dGTP, dCTP) added to T etc. 3ul ELPT mixture was placed on the wall of each tube, spun for 2sec to mix, and incubated at room temperature for 15min. 1ul chase solution (dATP, dTTP, dGTP, dCTP in solution) was placed on the wall of each tube, spun for 2sec, and incubated at room temperature for a further 15min. Finally, 3ul sequencing stop mix (10mM EDTA, 0.3% w/v xylene cyanol,
0.3% w/v bromophenol blue in formamide) was added to each tube. Samples were boiled at 100°C for 3min and placed on ice. 3ul of each sample was loaded onto an 8% denaturing polyacrylamide gel (63g urea, 11.4g polyacrylamide, 0.6g bis-acrylamide, 15ml 10x Tris borate buffer, made up to 150ml with water). The gel was electrophoresed in 1x Tris borate buffer at 1200v for 2-3hrs, after which it was fixed by soaking in several volumes of 5% (v/v) acetic acid, 5% (v/v) methanol for 30min. The gel was then transferred to Whatman 3MM paper, dried in a slab dryer at 80°C for 2hrs, and finally exposed to Fuji RX X-ray film.

### 2.20 SDS-polyacrylamide gel electrophoresis (PAGE).

Protein samples were analyzed on 10% SDS-polyacrylamide gels as described by Laemmli (1970). Resolving gel was made by mixing 10ml 30% stock acrylamide solution (30g acrylamide, 0.8g bis-acrylamide made up to 100ml with water), 7.5ml pH8.8 buffer (18.17g Tris/base pH8.8, 4ml 10% w/v SDS in 100ml water), and 11.0ml water. (18% SDS-polyacrylamide gels were made by adding 18ml 30% stock acrylamide to the resolving gel). 1.5ml 15mg/ml ammonium persulphate and 22.7ul TEMED were then added, the mixture poured between two glass plates, and allowed to polymerise. Stacking gel was made by mixing 1.3ml 30% stock acrylamide solution, 3.3ml pH6.8 buffer (6g Tris/base pH6.8, 4ml 10% w/v SDS in 100ml water), and 8.0ml water. 0.67ml 15mg/ml ammonium persulphate and 13.3ul TEMED were then added, the mixture poured on top of the resolving gel, and allowed to polymerise.

Protein samples were denatured by mixing with an equal volume of protein sample buffer (5% v/v β-mercaptoethanol, 3.4% w/v SDS, 15% v/v glycerol, 0.01% w/v bromophenol blue, 47mM Tris/HCl pH6.8) and boiled for 3min. Samples were loaded onto SDS-polyacrylamide gels and electrophoresed at 35mA for 3-4hrs or overnight at 7mA in pH8.3 electrophoresis buffer (3.1g Tris/base, 1g w/v SDS, 14.2g glycine, in 1l water). Protein bands were visualized by staining the gel in Coomassie brilliant blue solution (0.126g Coomassie brilliant blue R, 4.6% v/v
acetic acid, 25% v/v methanol) for 1-2hrs and destaining with 7.5% (v/v) acetic acid, 5% (v/v) methanol.

Urea SDS-PAGE was done according to Swank and Munkres (1971). The resolving gel was made by mixing 15ml 25% stock acrylamide solution (25g acrylamide, 2.5g bis-acrylamide in 100ml water), 3.0ml pH5.5 buffer (1% w/v SDS, 1.0M H₃PO₄ pH5.0), 14.4g urea, and made up to 30ml with water. 0.3ml 6% w/v ammonium persulphate and 20ul TEMED were added, the mixture poured between two glass plates and allowed to polymerise. Stacking gel was made by combining 1.3ml 30% stock acrylamide, 3.3ml pH5.5 buffer, 8ml water, 0.67ml 15mg/ml ammonium persulphate, and 13.3ul TEMED. Protein samples were resuspended in an equal volume of urea protein sample buffer (1% w/v SDS, 8M urea, 1% β-mercaptoethanol, 0.01% w/v bromophenol blue, 0.01M H₃PO₄ pH6.8), boiled for 3min, and electrophoresed at 35mA overnight in pH6.8 electrophoresis buffer (0.1% w/v SDS, 0.1M H₃PO₄ pH6.8). Staining and destaining were done as described above. Molecular weight markers used for SDS-PAGE were obtained from Sigma (MW-SDS-70K or MW-SDS-175).

2.21 In vitro coupled transcription/translation.

In vitro coupled transcription/translation of DNA was done using a prokaryotic DNA-directed translation kit obtained from Amersham based on the method developed by Zubay (1973). Essentially, 5ul (0.5-5.0ug) of plasmid DNA was incubated with 7.5ul supplement solution (containing nucleotides, tRNA, inorganic salts, energy regenerating system), 3ul amino acid mix minus methionine, 2ul L-[³⁵S] methionine, 7.5ul dilution buffer, and 5ul E. coli S30 extract, for 1hr at 37°C. 5ul methionine chase solution was then added and the samples incubated for a further 5min at 37°C. Equal volumes of protein sample buffer was then added and the samples boiled for 3min. Radioactive protein samples were analyzed by electrophoresing on SDS-polyacrylamide gels (Section 2.20) which were then fixed by soaking in destain for 1/2hr, dried, and exposed to Fuji RX X-ray film.
2.22 Protein extraction from *E. coli*.

50ml nutrient broth containing 2mM IPTG and 50ug/ml ampicillin was inoculated with *E. coli* carrying the required plasmid and grown overnight at 37°C with shaking. The cells were spun down in a Sorvall centrifuge (8000rpm, 10min), the pellet resuspended in 1ml extraction buffer (10mM HEPES pH7.9, 50mM NaCl, 1.5mM MgCl$_2$, 0.5mM PMSF), and sonicated for 10x15sec with 15sec intervals at amplitude 14. The sonicated samples were spun (13000rpm, 5min), and the supernatant retained. Protein samples were stored at -20°C unless otherwise indicated. Protein concentration was measured using the Bio-Rad protein assay. *using the Scuvrep 150 MSE (with exponential microprobe 3mm)*

2.23 Gel retardation assay.

Gel retardation assays using whole cell extracts were carried out essentially as described by Fried and Crothers (1981). 10ul (40ug) protein extract was preincubated with 1ul 5mg/ml poly(dI-dC).poly(dI-dC), 4ul 20% w/v Ficoll, 10ul extraction buffer, 10 to 14ul distilled water (depending on the amount of probe to be added), and incubated for 15min at room temperature. 10-20000cpm end-labelled DNA (usually in 1 to 5ul H$_2$O) was added and the mixture incubated for a further 15min at room temperature. 1ul of sample buffer (30% v/v glycerol, 0.05% w/v xylene cyanol, 0.05% w/v bromophenol blue in 0.125M Tris/HCl) was then added and samples electrophoresed on a 5% non-denaturing polyacrylamide gel (8.33ml 30% stock acrylamide, 2.5ml 10x Tris borate buffer, 0.4ml 10% w/v ammonium persulphate, 60ul TEMED, made up to 50ml water) in 1x Tris borate buffer at 150v for 2hrs. Gels were dried in a slab-dryer and exposed to Fuji RX X-ray film.

2.24 Competition assay.

For the competition assay, final protein concentration in the reaction was reduced to 2ul (8ug). To this was added 1ul 5mg/ml poly(dI-dC).poly(dI-dC), 4ul 20% (w/v) Ficoll, 18ul extraction buffer and the appropriate volume of H$_2$O ranging
from 14ul in the control (in which no cold DNA was to be added) to 0ul in the last reaction (in which 14ul of cold DNA was to be added). The samples were then incubated for 10min at room temperature. Increasing volumes of cold DNA was then added (up to a maximum of 14ul in the last tube) and the samples incubated for another 10min at room temperature. Finally, 10000cpm end-labelled DNA (in 1ul H$_2$O) was then added and the mixture incubated for a further 10min at room temperature. 1ul of sample buffer was added and samples electrophoresed on a 5% non-denaturing polyacrylamide gel at 150V for 2hrs. Gels were dried in a slab-dryer and exposed to Fuji RX X-ray film.

For competition reactions using 5ul (20ug) final protein concentration, the amount of extraction buffer added to the reaction was adjusted to 15ul. All other components were added as described above.

2.25 DNaseI protection assay (DNA footprinting).

This was done essentially as described by Galas and Schmitz (1978). 20ul of protein extract was preincubated with 2ul 5mg/ml poly(dI-dC).poly(dI-dC), 4ul 20% (w/v) Ficoll, 9-13ul distilled water, and incubated for 15min at room temperature. The control reaction contained all the above components except protein, the latter replaced with 20ul of extraction buffer. 20000cpm end-labelled DNA (in 1 to 5ul H$_2$O) was then added and the mixture incubated for a further 15min at room temperature. DNaseI (2mg/ml in 50mM NaCl, 5mM MgCl$_2$, 5mM CaCl$_2$, 20% v/v glycerol, 20mM HEPES pH7.9) was then added to a final concentration of either 5ug/ml (A), 0.5ug/ml (B), 0.05ug/ml (C), 0.005ug/ml (D), or 0.0005ug/ml (E). The reactions were incubated for 1min at room temperature before being terminated by adding 100ul stop buffer (1% w/v SDS, 60ug/ml tRNA, 400ug/ml Proteinase K, 100mM Tris/HCl 10mM EDTA pH8.0). The reactions were then incubated at 37°C for 5min and extracted several times with phenol/chloroform (until there was no denatured protein visible at the interface). The DNA was then precipitated by adding 2 volumes absolute alcohol and 1/10 volume 4M Na acetate pH6, and left on
ice for 20min. The precipitate was collected by centrifugation (13000rpm, 10min),
the pellet washed with 70% (v/v) alcohol, and then dried. The pellet was
resuspended in 4ul sequencing stop mix, boiled for 3min, and electrophoresed on an
8% denaturing polyacrylamide-urea gel (Section 2.19). The gel was dried in a slab-
dryer and exposed to Fuji RX X-ray film.

Abbreviations.
IPTG: Isopropyl β-D-thiogalactoside
X-gal: 5-Bromo-4-chloro-3-indoyl-β-D-galactopyranoside
EDTA: Diaminooethanetetra-acetic acid disodium salt
PEG: Polyethylene glycol
TES: N-tris [Hydroxymethyl]methyl-2-aminoethanesulfonic acid
MOPS: 3-[N-Morpholino]propanesulfonic acid
PIPS: 1,4-Piperazinediethanesulfonic acid
DTT: Dithiothreitol
HEPES: N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid
PMSF: Phenylmethylsulfonyl fluoride
TEMED: N,N,N',N'-Tetramethylethlenediamine
CHAPTER THREE.
Cloning, expression, and characterisation of the cop/korB gene from the Streptomyces plasmid pIJ101 in E. coli JM107.

The aims of this set of experiments were three-fold:

a) firstly, to reconfirm the sequence of the cop/korB gene,

b) secondly, to clone the cop/korB open reading frame in an appropriate expression vector in order to over-express the Cop/KorB protein in vivo in sufficient amounts for purification,

c) and finally, to characterize the biological function of the cop/korB gene product by binding studies (i.e. gel retardation assays and footprinting analysis) using specific DNA fragments from pIJ101.

3.1 Construction of pQR200 (cop/korB plus sti in pUC8).

The next two sections describe the strategy employed to clone the cop/korB gene from the Streptomyces plasmid pIJ101 into the E. coli pUC8 plasmid (Table 2.3). pUC8 was chosen as a cloning vector because its lacZ system should provide a strong inducible promoter from which the cop/korB gene can be expressed in E. coli.

5ug of pIJ101 DNA (Fig. 1.2) isolated from S. lividans TK24 was digested to completion with the restriction endonuclease BclI. This digestion resulted in the production of five fragments on a 1% agarose gel of the following sizes: 3.0kb [BclI(10) to BclI(29)], 1.4kb [BclI(29) to BclI(41)], 1.8kb [(BclI(41) to BclI(49)], 1.2kb [(BclI(49) to BclI(57)], and 1.5kb [(BclI(58) to BclI(10)]. The 1.2kb BclIE fragment [that is, from BclI(49) to BclI(57)] containing the cop/korB gene, sti determinant and orf85/79 genes, was extracted from a 1% agarose gel by electrophoresis onto a dialysis membrane (Section 2.13) and ligated to 2ug of the E. coli plasmid pUC8 cut with BamHI. A sample of the ligation mixture was analysed
on an agarose gel to ensure that ligation had occurred. The ligated DNA was then used to transform competent *E. coli* JM107 cells (Section 2.5) and the transformation mixture plated out onto ampicillin, IPTG, X-gal nutrient agar plates and left overnight at 37°C.

Twelve recombinant (white) colonies were selected and plasmid mini-preparations made in order to examine the constructs obtained (Section 2.6). Plasmid DNA was digested with *EcoRI* and *HindIII* in order to ensure that the 1.2kb *BclIIE* insert was present. Since *Streptomyces* promoters are not normally recognized in *E. coli* (Bibb and Cohen 1982; Bibb et al. 1985a), it is essential to have the reading frame of the putative *cop/korB* gene in the correct orientation with respect to the *lacZ* promoter. This would enable transcription from the *lacZ* promoter which should allow expression of the *cop/korB* gene product in *E. coli*. A secondary *EcoRI/Apal* digest was done to check the orientation of the *BclIIE* fragment with respect to the pUC8 *lacZ* promoter. Thus, once a construct produced a 1.2kb insert after an *EcoRI/HindIII* digest and resulted in 2.9kb and 1.0kb size fragments after an *EcoRI/ApaI* digest, a large-scale plasmid preparation was made (Section 2.7). This construct was named pQR200 (Fig. 3.1).

### 3.2 Construction of pQR206 (*cop/korB* in pUC8).

In order to remove the intervening *sti* region and *orf85/79* genes from pQR200, 5ug of pQR200 (Fig. 3.1) was digested with *EcoRI* and *SpeI* to completion which resulted in two bands: a 3.2kb fragment (pUC8 plus *cop/korB*) and a 0.7kb fragment (*sti*). The digest was run on a 1% agarose gel, the 3.2kb band recovered by electroelution (Section 2.13), and the *EcoRI* and *SpeI* 5' overhanging ends made flush with DNA Polymerase I (Poll). The DNA was then religated overnight and used to transform *E. coli* JM107 (Section 2.5).

As expected, all colonies obtained from this transformation were white. Twelve colonies were selected and their DNA digested with *EcoRI* and *SpeI*. Those constructs which did not digest with *EcoRI* and *SpeI* were chosen for further
1) Cut with SpeI and EcoRI to remove 0.7kb sti fragment.
2) Blunt-end 3.2kb fragment containing korB sequence.
3) Religate.
analysis since presumably these constructs had had the *sti* region successfully removed. These constructs were further digested with *SacI* and *HindIII* respectively, and compared to pQR200 digested with these same enzymes by running them on a 1% agarose gel. Only plasmid DNA which did not cut with *SacI* (indicating that the *sti* region had been removed as the only *SacI* site present on the construct is located within the *sti* region) and which was smaller than pQR200 by approximately 0.7kb after digesting with *HindIII*, was selected for large-scale plasmid preparation. This construct was named pQR206 (Fig. 3.1).

### 3.3 Sequencing of cop/korB (pQR206).

pQR206 (containing the cop/korB gene) was sequenced for two main reasons. Firstly, to analyse the *EcoRI*/SpeI junction in order to determine if the *EcoRI* and SpeI 5' overhanging ends had been removed or filled in by the PolII, or a combination of the two. This in turn would indicate if the cop/korB gene was in the correct reading frame with respect to the *lacZ* gene. Secondly, to reconfirm Kendall and Cohen's (1988) sequence data to ensure that there were no discrepancies that would result in a change in the length of the postulated ORF (if a base had been added or deleted) or a change in the projected amino acid sequence (if a base had been exchanged).

The entire 0.5kb *SpeI*(53)-*BclI*(49) fragment was sequenced by double-stranded sequencing from both ends using reverse and universal primers (Section 2.19). The *cop/korB* sequence data which resulted was completely homologous to that of Kendall and Cohen's (1988) sequence data and no discrepancies were found (Fig. 3.2). Frame analysis (Bibb *et al.* 1984) showed the presence of one ORF within this region. Kendall and Cohen (1987) have previously mapped a gene to this region and named it *korB* due to its ability to suppress the lethal effects of the *kilB* gene. Deng *et al.* (1988) have also mapped a gene to this region and named it *cop* because of its negative effects on copy number. As *cop* and *korB* both map at the same position and only one major ORF was found within this region, this ORF must
Figure 3.2  Sequence of korB gene in pQR206

+1lacZ
CCCAGGCTACGTTTTATGCTTCCGGCTCGGATGTGGAATTGTGAGCCGATAAACAAATTCACA

+1lacZ
S/DlacZ
CAGCAACAGCTATGACCATTGATACGAGATCTAGTGGCAAGACGCGCCGGAGACCCTGATCGCCGGAGGCATCAAC

-EcoR/SpeI
-35korB
CTAGTGGGCAAAGCTGACACGTTGGGCAAGAGGACGACCATCAGGACCGGTCG

+1korBts
XhoI
ATGACGCAAAAAGAACCAGGGGGGAGATCCGGGGGAGGGCCGCTGAAGCCGGGTCG

+1korBts
FspI
MetThrMetIleThrAsnLeuValAlaGlnThrAspThrValGlyGlnAspAspPhe

CGGATCAAGCTCCTCGCGAGCTGGACAGATTGAGAGAGCTGCGCCCCTGATTCGGGAGGCAGTC

SmaI
ArgIleLysLeuLeuValAlaGluLeuAspGluIleGluLysGluLeuArgProLeuIleAlaGluAlaVal

CGGATGGAGGTCCCTACCCGCGCATCAACGAGGTAGCGCGAGTTGCGCTCCGAAACGAGGGGCGTTGG

ArgMetGluValProTyrArgIleAsnGluValThrAlaValAlaProAsnThrAlaArgAlaTrp

GCCAAGCCGAGGAGCCGAAAGCTGCGGCTCTGGACAGGCCTGACCTCGCGCGCAGCGTA

AlaLysAlaGluAlaGluLysGlySerGlySerXXX

Apal
GGGGCCCTCTCTGGAAACCGCTCGAAACCTTGCAACTTCGAGTCAAGGACCGGTCTCGCGGAGGTCAACC

GGGGTGCAACGGCTCAACCGGGTGCGCAACCCCGGAAACCGTCGTCTGGACGTGAAGAAGCCTCGT

CGGGCGGCTGCTGACCGGGGCTTCTCAATGGCGAGTATCG
encode a protein which has a dual function; firstly as a negative regulator of copy number (possibly through its interaction with sti ) during plasmid replication, and secondly as a repressor of kilB during plasmid spread.

Fig. 3.2 also shows the DNA sequence of pQR206 across the EcoRI/Spel junction. When originally constructing pQR206, it was expected that the EcoRI and SpeI 5' overhanging ends would be completely filled in with dNTP's by the PolII. This would have resulted in the LacZ protein being prematurely terminated eleven amino acids after the incorporation of the first methionine because of an artificially generated in-frame stop codon present at this position. If so, then the LacZ protein would have terminated before the start of the cop/korB gene and no LacZ-Cop/KorB in-frame fusion proteins or out-of-frame nonsense proteins would be generated.

However, the sequence data indicates that, although the SpeI 5' overhanging end had been completely filled in with dNTP's by the PolII, the last "T" nucleotide of the EcoRI 5' overhanging end had been removed before the remaining bases had been filled in. This combination of bases should result in the generation of a lacZ-cop/korB in-frame gene fusion with the production of a hybrid polypeptide. Thus, theoretically, E. coli RNA polymerase would recognize the lacZ promoter and a mRNA fusion transcript of approximately 500b should be produced if transcription starts at the lacZ transcription start site (+1lacZtss in Fig. 3.2) and terminates at the end of the cop/korB gene (which has been shown by Deng et al. (1987) to have transcriptional terminating activity in E. coli ). Ribosomes should then bind to this fusion transcript at the lacZ Shine-Dalgarno (AGGA) sequence which lies eight nucleotides upstream of the lacZ translation start site. When translated by the ribosomes, this fusion transcript should give rise to a LacZ-Cop/KorB in-frame fusion protein with an additional 19 amino acids incorporated at the N-terminus of the Cop/KorB protein.
By analysing the DNA sequence of the \textit{cop/korB} ORF, the amino acid sequence of the putative protein can be deduced and is shown in Fig. 3.2. The native Cop/KorB protein would possess 80 amino acids and have a molecular weight of 8.8kDa. The average pI of the native protein would be 6.23 and thus it would be slightly acidic in nature. However, the LacZ-Cop/KorB fusion protein generated from pQR206 would possess 99 amino acids (due to the additional 19 amino acids incorporated at its N-terminal end) and would have an increased molecular weight of 10kDa. The average pI of the LacZ-Cop/KorB fusion protein would be 6.10 and thus it would also be intrinsically acidic.

It is important to realize that if the LacZ-Cop/KorB fusion protein is produced from pQR206, then the additional amino acids incorporated at its N-terminal end may interfere with the correct folding of the Cop/KorB protein which in turn will affect activity. However, the advantage of this construct is that being both a transcriptional and a translational fusion, it has the optimum location of control regions. This should maximise the probability of protein production from the cloned \textit{cop/korB} gene, as the \textit{cop/korB} gene has been placed directly under the control of a strong promoter and ribosome binding site which are efficiently recognised by the \textit{E. coli} host machinery for transcription and translation respectively.

3.4 Expression of Cop/KorB protein \textit{in vitro}.

The bacterial cell-free transcription-translation (CTT) system (Zubay 1973) allows the \textit{in vitro} expression of genes from bacterial plasmids, provided that the relevant control signals are present such as the Pribnow box for the initiation of transcription and the Shine-Dalgarno sequence for translation. The system utilises cell-free extracts from \textit{E. coli} containing the necessary enzymes and components for coupled transcription-translation (for example, ribosomes, RNA polymerases, etc.), although this extract has to be supplemented with amino acids, an energy
regeneration system (ATP, CTP, GTP, UTP, PEP), and several cofactors (tRNA, cyclic AMP).

This system was used to study cop/korB expression in vitro. The CTT system identifies if any protein is produced specifically from the cloned cop/korB gene, thereby mapping the protein product to a defined DNA fragment. This application represents a major advantage over the use of in vivo methods, since matching of cloned DNA fragments with unidentified protein products can be difficult in systems such as minicells or maxicells. Furthermore, the incorporation of radioactive label into protein is relatively more efficient than using in vivo methods.

Fig. 3.3 shows in vitro generated CTT protein products (Section 2.21) from pUC8 and pQR206 run on an 18% SDS-polyacrylamide gel (Section 2.20). When pUC8 was used in the in vitro CTT system, the major protein product produced was the 31kDa β-lactamase protein from the ampicillin resistance (bla) gene. However, when pQR206 (Fig. 3.1) was used to initiate synthesis of proteins in the in vitro E. coli CTT system, an additional protein of approximately 10kDa was produced. As can be seen by the rather diffuse nature of this 10kDa band, polypeptides with a molecular weight below 10kDa are not well resolved by PAGE when SDS is the sole dissociating agent and the accurate sizing of small proteins becomes difficult, even when using 18% gels. Thus, an amended gel system was developed based on the method of Swank and Munkres (1971) whereby 8M urea was incorporated in 12.5% SDS-polyacrylamide gels containing a high ratio of bisacrylamide crosslinker to acrylamide (1:10) (Section 2.20). Under these conditions, electrophoresis gives a good linear relationship between log protein molecular weight and mobility over the molecular weight range of 2.5kDa to 17kDa (Swank and Munkres 1971).

When pQR200 (Fig. 3.1) was used in the CTT system and the 35S-methionine labelled proteins were run on an 8M urea 12.5% SDS-polyacrylamide gel, only the β-lactamase protein (31kDa) was produced (Fig. 3.4, lane 3), and the pattern
Figure 3.3

*In vitro* coupled transcription-translation of pQR206 resolved on an 18% SDS-polyacrylamide gel.

CTT reactions were done as described in Section 2.21. After electrophoresis (Section 2.20) of the $^{35}$S-methionine labelled protein products, the gel was dried down and exposed to Fuji RX X-ray film. The arrow indicates the production of a new protein band of approximately 10kDa present in the pQR206-generated protein products compared to the pUC8-generated protein products (as a control). This new band corresponds to the expected size of the Cop/KorB fusion protein which would be produced from pQR206 (see Fig. 3.2).
Figure 3.4

*In vitro* coupled transcription-translation of pQR206 resolved on an 8M urea 12.5% SDS-polyacrylamide gel.

CTT reactions were done as described in Section 2.21. After electrophoresis (Section 2.20) of the $^{35}$S-methionine labelled protein products, the gel was dried down and exposed to Fuji RX X-ray film.

**Lanes**

1) pUC8: 31kDa β-lactamase protein produced (arrowed).
2) pUC8 cut with *Scal*: β-lactamase protein not produced.
3) pQR200: 31kDa β-lactamase protein produced.
4) pQR206: 10kDa KorB protein produced (arrowed) in addition to β-lactamase.
5) pQR206 cut with *Sma*I: 10kDa KorB protein not produced.
obtained was very similar to that observed for pUC8 alone (lane 1). This presumably was because the cop/korB gene was not transcribed from its own promoter. It therefore suggests that the cop/korB promoter is not recognised by the E. coli RNA polymerase, as is usually the problem with most Streptomyces promoters (Bibb and Cohen 1982; Bibb et al. 1985a). If the lacZ promoter is driving transcription then the intervening sti sequence (which is located between placZ and cop/korB) may act as a transcriptional terminator due to its inverted repeat structure which could form a stem-loop in vivo (Chapter 5, Section 5.6). Thus, E. coli RNA polymerase could terminate transcription at the sti site before any cop/korB transcript can be synthesized. Even if efficient transcription from pcop/korB or placZ was taking place, E. coli ribosomes a) may not recognize or bind only very weakly to the potential E. coli-like AGGA ribosomal binding site located 11bp upstream from the cop/korB AUG start codon or b) may not recognize the AUG translational start codon since the latter is located on the verge of the acceptable limit between a ribosomal binding site and a translational start site (Shine and Dalgarno 1974).

When pQR206 (Fig. 3.1) was used in the in vitro CTT system, an additional protein of approximately 10kDa was produced (Fig. 3.4, lane 4), since in this construct the cop/korB gene was placed immediately adjacent to the lacZ promoter and fused with the lacZ gene by deleting the intervening EcoRI/SpeI sti region (see Fig. 3.2). Digestion of pQR206 with Smal, which by sequence analysis cleaves the cop/korB coding region 45bp (14 amino acids) from the C-terminus end (see Fig. 3.2), eliminated the production of the 10kDa protein when placed in the in vitro CTT system (Fig. 3.4, lane 5). This result verified that the designated cop/korB region does encode for a protein with an approximate size of 10kDa from sequence analysis. As a control, pUC8 was digested with Scal (which cleaves within the bla gene) thereby disrupting the production of the 31kDa β-lactamase protein (Fig. 3.4, lane 2).
3.5 Expression of Cop/KorB protein *in vivo* in *E. coli*.

Although the *in vitro* CTT system confirmed that the pQR206 plasmid construct produces the 10kDa Cop/KorB protein, only very small quantities of this protein were made. It is desirable to over-express the protein *in vivo* so that relatively large quantities can be obtained for further experimentation and eventual purification.

Initial attempts using SDS-PAGE to detect Cop/KorB protein production *in vivo* from 2hr induced *E. coli* cultures (OD$_{550}$=0.2) harbouring pQR206 were not successful. In order to ensure that there was no barrier to *cop/korB* expression at the transcriptional level, a Northern blot was done to confirm that transcription was taking place at a detectable frequency and that any message produced *in vivo* was not unstable.

Total mRNA was isolated from *E. coli* JM107 containing pQR206 and pUC8 (as a control) grown overnight in the presence of IPTG and ampicillin (Section 2.16), run on a 1% agarose gel containing formaldehyde, and blotted directly onto a Zeta-probe nylon membrane (Section 2.17). The nylon membrane was then probed with $[^\alpha-\text{32P}]$dCTP-labelled 0.5kb *cop/korB* fragment isolated from pQR200 (Fig. 3.1) cut with *Spel/HindIII*. A message of approximatly 500b in length was detected in total mRNA isolated from *E. coli* harbouring pQR206 but not for mRNA isolated from *E. coli* harbouring pUC8 (Fig. 3.5). This band corresponds to the size of the message expected from pQR206 if transcription had commenced at the *lacZ* promoter and terminated at the end of the *cop/korB* gene which has been shown by Deng *et al.* (1987) to have efficient transcriptional terminator activity in *E. coli*. It should be noted that the Northern blot simply confirms the presence and size of the *cop/korB* mRNA transcript and that the relative abundance or half-life of the mRNA cannot be inferred from this data.

In order to re-examine the expression of the Cop/KorB protein product *in vivo*, cultures of *E. coli* JM107 containing pQR206 and pUC8 (as a control) respectively
Total RNA was isolated from induced cultures of *E. coli* JM107 harbouring pQR206 and pUC8 (as a control) respectively (Section 2.16). 20μg of RNA from each sample prepared above was resolved on a 1% agarose gel containing formaldehyde and blotted onto a BioRad Zeta-probe nylon membrane (Section 2.17). Baked membranes were probed with [α-32P]dCTP nick-translated 0.5kb *cop/korB* fragment isolated from pQR200 (Fig. 3.1) cut with SpeI and HindIII. The arrow indicates a message of approximately 500b in size present in total RNA isolated from *E. coli* containing pQR206 but not from *E. coli* containing pUC8.
Figure 3.6

*In vivo* synthesis of *cop/korB*-encoded protein product from pQR206 in *E. coli* resolved on an 8M urea 12.5% SDS-polyacrylamide gel.

Induced overnight cultures of *E. coli* JM107 containing pQR206 and pUC8 (as a control) respectively were sonicated in extraction buffer (Section 2.22). Approximately 80ug of each total protein sample prepared above was analysed by electrophoresis on an 8M urea 12.5% SDS-polyacrylamide gel, after which the gel was stained with Coomassie blue and destained (Section 2.20).

Lane 1) Total protein extract from *E. coli* containing pUC8 (control).
Lane 2) Total protein extract from *E. coli* containing pQR206 showing an additional 10kDa band (arrowed) corresponding to the Cop/KorB protein.
In *vivo* synthesis of *cop/korB*-encoded protein product from pQR206 in *E. coli* resolved on an 18% SDS-polyacrylamide gel.

Induced overnight cultures of *E. coli* JM107 containing pQR206 and pUC8 (as a control) respectively were sonicated in extraction buffer (Section 2.22). Approximately 80µg of each total protein sample prepared above was analysed by electrophoresis on an 18% SDS-polyacrylamide gel, after which the gel was stained with Coomassie blue and destained (Section 2.20).

Lane 1) Total protein extract from *E. coli* containing pUC8 (control).
Lane 2) Total protein extract from *E. coli* containing pQR206 showing an additional 10kDa band (arrowed) corresponding to the Cop/KorB protein.
Figure 3.8

Densitometry scan.

Scanning of the negative obtained for the stained gel in Fig. 3.7, lane 2 (showing total protein from *E. coli* JM107 containing pQR206) indicates that the Cop/KorB protein accounts for at least 2% of total intracellular protein. The shaded peak corresponds to the Cop/KorB protein band.
were grown overnight in the presence of IPTG and ampicillin (Section 2.22), and sonicated samples run on 8M urea 12.5% SDS-polyacrylamide (Fig. 3.6) and 18% SDS-polyacrylamide gels (Fig. 3.7) respectively. Under both gel systems, an additional protein band corresponding to 10kDa in size was present in protein extracts made from *E. coli* JM107 harbouring pQR206 but not in extracts made from *E. coli* harbouring pUC8 alone. Thus, it was necessary to grow the *E. coli* culture harbouring pQR206 until *stationary phase* was reached (typically after an overnight incubation) so that sufficient Cop/KorB protein was produced *in vivo* for detection by SDS-PAGE. Densitometry scanning of the stained gel in Fig. 3.7 indicated that production of the Cop/KorB protein accounted for at least 2% of total intracellular protein (Fig. 3.8).

### 3.6 Cloning of *cop/korB* into other expression vectors.

In order to obtain higher yields of Cop/KorB protein for purification purposes, the *cop/korB* gene was subcloned into the *E. coli* plasmid pKK223.3 (Table 2.3) which contains the strong *trp-lac (tac)* promoter for controlled, high-level expression.

pQR200 (Fig. 3.1) was digested with *PstI* and *SpeI*, and the 0.5kb *cop/korB* fragment extracted from a gel (Section 2.13). This 0.5kb *cop/korB* fragment was treated with *P*oll to make the *PstI* and *SpeI* overhanging ends flush before ligating it in both orientations to pUC8 cut with *SmaI*, creating pQR412 and pQR413 respectively (Fig. 3.9). pQR412 was then digested with *EcoRI* and *PstI*, and ligated to pKK223.3 cleaved with the same enzymes, creating pQR414 (Fig. 3.9). pUC8 was used as an intermediate construct for the subcloning of *cop/korB* because pKK223.3 does not have a convenient blue/white screening system, and thus every ampicillin resistant pKK223.3 colony would need to be analysed to determine if it contained an insert. However, since pKK223.3 does not contain the *lacZ* gene, no fusion proteins can result from the construction of pQR414. As commercial primers cannot bind to pKK223.3, pQR412 rather than pQR414 was used for sequencing to
Construction of pQR412, pQR413 and pQR414

1) Cut pQR200 with PstI and SpeI
2) Blunt-end 0.5kb korB fragment and ligate to pUC8 cut with SmaI

1) Cut pQR412 with EcoRI and PstI
2) Ligate 0.5kb korB fragment into pKK223.3 cut with EcoRI and PstI

1) Cut pQR412 with EcoRI and PstI
2) Ligate 0.5kb korB fragment into pKK223.3 cut with EcoRI and PstI
determine the exact position of the cop/korB gene in relation to the tac promoter. This is possible because pQR414 has the same EcoRI-PstI insert as pQR412 which was confirmed by restriction map analysis. Fig. 3.10a shows the sequence across the junction of the tac promoter and the cop/korB gene in pQR414, and verifies that the SpeI site had been completely filled in with nucleotides by the PolI.

Total mRNA was isolated from *E. coli* containing pQR414 and pUC8 (as a control) (Section 2.16), run on an agarose-formaldehyde gel, and blotted directly onto a Zeta-probe nylon membrane (Section 2.17). The nylon membrane was probed with [α-32P]dCTP-labelled 0.5kb cop/korB fragment isolated from pQR200 (Fig. 3.1) cut with SpeI and HindIII. As for mRNA isolated from *E. coli* containing pQR206 (Fig. 3.5), one band corresponding to a message of approximately 500b in size was detected in total mRNA isolated from *E. coli* containing pQR414 whilst no bands were detected in mRNA isolated from *E. coli* containing pUC8 (data not shown). This result confirms that transcription did take place at a detectable level and that a transcript of the expected size of 500b was actively synthesized.

However, when total protein extract made from *E. coli* harbouring pQR414 was run on an 18% SDS-polyacrylamide gel, no new protein band corresponding to 10kDa was visible when compared to *E. coli* total protein extract harbouring pKK223.3 (Fig. 3.11). Thus, production of cop/korB mRNA is not a guarantee that large quantities of protein will be synthesized. This result suggests that the *E. coli*-like AGGA Shine-Dalgarno sequence situated 11b upstream of the cop/korB ATG start-site may not be recognized by the host ribosomes. The Shine-Dalgarno site on pKK223.3 located downstream from the tac promoter is situated 55b away from the cop/korB ATG start codon. This difference far exceeds the optimum alignment of a Shine-Dalgarno sequence to the translational start of a gene, which is typically separated by 3 to 11b (Shine and Dalgarno 1974). Thus, although ribosomes will bind to the pKK223.3 AGGA Shine-Dalgarno sequence, they will eventually dissociate from the mRNA transcript because there is no AUG start codon located within 3 to 11b from the ribosomal binding site. In order for Cop/KorB protein
Figure 3.10
Sequence of pQR414 across the junction of the tac promoter and korB gene

A)

B)
Induced overnight cultures of *E. coli* JM107 containing pQR414 and pKK223.3 (as a control) respectively were sonicated in extraction buffer (Section 2.22). Approximately 80ug of each total protein sample prepared above was analysed by electrophoresis on an 18% SDS-polyacrylamide gel, after which the gel was stained with Coomassie blue and destained (Section 2.20).

Lane 1) Total protein extract from *E. coli* containing pKK223.3 (control).
Lane 2) Total protein extract from *E. coli* containing pQR414 showing a similar protein pattern to the control with no new band produced in the 10kDa range.
expression to take place in pKK223.3, the intervening 55bp sequence between the pKK223.3 ribosomal-binding site and the \textit{cop/korB} gene has to be removed to bring the ATG start-site of the \textit{cop/korB} gene within 3 to 11bp of the pKK223.3 ribosomal binding site.

Alternatively, even if the \textit{E. coli-like} AGGA Shine-Dalgarno sequence located on the cloned \textit{cop/korB} fragment is recognized by \textit{E. coli} ribosomes, the presence of an inverted repeat structure overlapping the AGGA sequence may cause the mRNA to form an internal secondary structure \textit{in vivo} which would make this potential Shine-Dalgarno sequence inaccessible to the ribosomes (Fig. 3.10b). Although this intra-strand annealing of the mRNA transcript might also occur for pQR206, the presence of the exposed \textit{llacZ} Shine-Dalgarno sequence to which the ribosomes can attach and their subsequent synthesis of a LacZ fusion protein should allow the ribosomes to continue to translate the mRNA transcript and unravel any internal secondary structure forming along the \textit{cop/korB} sequence. Because of the advantages of using pQR205 discussed in Section 3.3 and the production of relatively high levels of Cop/KorB protein from this construct, it was used in tests for binding activity.

3.7 Analysis of Cop/KorB binding.

Since indirect genetic evidence suggests that Cop/KorB binds to the pIJ101 \textit{kilB} and \textit{cop/korB} promoters \textit{in vivo} (Stein et al. 1989, Stein and Cohen 1990), it was desirable to determine the precise DNA sequences needed for recognition and binding. Initial attempts were made to purify Cop/KorB protein by running crude \textit{E. coli} extracts containing Cop/Kor3 down a Superose size fractionation column. However, when collected fractions were examined by SDS-PAGE, the relatively small Cop/KorB protein was co-eluting with several high molecular weight proteins. This may have been due to \textit{Cop/KorB} binding non-specifically to DNA which would affect its migration through the column. However, because of the lack of a suitable assay for detecting the Cop/KorB protein and thus the need to examine each
fraction from a column on SDS-PAGE systems which were relatively difficult to run due to the small size of Cop/KorB, it was decided to utilize crude protein extracts containing Cop/KorB for the binding studies. Once able to determine the exact sequence of the bound DNA, it would then theoretically be possible to devise a one-step purification procedure using affinity chromatography.

3.7.1 Gel retardation assay.

The gel retardation assay is based on the retarded mobility of protein-DNA complexes versus free DNA under non-denaturing PAGE (Fried and Crothers 1981). The gel retardation assay in this case allows a) the detection of the Cop/KorB DNA binding protein in *E. coli* total crude protein extracts, b) the general characterisation of the *kilB* and *cop/korB* DNA recognition sequences needed for Cop/KorB protein binding, and finally c) the definition of the optimal binding conditions for more detailed footprint analysis of the Cop/KorB protein-DNA complex.

The *kilB* gene was extracted on a 2.8kb *BcII*(10) to *BcII*(29) fragment isolated from pIJ101 (Fig. 1.2) and cloned into the *BamHI* site of pUC19, creating pQR427 (Fig. 3.12). pQR427 was then digested with *SalGI* (which removes a total of 1kb of DNA containing the *spdB* and *orf66* ORF's), and the largest 4.5kb band extracted from an agarose gel and religated, creating pQR428 (Fig. 3.12). pQR428 was then digested with *BalI* and *SmaI* (which removes a 1kb fragment containing the remaining part of the *rep* ORF and a non-coding region of pIJ101 DNA), and the larger 3.5kb fragment containing the *kilB* gene was extracted from a gel and religated, creating pQR456 (Fig. 3.12). pQR456 was then digested with *SalGI*, end-labelled with [γ-32P]ATP (Section 2.14) and finally digested with *SacII*, resulting in the 0.65kb *kilB* gene labelled at the *SalGI* end. The end-labelled *kilB* gene was isolated from a 1% agarose gel to remove any unincorporated nucleotides and to separate the 0.65kb *kilB* fragment from the 2.85kb fragment also present. The purified 0.65kb *SalGI*(19)-*SstII*(16) *kilB* probe was then incubated with crude
Figure 3.12  Construction of pQR427, pQR428 and pQR456

1) Cut pQR427 with SalGI
2) Religate 4.5kb fragment

pQR427 5.5kb

1) Cut pQR428 with Ball and Smal
2) Religate 3.5kb fragment

pQR428 4.5kb

1) Cut pQR456 with Ball and Smal
2) Religate 3.5kb fragment

pQR456 3.5kb
protein extract made from *E. coli* containing pQR206 and pUC8 respectively and run on a 5% acrylamide gel under non-denaturing conditions (Section 2.23).

Fig. 3.13(1) shows that after incubation with pUC8 and pQR206 crude protein extracts respectively, only the labelled 0.65kb *Sal*I(19)-*Sst*II(16) *kilB* probe incubated with Cop/KorB protein is retarded but not the probe incubated with pUC8 crude protein extract. This strongly suggests that the Cop/KorB protein is recognizing and specifically binding to the *kilB* fragment, and thus causing the probe to become retarded in its mobility resulting in a band shift. It should be noted that *E. coli* crude extracts containing Cop/KorB protein loose their activity when stored at -20°C even in the presence of 30% glycerol. Thus, protein extracts were always prepared fresh on the day when needed from induced overnight cultures of *E. coli* harbouring pQR206.

The Cop/KorB protein was next tested for its ability to bind to its own promoter. pQR412 (Fig. 3.9) was digested with *Eco*RI, end-labelled with [*γ-32P*]ATP (Section 2.14) and finally digested with *Hind*III, resulting in the 0.5kb *cop/korB* gene labelled at the *Eco*RI end. Gel-purified 0.5kb *Spe*I-cleaved *cop/korB* probe was incubated with *E. coli* crude extract containing pQR206 and with *E. coli* crude extract containing pUC8 (as a control), respectively. Fig. 3.13(2) shows that the Cop/KorB protein specifically retards the labelled 0.5kb *Spe*I-cleaved *cop/korB* probe although not as efficiently as with the 0.65kb *Sal*I(19)-*Sst*II(16) *kilB* probe, since there is still some unbound *cop/korB* probe remaining in the assay.

Initially, it seemed that the reduced ability of Cop/KorB protein to bind to the *cop/korB* probe was due to the manner in which pQR412 was constructed (Fig. 3.9); that is, using the *Spe*I(53) site to subclone the *cop/korB* fragment inadvertently cleaved too close to the *cop/korB* promoter region involved in Cop/KorB recognition and binding. In order to determine if this "truncated" *cop/korB* promoter was the cause of the reduced affinity of the Cop/KorB protein for pQR412, the *cop/korB* promoter was subcloned in such a manner as to allow for the
Gel retardation assay using crude Cop/KorB protein extract incubated with kilB and cop/korB (SpeI-cleaved) probes respectively.

10µl (approx. 40µg) of *E. coli* crude extract containing pUC8 and pQR206-generated protein products were incubated with 10-20,000cpm labelled kilB and cop/korB (SpeI-cleaved) probes respectively for 15min at room temperature before electrophoresing on a non-denaturing 5% acrylamide gel (Section 2.23). The gel was then dried down and exposed to X-ray film.

**Lanes**

1a) Free 0.65kb SalGI(19)-SacII(16) kilB probe (no protein present).
1b) 0.65kb kilB probe incubated with pUC8 crude protein extract showing no binding.
1c) 0.65kb kilB probe incubated with Cop/KorB crude protein extract showing complete binding and retardation of kilB probe (arrowed).

**Lanes**

2a) Free 0.5kb SpeI(53)-BclI(49) cop/korB probe (no protein present).
2b) 0.5kb cop/korB probe incubated with pUC8 crude protein extract showing no binding.
2c) 0.5kb cop/korB probe incubated with Cop/KorB crude protein extract showing only partial binding and retardation (top arrow). Bottom arrow shows residual amounts of unbound cop/korB probe.
presumed cop/korB promoter to be flanked by additional DNA sequence from pIJ101.

pQR200 (Fig. 3.1) was cut with EcoRI and HindIII, and the 1.2kb fragment containing the cop/korB gene isolated from an agarose gel (Section 2.13). This 1.2kb fragment was then digested to completion with Sau3A which resulted in the production of seven bands: 500bp, 200bp, 140bp, 110bp, 90bp, and two 40bp fragments. The 500bp Sau3A fragment containing the cop/korB promoter region was extracted from a gel and subcloned into the BamHI site of pUC19, creating pQR452 (Fig. 3.14). pQR452 was digested with HindIII, end-labelled with [γ-32P]ATP (Section 2.14) and finally digested with EcoRI, resulting in the 0.5kb cop/korB fragment labelled at the HindIII end. The 0.5kb Sau3A-cleaved cop/korB probe was then incubated with pQR206 and pUC8 crude protein extracts respectively and run on a non-denaturing 5% acrylamide gel (Section 2.23).

Fig. 3.15(2) shows that as with the 0.5kb Spel-cleaved cop/korB fragment (Fig. 3.13), this 0.5kb Sau3A-cleaved cop/korB fragment is only partially retarded when incubated with crude protein extract from E. coli harbouring pQR206 and is not retarded with pUC8 crude protein extracts. Thus, even with the presence of additional sequence adjacent to the cop/korB promoter region, Cop/KorB protein still does not bind to all the cop/korB probe in the assay. This effect is not due to excess cop/korB probe present compared to kilB probe since equivalent counts were added in each case (although it should be noted that the proportion of unlabelled DNA mixed in with the kilB and korB probes may differ). Furthermore, even with the reduction of added probe, there were still traces of unbound cop/korB probe present in the assay (Fig. 3.16). The differences in Cop/KorB protein concentration in crude samples is also not a viable explanation since the same batch of protein samples were used simultaneously to do cop/korB and kilB binding assays, with the latter fragment always being completely retarded while the former was on most occasions only partially retarded. It should be noted that the residual
Figure 3.14  Construction of pQR452

1) Cut pQR200 with EcoRI and HindIII
2) Isolate 2.0kb fragment and cut with Sau3A
2) Ligate 0.5kb Sau3A fragment into pUC19 cut with BamHI

![Diagram showing the construction process of pQR452]
Gel retardation assay using crude Cop/KorB protein extract incubated with kilB and cop/korB (Sau3A-cleaved) probes respectively.

10μl (approx. 40μg) of E. coli crude extract containing pUC8- and pQR206-generated protein products were incubated with 10-20,000cpm labelled kilB and cop/korB (Sau3A-cleaved) probes respectively for 15min at room temperature before electrophoresing on a non-denaturing 5% acrylamide gel (Section 2.23). The gel was then dried down and exposed to X-ray film.

Lanes

1a) Free 0.65kb SalGI(19)-SacII(16) kilB probe (no protein present).
1b) 0.65kb kilB probe incubated with pUC8 crude protein extract showing no binding.
1c) 0.65kb kilB probe incubated with Cop/KorB crude protein extract showing complete binding and retardation of kilB probe (arrowed).

Lanes

2a) Free 0.5kb Sau3A cop/korB probe (no protein present).
2b) 0.5kb cop/korB probe incubated with pUC8 crude protein extract showing no binding.
2c) 0.5kb cop/korB probe incubated with Cop/KorB crude protein extract showing only partial binding and retardation (top arrow). Bottom arrow indicates residual unbound cop/korB probe.
Gel retardation assay using crude Cop/KorB protein extract incubated with different concentrations of cop/korB (Sau3A-cleaved) probe.

10μl (approx. 40μg) of E. coli crude protein extract containing pUC8- and pQR206-generated protein products were incubated with a) 5μl (20,000cpm) labelled cop/korB (Sau3A-cleaved) and b) 1μl (4000cpm) labelled cop/korB (Sau3A-cleaved) probes respectively for 15min at room temperature before electrophoresing on a non-denaturing 5% acrylamide gel (Section 2.23). The gel was then dried down and exposed to X-ray film.

Lanes
1a) 5μl (20,000cpm) free 0.5kb Sau3A cop/korB probe.
1b) 5μl 0.5kb cop/korB probe incubated with pUC8 crude protein extract.
1c) 5μl 0.5kb cop/korB probe incubated with Cop/KorB crude protein extract showing partial binding and retardation of cop/korB probe (arrow 1). Arrow 2 shows that approximately half of the added cop/korB probe remains unbound.

Lanes
2a) 1μl (5000cpm) free 0.5kb Sau3A cop/korB probe.
2b) 1μl 0.5kb cop/korB probe incubated with pUC8 crude protein extract.
2c) 1μl 0.5kb cop/korB probe incubated with Cop/KorB crude protein extract showing incomplete binding and retardation of cop/korB probe (arrow 1). Arrow 2 indicates that only residual amounts of cop/korB probe remains unbound with the reduction of added probe.
unbound probe in the cop/korB assay migrates somewhat slower than free probe in some cases. The reason for this is unknown.

Deng et al. (1988) have stated that the cop gene product decreases the copy number of Sti+ plasmids by inhibiting the initiation of second-strand synthesis at the sti locus. If this is correct, it is assumed that Cop protein would bind to the double-stranded form of sti rather than the single-stranded sti intermediate, as the latter case would result in the undesirable accumulation of ssDNA in the cell which is not normally observed in pIJ101. The sti determinant was subcloned on a 0.7kb SpeI-EcoRI fragment from pQR200 into pUC19 cut with XbaI and EcoRI, creating pQR417 (Fig. 3.17). pQR417 was digested with EcoRI and HindIII, the 5' overhanging ends filled in with [α-32P]dCTP using T7 DNA Polymerase, and the 0.7kb labelled sti fragment purified by extraction from an agarose gel (Section 2.14).

Fig. 3.18 shows this 0.7kb SpeI(53)-BclI(57) sti probe incubated with Cop/KorB crude protein extract and pUC8 crude protein extract (as a control) resolved on a non-denaturing 5% acrylamide gel (Section 2.23). There is no detectable retarded band present when sti is incubated with Cop/KorB protein, suggesting that Cop/KorB protein does not recognize or bind to the double-stranded form of the sti determinant. (It should be noted that a sample from the same batch of Cop/KorB crude protein extract used in the sti retardation assay was able to completely retard a 0.65kb kilB probe, confirming that the crude protein extract contained active Cop/KorB protein.) Since sti has no effect on Cop/KorB, these results also reconfirm that binding to labelled kilB and cop/korB target DNA is sequence-specific.

In conclusion, Cop/KorB protein has specific recognition for the 0.65kb SalGI(19)-SstII(16) kilB fragment, the 0.5kb SpeI(53)-BclI(49) cop/korB fragment, and the 0.5kb Sau3A cop/korB fragment, but not for the double-stranded 0.7kb SpeI(53)-BclI(57) sti fragment. Secondly, Cop/KorB seems to have a greater
Figure 3.17  Construction of pQR417, pQR437, and pQR438.

1) Cut pQR200 with Spel and EcoRI
2) Ligate 0.7kb sti fragment into pUC19 cut with Xbal and EcoRI

1) Cut pQR417 with SacI and HindIII
2) Ligate 0.6kb sti fragment into pUC19 cut with SacI and HindIII

1) Cut pQR437 with SacI and SacII
2) Ligate 3.23kb fragment to 14bp SacI-SacII pBS polylinker

pQR417
3.4kb

Ap^r

pQR437
3.3kb

Ap^r

pQR438
3.23kb

placZ

<orf85
orf79>
Figure 3.18
Gel retardation assay using crude Cop/KorB protein extract incubated with sti probe.

10μl (40μg) of E. coli crude extract containing pUC8- and pQR206-generated protein products were incubated with 10-20,000cpm labelled sti probe for 15min at room temperature before electrophoresing on a non-denaturing 5% acrylamide gel (Section 2.23). The gel was then dried down and exposed to X-ray film.

Lanes
a) Free 0.7kb SpeI(53)-BclII(57) sti probe (no protein present).
b) 0.7kb sti probe incubated with pUC8 crude protein extract showing no binding.
c) 0.7kb sti probe incubated with Cop/KorB crude protein extract also showing no binding of Cop/KorB protein to sti.
affinity for the kilB fragment than for the cop/korB fragments. Finally, the SpeI(53) site in the cop/korB fragment probably lies outside the boundaries of the operator sequence of the Cop/KorB protein since there is no marked difference in the binding pattern of Cop/KorB protein to cop/korB DNA containing or lacking the SpeI site.

3.7.2 Competition assay.

Sequence-specificity of Cop/KorB binding to kilB and cop/korB DNA fragments can be demonstrated in a competition assay. The concentration of Cop/KorB protein in the reaction has to be the limiting factor in order for the competition assay to work. The Cop/KorB protein is then pre-incubated with increasing amounts of cold target DNA prior to the addition of labelled probe.

Fig. 3.19 shows the results of a competition assay using 2μl of crude protein extract incubated with varying concentrations of kilB and cop/korB DNA (Section 2.24). Lane A)0 shows that when limiting amounts of Cop/KorB protein are incubated with the labelled 0.65kb SacII(16)-SalGI(19) kilB probe, a faint retarded band can be seen which represents the Cop/KorB protein recognizing and binding to some of the kilB probe. Since the majority of the kilB probe is unbound, this indicates that the concentration of Cop/KorB protein in the assay is the limiting factor, and it follows that should Cop/KorB protein concentration be increased, the amount of free probe in the assay should decrease as more of the probe becomes bound and thus retarded by the protein. Lane A)10X shows that when approximately 10-times unlabelled kilB DNA over kilB probe is added to the assay, this cold kilB DNA competes with the labelled kilB probe for Cop/KorB binding and thus reduces the amount of labelled kilB probe retarded in the gel. Lane A)50X shows that when 50-times unlabelled kilB DNA has been added to the assay, the cold DNA has completely "out-competed" labelled kilB probe for Cop/KorB binding as there is no detectable retardation of labelled kilB probe.
Figure 3.19

**Competition assay using 2μl of crude Cop/KorB protein extract preincubated with increasing concentrations of unlabelled kilB and cop/korB (Sau3A-cleaved) DNA.**

2μl (8μg) of *E. coli* crude extract containing pQR206-generated protein products was preincubated with increasing amounts of unlabelled kilB and cop/korB DNA (ranging from 10 to 50-fold excess over labelled probe) for 10min at room temperature prior to the addition of labelled kilB and cop/korB probe respectively, and incubated for a further 10mins (Section 2.24). Samples were electrophoresed on a non-denaturing 5% acrylamide gel which was then dried down and exposed to X-ray film.

Lane 0 in both the kilB (A) and cop/korB (B) reactions is the control in which no cold DNA was added to the Cop/KorB crude protein extract. The arrows indicate that some of the 0.65kb *SalG1*19-*SacI*16 kilB probe and 0.5kb Sau3A cop/korB probe respectively is recognized and bound by the limiting concentrations of Cop/KorB protein in the assay. The addition of 50X unlabelled kilB DNA is sufficient to completely out-compete labelled kilB probe for Cop/KorB binding, whereas the addition of only 10X unlabelled cop/korB DNA is needed to completely out-compete labelled cop/korB probe for Cop/KorB binding.
Lane B)0 (Fig. 3.19) shows the labelled 0.5kb Sau3A cop/korB probe incubated with limiting amounts of Cop/KorB protein in this control track. Although there is a distinct retarded band present when kilB probe is incubated with limiting amounts of Cop/KorB protein, the equivalent amount of cop/korB probe incubated with the same limiting concentration of Cop/KorB protein results in a much weaker retardation in the control track. When 10-times unlabelled cop/korB DNA is added to the assay, the cold DNA has completely out-competed labelled cop/korB probe for Cop/KorB binding as shown by the absence of a detectable retarded band in this track. Thus, the addition of only 10-times unlabelled cop/korB DNA to the assay is sufficient to completely abolish Cop/KorB binding to labelled probe, while the addition of at least 50-times unlabelled kilB DNA is needed to efficiently compete out Cop/KorB binding to kilB probe. This suggests that Cop/KorB protein may have a relatively lower affinity for the cop/korB DNA compared to the kilB DNA.

The competition assay (Section 2.24) was repeated for the labelled 0.5kb Sau3A cop/korB probe using 5ul of crude protein extract instead of 2ul, in order to obtain a more definitive competition pattern for cop/korB (Fig. 3.20). Track C)0 shows the control where Cop/KorB protein has been incubated with 0.5kb Sau3A cop/korB probe only. Because a relatively high amount of protein has been used compared to the concentration of probe, the cop/korB probe has been completely bound and retarded by the Cop/KorB protein. The cop/korB probe only starts to become out-competed by the unlabelled cop/korB DNA when 30-times of the latter has been added to the assay. It should be stressed that these competition experiments are not quantitative since the absolute quantities of Cop/KorB protein in the E. coli crude extract and the exact ratio of labelled:unlabelled DNA in the probe preparation are not known.

In conclusion, the competition assay demonstrates that Cop/KorB binding to kilB and cop/korB probes is sequence-specific since binding can be successfully out-competed by the addition of unlabelled kilB and cop/korB DNA respectively. Furthermore, the Cop/KorB protein may have a stronger affinity for the kilB gene
Figure 3.20

Competition assay using 5μl of crude Cop/KorB protein extract preincubated with increasing concentrations of cop/korB (Sau3A-cleaved) DNA.

5μl (20μg) of *E. coli* crude protein extract containing pQR206-generated protein products was preincubated with increasing amounts of unlabelled cop/korB DNA for 10min at room temperature prior to the addition of labelled cop/korB probe and incubated for a further 10mins (Section 2.24). Samples were electrophoresed on a non-denaturing 5% acrylamide gel which was then dried down and exposed to X-ray film. Lane 0 is the control in which no cold DNA was added to the Cop/KorB crude extract. The arrow indicates that the Cop/KorB protein has bound to all the 0.5kb Sau3A cop/korB probe in the reaction. Preincubation of Cop/KorB protein with approximately 10 to 50X unlabelled cop/korB DNA prior to the addition of labelled cop/korB probe results in the decrease of Cop/KorB binding to labelled probe.
compared to the cop/korB gene since Cop/KorB binding appears to be more rapidly titrated out with the addition of unlabelled cop/korB DNA than with the addition of unlabelled kilB DNA respectively.

3.7.3 Identification of cop/korB and kilB recognition sequences.

Several plasmid constructs were made in order to estimate where Cop/KorB protein was binding along the 0.65kb SalGI(19)-SacII(16) kilB and 0.5kb Sau3A cop/korB DNA fragments used for the gel retardation assays (Section 3.7.1). This information was important for the subsequent DNA footprinting experiments since only a maximum of 250 bases can be clearly resolved on sequencing gels at any one time, and thus DNA probes must contain the recognition sequence for Cop/KorB protein binding within 250 bases or less from their labelled end.

Since the SpeI(53) site was previously demonstrated to be situated outside the cop/korB recognition sequence for Cop/KorB binding (Section 3.7.1), it was essential to investigate whether the FspI(52) site spanning the putative -35 region of the cop/korB promoter and lying immediately adjacent to the SpeI(53) site was necessary for recognition and binding. pQR412 (Fig. 3.9) was digested with FspI and HindIII, and the 0.5kb cop/korB fragment cloned into pUC18 cut with Smal and HindIII, creating pQR458 (Fig. 3.21). pQR458 was cleaved with EcoRI and HindIII, the 0.5kb cop/korB fragment with a truncated promoter was extracted from an agarose gel (Section 2.13) and the 5' overhanging ends filled in with [α-32P]dATP using T7 DNA polymerase (Section 2.14). In order to remove the entire proposed cop/korB promoter region, pQR458 was digested with XhoII and the 0.45kb fragment cloned in both orientations into pUC18 cut with BamHI, creating pQR459 (Fig. 3.21) and pQR460 respectively. pQR459 was digested with EcoRI and HindIII, the promoterless 0.45kb cop/korB fragment extracted from a gel (Section 2.13) and the ends labelled with [α-32P]dATP using T7 DNA polymerase (Section 2.14).
Construction of pQR458 and pQR459

1) Cut pQR412 with FspI and HindIII
2) Ligate 0.5kb korB fragment into pUC18 cut with SmaI and HindIII

1) Cut pQR458 with Xhol II
2) Ligate 0.45kb fragment to pUC18 cut with BamHI

pQR458
3.2kb

pQR459
3.15kb
Both the labelled 0.5kb cop/korB probe with the FspI-truncated promoter and the 0.45kb promoterless cop/korB probe were incubated with Cop/KorB crude protein extract and run on a non-denaturing 5% acrylamide gel (Section 2.23). Fig. 3.22 shows that neither of these probes are recognized by Cop/KorB protein as there are no retarded bands visible in lanes B4 and B6 respectively. Lane B2 is a control showing Cop/KorB protein binding to the original 0.5kb cop/korB Sau3A probe which confirms that the crude extract contains active Cop/KorB protein. Thus, this result indicates that the DNA sequence lying immediately adjacent to the FspI(52) site is important in recognition and binding of the Cop/KorB protein, and that binding is taking place within a 50bp region located between the SpeI(53) and the XhoII sites and encompassing the cop/korB promoter region (Fig. 3.2). Stein et al. (1989) have also identified, by deletion analysis, a similar region needed for promoter activity and regulation in vivo spanning from the SpeI site to a Sau3A site (which overlaps the XhoII site). Furthermore, Stein et al (1989) have reported that insertion of a ClaI linker into the FspI site resulted in loss of korB promoter activity in vivo.

In an attempt to identify the boundaries of the kilB recognition site, pQR456 (Fig. 3.12) was digested with FspI and EcoRI and the 0.73kb kilB fragment cloned into pUC19 cut with SmaI and EcoRI, creating pQR457 (Fig. 3.23). As with the cop/korB promoter, the FspI(18) site lies within the putative -35 region of the kilB promoter, and is approximately 120bp from the SalGI(19) site which defines one end of the upper limit of the kilB promoter in pQR456. pQR457 was cleaved with EcoRI and HindIII, the 0.73kb kilB gene with a truncated promoter isolated from a gel (Section 2.13) and the 5' overhanging ends labelled with [$\alpha$-32P]dATP using T7 DNA polymerase (Section 2.14). To remove the entire proposed promoter region of kilB, pQR457 was cleaved with BstEII and EcoRI. This 0.66kb promoterless fragment was extracted from a gel (Section 2.13) and labelled at its EcoRI site with [$\alpha$-32P]dATP using T7 DNA polymerase (Section 2.14).
Gel retardation assay using crude Cop/KorB protein extract incubated with \( \text{kilB} \) and \( \text{cop/korB}-\text{derived probes} \).

10\( \mu l \) (40\( \mu g \)) of \( \text{E. coli} \) crude extract containing pQR206-generated protein products were incubated with 10-20,000 cpm labelled \( \text{kilB} \) and \( \text{cop/korB}-\text{derived probes} \) for 15 min at room temperature before electrophoresing on a non-denaturing 5% acrylamide gel (Section 2.23). The gel was then dried down and exposed to X-ray film.

Lanes

A1) Free 0.65kb \( \text{SalI(19)-SacII(16)} \) \( \text{kilB} \) probe (no protein present).
A2) 0.65kb \( \text{kilB} \) probe incubated with Cop/KorB crude protein extract showing complete binding and retardation of the \( \text{kilB} \) probe.
A3) Free 0.73kb \( \text{FspI(18)-cleaved} \) \( \text{kilB} \) probe (no protein present).
A4) 0.73kb \( \text{kilB} \) probe incubated with Cop/KorB crude protein extract showing no binding since cleavage with \( \text{FspI} \) disrupts the -35 region of \( \text{kilB} \) promoter.
A5) Free 0.66kb \( \text{BstEII(17)-cleaved} \) \( \text{kilB} \) probe (no protein present).
A6) 0.66kb \( \text{kilB} \) probe incubated with Cop/KorB crude protein extract showing no binding since cleavage with \( \text{BstEII} \) removes the entire \( \text{kilB} \) promoter.

Lanes

B1) Free 0.5kb \( \text{Sau3A cop/korB} \) probe (no protein present).
B2) 0.5kb \( \text{cop/korB} \) probe incubated with Cop/KorB crude protein extract showing partial binding and retardation of \( \text{cop/korB} \) probe.
B3) Free 0.5kb \( \text{FspI(52)-cleaved cop/korB} \) probe (no protein present).
B4) 0.5kb \( \text{cop/korB} \) probe incubated with Cop/KorB crude protein extract showing no binding since cleavage with \( \text{FspI} \) disrupts the -35 region of the \( \text{cop/korB} \) promoter.
B5) Free 0.45kb \( \text{XhoII-cleaved cop/korB} \) probe (no protein present).
B6) 0.45kb \( \text{cop/korB} \) probe incubated with Cop/KorB crude protein extract showing no binding since cleavage with \( \text{XhoII} \) removes the entire \( \text{cop/korB} \) promoter.
Figure 3.23 Construction of pQR457

1) Cut pQR456 with FspI and EcoRI
2) Ligate 0.73kb kilB fragment into pUC19 cut with SmaI and EcoRI
Both the labelled 0.73kb \textit{ki}l\textit{B} probe with the \textit{FspI}-truncated promoter and the 0.66kb promoterless \textit{ki}l\textit{B} probe were incubated with Cop/KorB crude protein extract and run on a non-denaturing 5\% acrylamide gel (Section 2.23). Fig. 3.22 indicates that Cop/KorB protein does not recognize or bind to either of these \textit{ki}l\textit{B} probes which is shown by the lack of retarded bands in lanes A4 or A6 compared to the free probe in lanes A3 and A5 respectively. Lane A2 shows the original 0.65kb \textit{SalGI}(19)-\textit{SacII}(16) \textit{ki}l\textit{B} fragment being recognized, bound, and thus retarded by active Cop/KorB protein as a control. Thus, these results suggest that the \textit{FspI}(18) site spanning the -35 region of the \textit{ki}l\textit{B} promoter is important for Cop/KorB recognition and binding, and that binding occurs in the region of the 180bp \textit{SalGI}(19) to \textit{BstEII}(17) section of DNA containing the \textit{ki}l\textit{B} promoter. Stein et al. (1989) have found that a fragment spanning from the first "T" of the \textit{FspI} site (bp2501 in Appendix I) to the \textit{BstEII} site can be recognized by KorB \textit{in vivo}, and insertion of a \textit{ClaI} or \textit{HpaII} linker into the \textit{FspI} site resulted in loss of \textit{ki}l\textit{B} promoter activity.

\section*{3.8 DNaseI protection analysis (DNA footprinting).}

DNaseI protection analysis (DNA footprinting) is a sensitive, accurate technique used to determine the precise location of protein-binding sites within DNA fragments (Galas and Schmitz 1978). Essentially, an end-labelled DNA fragment containing the specific recognition sequence is incubated with the DNA-binding protein. This nucleoprotein complex is then mildly digested with DNaseI which randomly cleaves the DNA except where inhibited by the binding of the protein. Electrophoretic comparison of the cleavage products generated from naked DNA and from the same DNA complexed with protein reveals "gaps" in the binding pattern of the latter (known as a footprint), where the DNA has been protected by the bound protein. It is essential for the DNA fragments of interest to be labelled at only one end and for the footprint to be less than 250b from the labelled end as
sequencing gel electrophoresis can only resolve a maximum of 200-250b as discussed in Section 3.7.3.

pQR456 (Fig. 3.12) was digested with SalGI, end-labelled with \( \gamma^{32}\text{P} \)ATP using T4 PNK (Section 2.14) and finally digested with SacII, resulting in the 0.65kb \( \text{kilB} \) gene labelled specifically at the SalGI end of the non-coding strand. In order to label the \( \text{kilB} \) coding strand, pQR456 was digested with HindIII, labelled with \( \alpha^{32}\text{P}\)dATP using T7 DNA polymerase (Section 2.14), and finally digested with SacII. pQR452 (Fig. 3.14) was digested with HindIII, end-labelled with \( \gamma^{32}\text{P} \)ATP using T4 PNK (Section 2.14) and finally digested with EcoRI, resulting in the 0.5kb \( \text{cop/korB} \) gene labelled specifically at the HindIII end of the coding strand. To label the \( \text{cop/korB} \) non-coding strand, pQR452 was digested with HindIII and the 5' overhanging end filled in with \( \alpha^{32}\text{P}\)dATP using T7 DNA polymerase (Section 2.14), and finally digested with EcoRI. Fig. 3.24 and 3.25 show the results of footprinting experiments (Section 2.25) carried out on the end-labelled 0.65kb \( \text{kilB} \) and 0.5kb \( \text{cop/korB} \) genes. The results of these footprinting experiments are summarized schematically in Fig. 3.26 and shows that the operator sites for the Cop/KorB protein correspond with the putative \( \text{kilB} \) and \( \text{cop/korB} \) promoters.

The pattern of protection obtained for the \( \text{cop/korB} \) gene indicates that the Cop/KorB protein binds to staggered 33b regions on the coding (Fig. 3.24) and non-coding (Fig. 3.25) strands respectively. This protected region corresponds to the location of the proposed \( \text{cop/korB} \) promoter. For both strands there exists a hypersensitive site which lies exactly halfway between the protected 33b regions, dividing them into two equal blocks of 16b each (although the hypersensitive site on the coding strand is more pronounced than the one present on the non-coding strand). Hypersensitive sites are often indicative of bending of the DNA induced by the bound protein. Thus, Cop/KorB protein when binding to its own promoter, could produce a "kink" in the DNA which is then exposed to DNaseI activity. The Cop/KorB protein protects the entire -10 region of its promoter and the mapped start of transcription on the coding strand, but only protects part of the -35 region. The
Figure 3.24 (cont’d overleaf)

DNasel protection analysis of pQR206-derived Cop/KorB protein extract incubated with 0.65kb kilB and 0.5kb cop/korB probes end-labelled with T4 Polynucleotide Kinase.

20μl (80μg) of E. coli crude extract containing pQR206-generated protein products were incubated with 20,000cpm 0.65kb kilB and 0.5kb cop/korB end-labelled probes respectively (see text for detail) for 15min at room temperature, then subjected to DNaseI digestion for 1min before reactions were terminated (Section 2.25). Purified samples were electrophoresed on a 8% sequencing gel using pQR412 (sequenced using the 17b reverse primer 5’dCAGGAAACAGCTATGAC) as a size marker (Section 2.19). Gels were dried down prior to autoradiography.

Lanes 1b) 0.65kb end-labelled kilB::Cop/KorB nucleoprotein complex digested with 0.5μg/ml final DNasel concentration.
1c) 0.65kb end-labelled kilB::Cop/KorB nucleoprotein complex digested with 0.05μg/ml final DNasel concentration.
1d) 0.65kb end-labelled kilB probe (no protein) digested with 0.005μg/ml final DNasel concentration.
1e) 0.65kb end-labelled kilB probe (no protein) digested with 0.0005μg/ml final DNasel concentration.

Lanes 1b and 1c show that a 52b region on the kilB non-coding strand (encompassing the kilB promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the band pattern generated by DNaseI cleavage of the kilB probe alone.

Lanes 2b) 0.5kb end-labelled cop/korB::Cop/KorB nucleoprotein complex digested with 0.5μg/ml final DNasel concentration.
2c) 0.5kb end-labelled cop/korB::Cop/KorB nucleoprotein complex digested with 0.05μg/ml final DNaseI concentration.
2d) 0.5kb end-labelled cop/korB probe (no protein) digested with 0.005μg/ml final DNaseI concentration.
2e) 0.5kb end-labelled cop/korB probe (no protein) digested with 0.0005μg/ml final DNaseI concentration.

Lanes 2b and 2c show that a 33b region of the cop/korB coding strand (overlapping the cop/korB promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the pattern generated by DNaseI cleavage of free cop/korB probe. Furthermore, there exists a protein-induced hypersensitive site situated in the centre of the 33b protected region, dividing it into two equal blocks of 16b each.
DNase I protection analysis of pQR206-derived Cop/KorB protein extract incubated with 0.65kb *kilB* and 0.5kb *cop/korB* probes filled-in with T7 DNA polymerase.

20 µl (80 µg) of *E. coli* crude protein extract containing pQR206-generated protein products were incubated with 20,000 cpm 0.65kb *kilB* and 0.5kb *cop/korB* filled-in probes respectively (see text for detail) for 15 min at room temperature, then subjected to DNase I digestion for 1 min before reactions were terminated (Section 2.25). Purified samples were electrophoresed on an 8% sequencing gel using pQR412 (sequenced using the 17b reverse primer 5' dCAGGAAACAGCTATGAC) as a size marker (Section 2.19). Gels were dried down prior to autoradiography.

Lanes 1a) 0.65kb filled-in *kilB*:Cop/KorB nucleoprotein complex digested with 5 µg/ml final DNase I concentration.
1b) 0.65kb filled-in *kilB*:Cop/KorB nucleoprotein complex digested with 0.5 µg/ml final DNase I concentration.
1c) 0.65kb filled-in *kilB*:Cop/KorB nucleoprotein complex digested with 0.05 µg/ml final DNase I concentration.
1d) 0.65kb filled-in *kilB* probe (no protein) digested with 0.005 µg/ml final DNase I concentration.
1e) 0.65kb filled-in *kilB* probe (no protein) digested with 0.0005 µg/ml final DNase I concentration.

Lanes 1b and 1c show that a 60 bp region of the *kilB* coding strand (encompassing the *kilB* promoter) has been protected from DNase I digestion by the bound Cop/KorB protein compared to the band pattern generated by DNase I cleavage of the *kilB* probe alone.

Lanes 2a) 0.5kb filled-in *cop/korB*:Cop/KorB nucleoprotein complex digested with 5 µg/ml final DNase I concentration.
2b) 0.5kb filled-in *cop/korB*:Cop/KorB nucleoprotein complex digested with 0.5 µg/ml final DNase I concentration.
2c) 0.5kb filled-in *cop/korB*:Cop/KorB nucleoprotein complex digested with 0.05 µg/ml final DNase I concentration.
2d) 0.5kb filled-in *cop/korB* probe (no protein) digested with 0.005 µg/ml final DNase I concentration.
2e) 0.5kb filled-in *cop/korB* probe (no protein) digested with 0.0005 µg/ml final DNase I concentration.

Lanes 2a, 2b, and 2c show that a 33 bp region on the *cop/korB* non-coding strand (overlapping the *cop/korB* promoter) has been protected from DNase I digestion by the bound Cop/KorB protein compared to the pattern generated by DNase I cleavage of free *cop/korB* probe. There also exists a protein-induced DNase I hypersensitive site situated in the middle of the 33 bp protected region, although it is not as pronounced as the site present on the coding strand (Fig. 3.24).
Figure 3.26  Summary of kilB and korB footprint data

A) kilB

2532
c CCGGAGCCCGGACCCGACCCGACCGCGACCGACTGACGAGTCACTG
n-c GGGGGGACGGGGCGACGGGGCGACGGGGCGACGGGGCGAGTCACTG

B) korB

7274
c CAACATGGGCGAATTTTGGACGATTGACGGGACCTAGACGACGAGTCACTG
n-c GTGCACTTGTGACGACGACGACGACGAGTCACTG

Nucleotides in spotted boxes show area of DNA protected by KorB binding and indicate KorB operator sequence on kilB and korB DNA respectively. The asterisk shows the location of the hypersensitive sites in korB. Lined boxes with roman numerals show regions of homology between kilB and korB operator sites. Horizontal arrows indicate inverted repeats. Vertical arrows indicate restriction enzyme cleavage sites (in italics). Horizontal arrows with filled circles show mapped transcription start sites.
protected pattern of the non-coding strand is exactly the same as the coding strand except that it is shifted 3b upstream in relation to the latter, and thus the entire -10 and all but the last base of the -35 region is 

Thus, cleaving the cop/korB promoter region at the FspI(52) site which lies within the -35 region will disrupt the Cop/KorB recognition sequence and eliminate Cop/KorB binding since Cop/KorB protein can no longer recognize the truncated operator site. This data supports the negative results obtained from the gel retardation assay when Cop/KorB protein was incubated with FspI-cleaved cop/korB fragment (Section 3.7.3), and verifies the conclusion that Cop/KorB protein was unable to recognize the truncated cop/korB promoter. However, since the SpeI(53) site is situated outside the operator sequence for Cop/KorB binding, SpeI-cleaved DNA fragments are still being actively recognized and bound by Cop/KorB protein as demonstrated in the gel retardation assays (Section 3.7.1). The footprint data for cop/korB also supports the conclusions reported by Stein et al. (1989), which state that KorB protein is able to repress transcription in vivo from a 61bp SpeI-Sau3A fragment possessing the korB promoter and that the insertion of a Clal linker into the FspI site in the putative -35 region of the korB promoter results in loss of promoter activity.

The pattern of protection obtained for the kilB gene suggests that the Cop/KorB protein protects a relatively large region spanning 52b on the non-coding strand (Fig. 3.24) and 60b on the coding strand (Fig. 3.25). The 60b protected region of the coding strand totally encompasses the 52b protected region on the opposite strand. Both the proposed -10 and -35 regions are completely protected as well as the major transcription start site. It is interesting to note that the kilB protected area is almost twice the size of the cop/korB protected area. One reason for the larger protected area for kilB may be because of the existence of a weaker secondary transcription start site upstream of the putative -35 region (Stein et al. 1989), which would require Cop/KorB protein to block transcription from both potential promoters in order to efficiently regulate kilB expression.
The footprinting data for the *kilB* gene confirms that the *FspI*(18) site lies within the operator region and is important if the integrity of Cop/KorB protein binding to the *kilB* promoter is to be maintained. This is also demonstrated by the gel retardation assays where Cop/KorB protein was unable to recognize and bind to a *kilB* DNA fragment cleaved at the *FspI*(18) site (Section 3.7.3). However, this footprinting data contradicts the conclusion made by Stein *et al.* (1989) that the KorB protein was able to repress transcription *in vivo* from a 76bp *FspI-BstEII* fragment containing the *kilB* promoter, since the protected region stretches at least 12b beyond the last "T" of the *FspI* site to include the weaker secondary transcription start site.

The *kilB* and *cop1korB* operator regions show significant sequence similarity (as would be expected if the same protein would have to bind and regulate both promoters), and there are at least nine blocks of completely homologous sequence which span the inverted repeat along the promoter region of these genes (Fig. 3.26). This conserved region comprising of the nine homologous blocks is 36 bp long in total; the *cop1korB* operator site seems to be virtually enclosed within this region while the *kilB* operator site extends 4 to 14b beyond the boundaries of the conserved region (Fig. 3.26). It is interesting to note that block I in the *kilB* operator (Fig. 3.26) is completely protected by Cop/KorB binding whereas the identical block in *cop1korB* is only partially protected, even though the sequence of the *kilB* and *cop1korB* DNA is identical in this region. It is also interesting to note that the hypersensitive sites in *cop1korB* lie almost in the centre of the inverted repeat located in the conserved region, while there are no observable hypersensitive sites at the equivalent points in *kilB*. The inverted repeat of the *kilB* gene has a higher degree of complementary intra-strand base-pairing (10/12 perfect matches) compared to the inverted repeat of the *cop1korB* gene (8/12 perfect matches), and it may be this fact together with the differences in protected sequence bordering the conserved region in *kilB* which result in the latter operator have a relatively higher
affinity for Cop/KorB binding as suggested by the competition assay (Section 3.7.2).

It should be noted that Stein and Cohen (1990) have cloned the korB gene on a 1.1kb BglIII-SpeI fragment into pUC19 cut with BamHI and XbaI, creating pDSS172, which they have shown to produce a 10kDa fusion protein in maxicells. More importantly, the KorB fusion protein product from pDSS172 was shown to be functionally active and able to repress transcription from the kilB promoter in vivo, since E. coli cells containing a lacZ gene under the control of the kilB promoter showed decreased LacZ expression when pDSS172 was present (Stein and Cohen 1990). Thus, the existence of the additional 27 amino acids at the N-terminal end of the pDSS172-derived KorB fusion protein does not seem to affect repressor activity.

Although the pQR206-generated Cop/KorB fusion protein has also been shown to be active and capable of specifically recognizing and binding to the kilB and cop/korB promoter sequences, it was possible that the additional 19 amino acids located at the N-terminal end of the protein could interfere with the footprinting pattern. Since the pDSS172-derived KorB fusion protein with an additional 27 amino acids located at the N-terminal end of the protein has been shown to be functionally active and stable in vivo (Stein and Cohen 1990), a similar plasmid construct was made in order to compare the footprinting pattern generated from both types of LacZ-Cop/KorB fusion proteins. Should the foreign amino acids be involved in non-specific binding, the Cop/KorB fusion protein with the longer foreign amino acids chain will cause a larger area of DNA to be protected from DNaseI digestion. Should the foreign amino acids not interfere with binding, the protected pattern emerging from DNaseI digestion will be identical for both types of LacZ-Cop/KorB fusion proteins.

pIJ101 (Fig. 1.2) was digested with BglIII and SpeI and the 1.1kb BglII(45)-SpeI(53) fragment force cloned into pUC19 cut with BamHI and XbaI, creating pQR461 (Fig. 3.27a). pQR461 produces an in-frame fusion protein with an additional 27 amino acids at its N-terminal end calculated from sequence analysis.
Figure 3.27b

Sequence across the pQR461 SpeI/XbaI junction

ATGACCATGATTACGCCAAGCTTGCAT
MetThrMetIleThrProThrLeuHis

XbaI/SpeI
lacZ\korB>
GCCTGCAGGTCGACTCTAGTTGCGCAG
AlaLysArgSerThrLeuValAlaGln

ACTGACACAGTCGGTCAGGATGACTTCATG
ThrAspThrValGlyGlnAspAspPheMet
DNasel protection analysis of pQR461-derived Cop/KorB protein extract incubated with 0.65kb *kilB* and 0.5kb *cop/korB* probes end-labelled with T4 Polynucleotide Kinase.

20μl (80μg) of *E. coli* crude extract containing pQR461-generated protein products were incubated with 20,000cpm 0.65kb *kilB* and 0.5kb *cop/korB* end-labelled probes respectively for 15min at room temperature, then subjected to DNaseI digestion for 1min before the reactions were terminated (Section 2.25). Purified samples were electrophoresed on a denaturing 8% polyacrylamide gel using pQR412 (sequenced using the 17b reverse primer 5’dCAGGAAACAGCTATGAC) as a size marker (Section 2.19). Gels were then dried down prior to autoradiography.

Lanes 1b) 0.65kb end-labelled *kilB*:Cop/KorB nucleoprotein complex digested with 0.5μg/ml final DNasel concentration.
1e) 0.65kb end-labelled *kilB* probe (no protein) digested with 0.0005μg/ml final DNasel concentration.

Lane 1b shows that a 52b region on the *kilB* non-coding strand (encompassing the *kilB* promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the band pattern generated by DNaseI cleavage of the *kilB* probe alone.

Lanes 2b) 0.5kb end-labelled *cop/korB*:Cop/KorB nucleoprotein complex digested with 0.5μg/ml final DNasel concentration.
2e) 0.5kb end-labelled *cop/korB* probe (no protein) digested with 0.0005μg/ml final DNasel concentration.

Lane 2b shows that a 33b region on the *cop/korB* coding strand (overlapping the *cop/korB* promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the pattern generated by DNaseI cleavage of free *cop/korB* probe. There also exists a protein-induced DNaseI hypersensitive site situated in the centre of the 33b protected region, dividing it into two equal blocks of 16b each.
Lane 1b shows promoter P1, lane 2b the promoter P2, lane 3b the probe alone.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.

Lane 2b shows the overlap of the bound Conf/KorB protein with the probes in lane 3b.
**Figure 3.29 (cont'd overleaf)**

DNaseI protection analysis of pQR461-derived Cop/KorB protein extract incubated with 0.65kb *kilB* and 0.5kb *cop/korB* probes filled-in with T7 DNA polymerase.

20µl (80µg) of *E. coli* crude extract containing pQR461-generated protein products were incubated with 20,000cpm 0.65kb *kilB* and 0.5kb *cop/korB* filled-in probes respectively for 15min at room temperature, then subjected to DNaseI digestion for 1min before reactions were terminated (Section 2.25). Purified samples were electrophoresed on a denaturing 8% polyacrylamide gel using pQR412 (sequenced using the 17b reverse primer 5'dCAGGAAACAGCTATGAC) as a size marker (Section 2.19). Gels were dried down prior to autoradiography.

Lanes 1b) 0.65kb filled-in *kilB:*Cop/KorB nucleoprotein complex digested with 0.5µg/ml final DNaseI concentration.
1e) 0.65kb filled-in *kilB* probe (no protein) digested with 0.0005µg/ml final DNaseI concentration.

Lane 1b shows that a 60b region on the *kilB* coding strand (encompassing the *kilB* promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the band pattern generated by DNaseI cleavage of the *kilB* probe alone.

Lanes 2b) 0.5kb filled-in *cop/korB:*Cop/KorB nucleoprotein complex digested with 0.5µg/ml final DNaseI concentration.
2e) 0.5kb filled-in *cop/korB* probe (no protein) digested with 0.0005µg/ml final DNaseI concentration.

Lane 2b shows that a 33b region on the non-coding strand (overlapping the *cop/korB* promoter) has been protected from DNaseI digestion by the bound Cop/KorB protein compared to the pattern generated by DNaseI cleavage of free *cop/korB* probe. There also exists a weak protein-induced DNaseI hypersensitive site situated in the middle of the 33b protected region.
(Fig. 3.27b). This is a total of eight amino acids more than the pQR206-derived Cop/KorB fusion protein. Fig. 3.28 and 3.29 show the footprint pattern obtained when crude *E. coli* extracts containing pQR461-derived Cop/KorB fusion protein was incubated with the same labelled *cop/korB* and *kilB* DNA probes used for the pQR206-derived Cop/KorB protein. The resulting protection pattern obtained for the pQR461-generated Cop/KorB protein is identical to that obtained for the pQR206-generated Cop/KorB protein. Thus, it can be concluded that the extra amino acids at the N-terminal end of the Cop/KorB protein do not interfere with *kilB* or *cop/korB* binding, and that the generated footprint pattern is an accurate representation of the Cop/KorB operator sites on *kilB* and *cop/korB* respectively.

3.9 Summary.

A 0.5kb *SpeI*(53)-*BciI*(49) fragment containing the pIJ101 *cop/korB* ORF was cloned into pUC8 under the control of the *E. coli* lacZ promoter, creating pQR206. Sequencing of pQR206 confirmed that a LacZ-Cop/KorB fusion protein (with an additional 19 amino acids at its N-terminus end) would be produced from this construct. This Cop/KorB fusion protein would have a molecular weight of 10kDa, compared to 8.8kDa for the native 80 amino acid-long Cop/KorB protein.

*In vitro* coupled transcription-translation of pQR206 identified a protein product of approximately 10kDa which corresponds to the predicted molecular weight deduced from the *cop/korB* nucleotide sequence. Furthermore, an additional band corresponding to 10kDa was observed in crude protein extracts made from *E. coli* containing pQR206, indicating that the Cop/KorB fusion protein was stably produced *in vivo*.

Crude extracts containing Cop/KorB fusion protein were shown to bind to a 0.65kb *SalGI*(19)-*SstII*(16) *kilB* fragment, a 0.5kb *SpeI*(53)-*BciI*(49) *cop/korB* fragment and 0.5kb *Sau3A* *cop/korB* fragment in gel retardation assays, strongly suggesting that the fusion protein was functionally active and was specifically recognizing the *kilB* and *cop/korB* putative promoter regions. Binding of
Cop/KorB protein to kilB probe was more efficient than binding to cop/korB probe since the kilB probe was always completely bound and retarded by Cop/KorB protein in band-shift assays while the cop/korB probe was usually only partially retarded. Furthermore, results from competition assays suggested that the Cop/KorB protein may have a higher affinity for the kilB promoter than for the cop/korB promoter, since binding of Cop/KorB to cop/korB and to kilB was more rapidly competed out with the addition of cold cop/korB DNA than with the addition of cold kilB DNA respectively.

DNaseI protection analysis indicated that the cop/korB operator site for the Cop/KorB protein covers staggered 33b regions on the coding and non-coding strands respectively, and is superimposed on the cop/korB promoter. Furthermore, a hypersensitive site divides each of these 33b regions into two equal 16b regions. The kilB operator site is approximately twice the size of the cop/korB operator site (covering 60b and 52b on the coding and non-coding strands respectively), and also corresponds with the kilB promoter. Sequence comparison shows strong homology between the kilB and cop/korB protected regions as would be expected if the same protein would need to bind and regulate both promoters. However, the existence of a stronger region of dyad symmetry on the kilB promoter sequence compared to the cop/korB promoter sequence may cause the Cop/KorB protein to have a relatively higher affinity for the former.
CHAPTER FOUR
Cloning and expression of the sti determinant and the rep gene from the Streptomyces plasmid pIJ101.

The aims of this set of experiments were three-fold:

a) To clone the sti determinant from pIJ101 in order to characterize its function as a site for second-strand synthesis in replication,

b) To investigate the role of Cop/KorB in negative control of plasmid copy number via the sti locus defined above,

c) And finally, to determine the minimal replicon of pIJ101 and study expression of the gene(s) involved in replication.

4.1 Cloning of cop/korB and sti into pIJ702.

sti has been defined as a non-coding region of DNA which causes strong incompatibility when present in its natural orientation with respect to the basic replicon (Deng et al. 1988). A pair of plasmids can co-exist in the same host if they both possess sti in the correct orientation (Sti+), or both possess sti in the reverse orientation or both lack sti (Sti-). A Sti+ and Sti- plasmid cannot co-exist in the same cell: if they occur together the Sti+ plasmid is retained while the Sti- plasmid is eventually lost. Sti- plasmids accumulate more ssDNA than Sti+ plasmids, implicating sti as the site where second-strand synthesis is initiated (Deng et al. 1988). It is thought that Cop is a trans-acting negative regulator which affects the copy number of Sti+ plasmids (Deng et al. 1988).

pIJ702, a derivative of pIJ101 which lacks sti and cop/korB, is known to accumulate ssDNA which are the presumed intermediates in plasmid replication (Deng et al. 1988). There must be present on pIJ702 one or more sequences which can act as weak or inefficient sites for second-strand synthesis, thereby allowing double-stranded pIJ702 to form but also keeping a pool of ssDNA present in the
Such ssDNAs are likely to be highly recombinogenic and it may be this property which leads to the rapid breakdown of some pIJ702 derivatives (Pigac et al. 1988).

Thus, if sti is indeed the site for second-strand synthesis, the reinsertion of sti into pIJ702 and pIJ702-pUC8 shuttle vectors (which have been shown to be structurally unstable and to accumulate ssDNA in S. lividans; R. Barallon, personal communication), should result in a decrease of ssDNA for pIJ702 and lead to a stabilization of pIJ702-pUC8 shuttle vectors if the production of ssDNA is the cause of these latter plasmids structural instability. Furthermore, sti when present should only be functional in the correct orientation with respect to the basic replicon in order to be recognized as the site for second-strand synthesis. The Cop/KorB protein product was also tested for its ability to inhibit second-strand synthesis in vivo via its interaction with the sti locus, as suggested by Deng et al. (1988).

The cop/korB:sti region was isolated from pIJ101 (Fig. 1.2) on a 1.2kb BclI fragment and cloned into pUC8 cut with BamHI, producing pQR200 (Fig. 3.1). In order to insert both the cop/korB and sti determinants into pIJ702, pQR200 was cut with PstI and ligated to pIJ702 also cut with PstI. Recombinant plasmids containing pIJ702 in both orientations were selected in E. coli and named pQR410a and pQR411a respectively (Fig. 4.1). To investigate whether the pUC8 component was responsible for conferring instability upon the pIJ702-pUC8 shuttle vectors, the 2.7kb pUC8 sequence was excised from pQR410a and pQR411a by an EcoRI-HindIII digestion. The remainder of the plasmid was religated using a 58bp EcoRI-HindIII polylinker isolated from pUC19, producing pQR410b and pQR411b respectively (Fig. 4.1). pQR410a, pQR411a, pQR410b, and pQR411b all successfully transformed S. lividans to thiostrepton resistance (Section 2.8) and were structurally stable from restriction map analysis.

In order to subclone only the sti determinant, pQR200 was cut with EcoRI and SpeI and the 0.7kb fragment containing sti was cloned into pUC19 cut with EcoRI and XbaI, producing pQR417 (Fig. 3.17). To reinsert the sti region into pIJ702,
Figure 4.1

Construction of pQR410a/b, pQR411a/b, pQR421a/b, and pQR422a/b.
pQR417 was ligated to pIJ702 via the PstI site. Again, constructs in both orientations were found in *E. coli*, producing pQR421a and pQR422a respectively (Fig. 4.1). To remove the pUC19 component, pQR421a and pQR422a were digested with *EcoRI* and *HindIII*, and religated to the *EcoRI-HindIII* pUC19 polylinker, producing pQR421b and pQR422b respectively (Fig. 4.1). pQR421a, pQR422a, pQR421b and pQR422b all successfully transformed *S. lividans* to thiostrepton resistance (Section 2.8) and were structurally stable from restriction map analysis.

4.2 Analysis of ssDNA from plasmids containing *sti* and *cop/korB*.

Total DNA (Section 2.9) was made from all eight constructs described above (that is, pQR410a, pQR410b, pQR411a, pQR411b, pQR421a, pQR421b, pQR422a, and pQR422b), electrophoresed on a 0.7% (w/v) agarose gel (together with lysates containing pIJ702 and pIJ303 as positive and negative controls respectively), and blotted directly onto a nylon membrane without a prior denaturation step (Section 2.15). The filter was then probed with [α-32P]dCTP-labelled pIJ702 (Fig. 4.2). ssDNA was detected in lysates containing pQR410a/b and pQR422a/b but not in lysates containing pQR411a/b or pQR421a/b. Thus, if *sti* is reinserted into pIJ702 in the correct orientation with respect to the basic replicon, a decrease in ssDNA production occurs as seen in lysates containing pQR411a/b and pQR421a/b, regardless of the presence or absence of *cop/korB*. This suggests that *sti* alone is sufficient for the efficient initiation of second-strand synthesis to form double-stranded plasmid derivatives, since no ssDNA was detected in pQR421a/b lysates.

Cop/KorB protein may act to negatively control plasmid copy number in two ways, firstly by binding to the double-stranded form of *sti* and thereby inhibiting the production of ssDNA, or secondly by binding to the single-stranded form of *sti* and thereby inhibiting the conversion of ssDNA to double-stranded plasmid DNA (which would lead to a pool of ssDNA in the cell). The former situation has been shown not to occur as Cop/KorB does not bind to double-stranded *sti* DNA in gel-
Figure 4.2
Detection of ssDNA in *S. lividans*.

A) Agarose gel electrophoresis of *S. lividans* whole cell lysates (Section 2.9); the upper band is the chromosomal DNA and the lower, faster migrating band is the plasmid DNA.

B) Southern blot (without pretreatment) of the same gel (Section 2.15), probed with [α-<sup>32</sup>P]dCTP nick-translated pIJ702 showing ssDNA accumulation for lysates in lanes 2, 3, 4, 9, and 10.

retardation assays (Section 3.7.1). The latter case also probably does not occur since there is no accumulation of ssDNA in lysates containing pQR411a/b. Furthermore, from yields of plasmid DNA obtained for Sti+ plasmids with and without Cop/KorB, it would seem that Cop/KorB does not significantly alter plasmid copy number.

If sti is reinserted in the reverse orientation to the basic replicon (pQR410a/b, pQR422a/b), then ssDNA accumulates in the cell presumably because sti cannot be recognised as the site for second-strand synthesis in its inactive orientation. It should be noted that Deng et al. (1988) have claimed that the SpeI(53) site (used to subclone sti in pQR417) lies within the sti determinant - although this is clearly not the case according to the results obtained from the blots. That is, if cleavage at the SpeI(53) site does disrupt Sti function, ssDNA should be detected in lysates containing pQR421a/b as the disrupted sti determinant would not be recognized as the site for second-strand synthesis.

These results firstly confirm that sti is the site for second-strand synthesis and is active only in its natural orientation with respect to the basic replicon. Secondly, Cop/KorB does not seem to play a role in the control of plasmid copy number through the sti locus since it does not bind to either the double-stranded form of sti (thereby inhibiting ssDNA initiation) or the single-stranded form of sti (thereby inhibiting ssDNA conversion to the double-stranded plasmid form). Thirdly, sti in either orientation and in the absence of Cop/KorB seems to stabilize the pIJ702-pUC8 shuttle vectors since no rearrangements were detected in S. lividans.

4.3 Location of the sti function.

Several further constructs were made in order to define more precisely where the sti determinant was located on the 0.7kb SpeI(53)-BclI(57) fragment (Fig. 4.3). pQR417 (Fig. 3.17) was cut with SacI and HindIII, and the 0.6kb sti fragment ligated to pUC19 cut with the same enzymes (thereby removing approximately
Figure 4.3

Construction of pQR441a/b and pQR443a/b.

1) Digest with EcoRI + HindIII
2) Ligate to 58bp EcoRI-HindIII pUC19 polylinker
100bp from the original sti fragment), creating pQR437 (Fig. 3.17). pQR437 and pIJ702 were digested with PstI, ligated, and used to transform E. coli (Section 2.5). Constructs containing pIJ702 in both orientations were identified and named pQR441a and pQR442a, respectively. Only pQR441a was used in subsequent experiments as it possessed sti in the same orientation with respect to the basic replicon (Fig. 4.3). To remove the pUC19 component, pQR441a was digested with EcoRI-HindIII and ligated to the 58bp EcoRI-HindIII pUC19 polylinker, producing pQR441b (Fig. 4.3). pQR441a and pQR441b both successfully transformed S. lividans to thiostrepton resistance (Section 2.8) and were structurally stable from restriction map analysis.

pQR437 was digested with SacI and SacII (removing a further 65bp from the original sti fragment) and the 3.23kb fragment ligated to the 14bp SacI-SacII pBluescript polylinker, creating pQR438 (Fig. 3.17). pQR438 was ligated to pIJ702 via the PstI site. Again constructs containing pIJ702 in both orientations were found in E. coli and named pQR443a and pQR444a, respectively. Only pQR443a was studied since it possessed sti in the same orientation to the basic replicon (Fig. 4.3). To remove the pUC19 component, pQR443a was digested with EcoRI and HindIII and ligated to the EcoRI-HindIII pUC19 polylinker, producing pQR443b (Fig. 4.3). Both pQR443a and pQR443b successfully transformed S. lividans to thiostrepton resistance (Section 2.8) and were structurally stable from restriction map analysis.

Total DNA (Section 2.9) was made from pQR441a/b and pQR443a/b, run on a 0.7% (w/v) agarose gel (with pIJ702 and pIJ303 as controls), and blotted directly onto a nylon membrane (Section 2.15). The membrane was probed with [α-32P]dCTP-labelled pIJ702 (data not shown). No ssDNA was detected in lysates containing pQR441a/b or pQR443a/b, indicating that the deletions made to the original 0.7kb SpeI(53)-BclI(57) sti fragment were located outside the functional site for second-strand synthesis. Thus, this 0.53kb SpeI(53)-SsrII(55) fragment defines the upper limit of the sti determinant.
4.4 Cloning of the rep gene.

This section describes the strategy employed to clone and define the rep gene from pIJ101 into the E. coli vector pBGS19- (Table 2.3). pBGS19- was originally selected as a cloning vector in order to test if the kanamycin resistance gene located on the plasmid could be expressed in Streptomyces, and thus provide a single selection system which would function in both E. coli and S. lividans.

pIJ101 (Fig. 1.2) isolated from S. lividans TK24 was digested to completion with SstII. This digestion resulted in the production of four fragments on a 1% agarose gel of the following sizes: 2.2kb [SstII(63) to SstII(16)], 2.4kb [SstII(16) to SstII(33)], 3.5kb [SstII(33) to SstII(55)], and 0.5kb [SstII(55) to SstII(63)]. The 2.2kb SstII(16)-SstII(63) fragment containing the rep and orf56 genes was isolated (Section 2.13), treated with PolII to make the overhanging ends flush, and finally ligated to pBGS19- cut with Smal, creating pQR420 (6.6kb). In order to remove the 200bp orf56 gene from pQR420, this construct was cut with Ball and PstI. The larger 6.4kb fragment was isolated, the overhanging ends made flush with PolII and religated, creating pQR430. To remove the entire 630bp non-coding region from pQR430, this construct was digested with NotI and SphI, the overhanging ends of the larger 5.8kb fragment made flush with PolII and religated, creating pQR431.

pQR420 [2.2kb SstII(16)-SstII(63) rep insert], pQR430 [2.0kb Ball(15)-SstII(63) rep insert], and pQR431 [1.4kb NotI(12)-SstII(63) rep insert] were used to transform S. lividans protoplasts (Section 2.8). Unfortunately, the E. coli kanamycin resistance gene on pBGS19- was not expressed efficiently in S. lividans. Consequently, the thiostrepton resistance gene was isolated from pIJ702 on a 1.0kb BclI fragment and ligated into pUC19 cleaved with SacI (both fragments having first been treated with PolII to make the overhanging ends flush), creating pQR433. pQR433 was digested with EcoRI and KpnI and the 1.0kb thiostrepton gene ligated into pQR420, pQR430, and pQR431 cut with the same enzymes, creating pQR434, pQR435, and pQR436 respectively (Fig. 4.4).
1) Cut pLJ101 with SstII.
2) Blunt-end 2.2kb rep fragment.
3) Ligate into pBGS19- cut with Smal.

1) Cut with Ball and PstI to remove orf56 gene.
2) Blunt-end and religate.

1) Cut with NotI and SpI to remove non-coding region.
2) Blunt-end and religate.
A further construct was also made to remove only half of the non-coding region, which had been completely removed in the construction of pQR436 (Fig. 4.4). pQR434 (Fig. 4.4) was cut with KpnI and PstI, the 2.2kb fragment isolated and partially digested with ApaI. This partial digest was ligated to pBluescript cut with ApaI and KpnI and recombinants screened for inserts of 1.7kb in size, creating pQR445 (Fig. 4.5). pQR445 was cut with KpnI and HindIII, and ligated to pBGS19- (containing the tsr insert) cut with the same two enzymes, creating pQR446 (Fig. 4.5).

pQR434 [2.2kb SstII(16)-SstII(63) rep insert], pQR435 [2.0kb BstII(15)-SstII(63) rep insert], pQR436 [1.4kb NstI(12)-SstII(63) rep insert], and pQR446 [1.7kb ApaI(14)-SstII(63) rep insert] were all used to transform S. lividans protoplasts (Section 2.8). Several transformants were obtained for pQR434 and pQR435 but no transformants were obtained for pQR436 and pQR446. The viability of pQR434 in S. lividans confirms the report by Kieser et al. (1982) that the 2.2kb SstII(16)-SstII(63) region of pIJ101 defines the upper limit of the minimal replicon since it is capable of autonomous replication in S. lividans. Furthermore, insertional inactivation at the BamHI(1), XhoI(3), SalGI(7), and BclI(10) sites within this 2.2kb region destroys the ability of pIJ101 to survive in S. lividans (Kieser et al. 1982) because these sites lie within the rep ORF (Kendall and Cohen, 1988). The survival of pQR435 in S. lividans suggests that the orf56 gene product is not necessary for plasmid maintenance and replication, and its function is currently unknown. Since pQR446 was non-viable in S. lividans, this indicates that the non-coding 280bp DNA segment between the ApaI(14) and BstII(15) sites is critical for survival of the replicon.

pQR446 may be non-viable in S. lividans for two main reasons. Firstly, the promoter of the rep ORF may have been disrupted (in which case no Rep protein will be produced), or secondly the origin of replication may have been removed (in which case Rep protein will be produced but it is unable to initiate replication at ori if the latter is missing). In order to determine which of these possibilities was the
1) Cut pQR434 with KpnI and PstI.
2) Isolate 2.2kb rep fragment and partially digest with ApaI.
3) Ligate partial digest into pBluescript cut with ApaI and KpnI.

1) Cut pQR445 with KpnI and HindIII.
2) Ligate 1.7kb rep fragment into pBGS19-(+tsr) cut with KpnI and HindIII.

127
case, an attempt was made to map the S1 transcription start site of the rep gene (Section 2.18). However, total mRNA made from S. lividans cultures harbouring pIJ702 grown for 6hrs, 12hrs, 24hrs, and 48hrs (Section 2.16) did not generate sufficient rep mRNA for detection in Northern blots or S1 mapping experiments. This may be because the rep mRNA is present at a relatively low copy number and/or expressed only at specific times during cell growth, thus making it difficult to detect.

It should be noted that although transformants were obtained for pQR434 and pQR435, these plasmids were highly unstable in S. lividans according to results obtained from restriction map analysis. In general, HindIII-KpnI double digests of pQR434 and pQR435 (Fig. 4.4) resulted in the production of the 2.2kb or 2.0kb rep fragment respectively. However, the HindIII-KpnI double digests usually generated a band smaller than the expected 5.4kb (pBGS19- plus tsr) band. This indicated that deletions were taking place in the pBGS19- part of the plasmid (as deletions in the tsr gene would result in loss of thiostrepton resistance). Thus, the essential rep region of the plasmid was maintained as any rearrangements in this section of the plasmid would make it non-viable in S. lividans, while the non-essential pBGS19-component was being successively deleted from the construct. It is also interesting to note that plasmid DNA isolated from each of the transformants analysed gave a unique restriction enzyme pattern which may indicate that the deletion events may not be site-specific.

The ability of the pQR434 and pQR435 rearranged plasmids to survive in E. coli may be affected by the deletions occurring in the pBGS19- part of the construct. To investigate this, a sample from the same S. lividans plasmid preparations used for restriction mapping was also used for re-transforming E. coli JM107 (Section 2.5). Apart from one pQR435 sample, all the pQR434 and pQR435 samples isolated from S. lividans were able to successfully re-transform E. coli JM107 to kanamycin resistance (although at a very low frequency). Surprisingly, when plasmid DNA from these colonies were restriction mapped, the restriction enzyme pattern
generated was identical to the control DNA (that is, the original pQR434 and pQR435 plasmids constructed in *E. coli* and used to transform *S. lividans* protoplasts). Thus, there must have been residual levels of non-rearranged plasmids in *S. lividans* (co-existing with the dominant rearranged variants), which were not detected in agarose gels but which were subsequently amplified in *E. coli*.

The presence of residual amounts of intact pQR434 and pQR435 in *S. lividans* suggests that these plasmids were able to replicate in the transformants and that the deletion events were occurring after at least one round of replication. Lee *et al.* (1986) have reported that *S. lividans* colonies were found to contain mixed groups of rearranged plasmids after an initial transformation with a pIJ702-pUC12 shuttle vector, and have suggested that the deletion events were occurring during the growth of the mycelium. Pigac *et al.* (1988) have also reported that *S. lividans* transformants contained a mixture of both intact pBR322-pIJ350 shuttle vectors plus various deleted forms; the pattern of deleted plasmids altering with the age of the transformant colony chosen for analysis.

As stated, Pigac *et al.* (1988) have suggested that plasmid structural instability in *S. lividans* may be due to the presence of ssDNA which are potentially highly recombinogenic. When total DNA was made from *S. lividans* harbouring the pQR435-rearranged plasmids (Section 2.9) and probed with \([\alpha^32P]dCTP\)-labelled 2.0kb *rep* DNA (Section 2.15), ssDNA was detected for all the lysates analysed (data not shown). Thus, as pQR435 was structurally unstable in *S. lividans* and accumulated ssDNA, *sti* was reinserted into pQR435 in an attempt to prevent the plasmid from rearranging (*sti* was shown in Sections 4.1 and 4.2 to stabilize the pIJ702-pUC8 shuttle vectors in *S. lividans*).

pQR438 [0.53kb SpeI(53)-SstII(55) *sti* fragment in pUC19; Fig. 3.17] was ligated to pBGS19- at the HindIII site, creating pQR453 (Fig. 4.6). pQR453 was then cut with *EcoRI*, the 0.53kb *sti* fragment isolated and ligated to pQR435 cut with *EcoRI*, creating pQR455 (Fig. 4.6). pQR455 was used to transform *S. lividans* TK24 protoplasts (Section 2.8). However, when plasmid DNA was isolated from *S.
Figure 4.6 Construction of pQR453 and pQR455

1) Cut pQR438 with HindIII.
2) Ligate into pBGS19- cut with HindIII.

1) Cut pQR453 with EcoRI.
2) Isolate 0.5kb sti fragment.
3) Ligate 0.5kb sti fragment into pQR435 cut with EcoRI.

1) Cut pQR453 with EcoRI.
2) Isolate 0.5kb sti fragment.
3) Ligate 0.5kb sti fragment into pQR435 cut with EcoRI.

pQR453 7.6kb

Km<sup>r</sup>

pQR455 7.93kb

Km<sup>r</sup>
lividans transformants and restriction mapped, it was found that in all cases the plasmid had structurally rearranged and the 0.53kb sti fragment was missing from the construct. Furthermore, when total DNA made from these transformants was probed with [α-32P]dCTP-labelled 2.0kb rep DNA, ssDNA was detected in all the lysates analysed (data not shown). Thus, the random reinsertion of sti into Sti-plasmids may not always prevent them from rearranging, and other factors such as the location of sti with respect to the rep gene and the type of plasmid used in the construct may be important.

4.5 Expression of Rep protein in vitro and in vivo in E. coli.

pQR431 [1.4kb NotI(12)-SstII(63) rep gene in pBGS19-] was sequenced using reverse primer (Section 2.19) to analyse the rep(NotI)/pBGS19(SphI) junction in order to determine how the Notl 5' overhanging end and Sphl 3' overhanging end had been made flush by the PolI before ligation. This, in turn, would indicate if the rep gene was in the correct reading frame with respect to the lacZ gene in pBGS19-.

Fig. 4.7 shows the DNA sequence across the Notl/Sphl junction of two pQR431 constructs. In pQR431a, the Sphl 3' overhanging end has been completely removed by the PolI while the Notl 5' overhanging end has been completely filled in with dNTPs. This combination of bases should result in the production of an in-frame LacZ-Rep fusion protein, with an additional 12 amino acids attached to the N-terminus of the Rep protein. In pQR431b, although the Sphl 3' overhanging end has been completely removed, the terminal "G" of the Notl 5' overhanging end has been removed before the remaining bases have been filled in by the PolI. Thus, an artificial in-frame stop codon has been generated 12 amino acids after the initial LacZ methionine. However, since the AUG start codon of the rep gene is located 42bp from the lacZ AGGA ribosomal binding site, the E. coli ribosomes are unable to recognize the rep start codon since it is located too far from the ribosomal
Figure 4.7  Sequence of pQR431 across the junction of the lacZ promoter and rep gene

pQR431a

-35lacZ
CCCAGGCTTTACACTTTATGCTTCCGGCTCGFTGTGTGTGGAAATTGTGAGCGGATAACAA

S/DlacZ
TTTCACACGGGAAACAGCTATGACCATTACCGGCGCCGCGGTGATG...
MetThrMetIleThrProSerLeuGlyArgAlaLeuMetrep

pQR431b

-35lacZ
CCCAGGCTTTACACTTTATGCTTCCGGCTCGFTGTGTGTGGAAATTGTGAGCGGATAACAA

S/DlacZ
TTTCACACGGGAAACAGCTATGACCATTACCGGCGCCGCGGTGATG...
MetThrMetIleThrProSerLeuAlaAlaProEnd
In vitro coupled transcription-translation of pQR431a/b resolved on a 10% SDS-polyacrylamide gel.

CTT reactions were done as described in Section 2.21. After electrophoresis (Section 2.20) of the $^{35}$S-methionine labelled protein products, the gel was dried down and exposed to Fuji RX X-ray film.

Lanes  
1) pQR431a: the arrow indicates the production of a 50kDa band from pQR431a corresponding to the Rep protein. 
2) pQR431b: no Rep protein produced as translation is prematurely terminated in this construct.
3) pBGS19-: control plasmid.
Figure 4.9

_E. coli_ total protein extracts containing pQR431a/b and pBGS19- resolved on a 10% SDS-polyacrylamide gel.

Induced overnight cultures of _E. coli_ JM107 containing pQR431a, pQR431b, and pBGS19- (as a control) respectively were sonicated in extraction buffer (Section 2.22). Approximately 40ug of each total protein sample prepared above was analysed by electrophoresis on a 10% SDS-acrylamide gel, after which the gel was stained with Coomassie blue and destained (Section 2.20).

Lanes 1, 2, and 3, show protein extracts from _E. coli_ harbouring pQR431a, pQR431b and pBGS19- respectively. No new band corresponding to the 50kDa Rep protein can be seen in lane 1.
binding site. Thus, theoretically, no Rep protein should be produced from pQR431b.

pQR431a, pQR431b, and pBGS19- (as a negative control) were used to initiate transcription/translation in the *E. coli* *in vitro* CTT system (Section 2.21), and the $^{35}$S-methionine labelled protein products visualized on a 10% SDS-polyacrylamide gel (Fig. 4.8). A new protein product of approximately 50kDa was produced for pQR431a when compared to protein products generated from pQR431b and pBGS19- respectively. This 50kDa protein corresponds to the expected size of the *rep* gene product from sequence analysis of pQR431a.

In order to detect Rep protein production *in vivo* (Section 2.22), cultures of *E. coli* JM107 containing pQR431a, pQR431b, and pBGS19- were grown overnight in the presence of IPTG and ampicillin, and sonicated samples were run on a 10% SDS-acrylamide gel (Fig. 4.9). However, no new protein product corresponding to 50kDa in size was observed for the pQR431a sample when compared to the pQR431b and pBGS19- samples. From *in vitro* analysis (Fig. 4.8), it is known that pQR431a is capable of producing Rep protein and that there is no barrier at the transcriptional or translational level. Thus, Rep protein produced from pQR431a may be unstable in *E. coli* since it is a relatively large protein and may be prone to proteolytic degradation *in vivo*. Alternatively, production of Rep protein *in vivo* from the pBGS19- *lacZ* promoter may not occur at high enough levels to allow for easy detection by SDS-PAGE. One solution to this may be to subclone the *rep* gene into an expression vector possessing the tac, λpL, or T7p10 promoters which would be more efficient than the *lacZ* promoter.

### 4.6 Summary.

The 1.2kb BclI(49)-BclI(57) *sti:cop/korB* and 0.7kb SpeI(53)-BclI(57) *sti* regions were isolated from the *Streptomyces* plasmid pIJ101 and cloned into pIJ702 (a derivative of pIJ101 which lacks *sti:cop/korB* and accumulates ssDNA) at the
PstI site in both orientations. No ssDNA was detected in constructs containing sti present in its correct orientation with respect to the basic replicon, with or without cop/korB. Constructs which contained sti in the reverse orientation were found to accumulate ssDNA. Thus, sti is only active as the site for second-strand synthesis in its natural orientation. Furthermore, sti inserted into the structurally unstable pIJ702-pUC8 shuttle vectors prevented them from rearranging in *S. lividans*. The sti determinant was defined to a 0.53kb StII(55)-SpeI(53) fragment.

The minimal replicon of pIJ101 was defined to a 2.0kb BglII(15)-StII(63) fragment containing the rep ORF and a non-coding region of DNA (pQR435). This non-coding region of DNA is crucial for plasmid maintenance and replication in *S. lividans*, and may either contain the origin of replication or the rep promoter (or both). pQR435 was unstable in *S. lividans* and accumulated ssDNA. The insertion of the 0.53kb StII-SpeI sti fragment into pQR435 did not stabilize this plasmid or prevent it from accumulating ssDNA.

The rep gene isolated on a 1.4kb NcoI(12)-SacII(63) fragment was treated with PolII and placed under the control of the pBGS19- lacZ promoter, creating pQR431. Sequencing of the rep/pBGS19- junction resulted in the identification of two types of plasmid constructs which differed by one base, leading to the production of a LacZ-Rep fusion protein in pQR431a and premature termination of the rep gene in pQR431b. *In vitro* CTT of pQR431a resulted in the production of a new 50kDa protein band which corresponds to the expected size of the Rep protein from sequence analysis. However, no new protein band was detected in crude extracts made from induced *E. coli* cultures containing pQR431a.
CHAPTER FIVE. General discussion.

5.1 Prologue.

The main aim of the project was to investigate the biological function of the cop/korB gene from the S. lividans plasmid pIJ101. This protein is of interest to study because it has been assigned two separate functions. Firstly, as a trans-acting negative regulator of copy-number via sti (the site for second-strand synthesis), since it has been reported that the copy number of a Sti+Cop⁻ plasmid decreases from about 1000 to normal levels (that is, 300/chromosome) when a pIJ101 Sti+Cop+ derivative is also present (Deng et al. 1988). Secondly, as a repressor of kilB, since the kilB gene cannot be introduced into S. lividans without the presence of the korB locus in cis or trans (Kendall and Cohen 1987). The kilB gene may be involved with plasmid spread during pock formation, since insertions into the kilB gene lead to the production of small pocks (Kendall and Cohen 1987).

Upon initiation of this work, there were no published reports on the specific cloning, sequencing, or overexpression of the pIJ101 cop/korB gene product, and indeed the existence of the cop/korB gene product had only been inferred from the indirect genetic evidence outlined above. To understand the regulatory interactions of Cop/KorB on sti, kilB, and cop/korB, it is necessary to determine the basic properties of the Cop/KorB protein. Thus, the approach taken to study Cop/KorB function was to subclone the cop/korB gene into an E. coli vector and express it in sufficient quantities to do various DNA binding studies in vitro.

5.2 Cloning and expression of the cop/korB gene product.

The pIJ101 cop/korB gene was initially isolated on a 1.2kb BclII(49)-BclII(57) fragment (also containing the sti determinant) and cloned into the E. coli vector pUC8, creating pQR200 (Section 3.1, Fig. 3.1). However, when pQR200 was placed in an E. coli in vitro coupled transcription-translation (CTT) system, no new protein products were detected (Section 3.4, Fig. 3.4). One reason for this may
be because the cop/korB promoter was not recognized by the E. coli RNA polymerase, even though the cop/korB promoter does have some sequence homology to the E. coli consensus-type promoters (Hawley and McClure 1983, Stein et al. 1989). This barrier to expression of Streptomyces genes in E. coli has been previously reported (Bibb and Cohen 1982, Bibb et al. 1985a), and has been attributed to the presence of different classes of transcriptional initiation signals in Streptomyces. These different promoters are transcribed by RNA polymerases which have varied recognition specificities depending on the type of sigma subunit associated with them (Westpheling et al. 1985, Buttner 1989, Buttner et al. 1990).

Assuming that transcription was taking place from the lacZ promoter, another plausible explanation for the lack of cop/korB expression from pQR200 may be that the inverted repeat sequences within the sti region (located between the lacZ promoter and the cop/korB gene) may act as transcriptional terminator(s) due to their secondary structure (Section 5.6). If stem-loop structures were to form in vivo, then RNA polymerase might terminate transcription at one or more of these structures before any cop/korB mRNA could be synthesized.

Even if full-length message was generated from either the lacZ promoter or the cop/korB promoter, the potential E. coli-like AGGA ribosomal binding site (located 11bp upstream from the cop/korB AUG translational start site) might not be efficiently recognized by the E. coli ribosomes if the secondary structure of the mRNA affects the access of ribosomes to the AGGA and AUG sequences. Alternatively, if this AGGA site is recognized, then the ribosomes may not identify the cop/korB translational start site due to its location at one extreme end of the optimal alignment between a ribosomal binding site and the translation start site (typically separated by 3 to 11bp with an optimum spacing of 7 to 9 nucleotides) (Shine and Dalgarno 1974, Stormo et al. 1982). One way to determine which of these possibilities is actually the case would have been to probe E. coli total mRNA with the cop/korB gene in order to assess the size of the message being produced from pQR200. This in turn would have indicated if full-length transcription was
taking place from the \textit{lacZ} or the \textit{cop/korB} promoter (in which case the barrier to \textit{cop/korB} expression is at the translational level), or if transcription was terminating at \textit{sti} or one of the other inverted repeat structures (in which case the barrier to \textit{cop/korB} expression is at the transcriptional level).

The \textit{cop/korB} was then placed directly under the control of the \textit{lacZ} promoter by deleting the intervening \textit{sti} sequence, creating pQR206 (Section 3.2, Fig 3.1). Sequencing of pQR206 indicated that a 10kDa LacZ-Cop/KorB fusion protein with an additional 19 amino acids incorporated at its N-terminal end should be produced from this construct (compared to the 8.8kDa native Cop/KorB protein) (Section 3.3, Fig. 3.2). \textit{In vitro} CTT of pQR206 identified a new protein band of 10kDa; digestion of pQR206 with \textit{Smal} (which cleaves within the \textit{cop/korB} ORF) resulted in the disappearance of this 10kDa band (Section 3.4, Fig. 3.4). Thus, this was the first direct evidence that the \textit{cop/korB} ORF does encode a polypeptide.

Analysis of crude protein extracts from induced cultures of \textit{E. coli} harbouring pQR206 confirmed that the Cop/KorB fusion protein was stably produced \textit{in vivo} and accounted for approximately 2\% of total intracellular protein (Section 3.5, Fig. 3.7 and 3.8). Since initial attempts to purify the Cop/KorB fusion protein were not successful, \textit{E. coli} crude extracts containing Cop/KorB protein were used to test for binding activity on the \textit{kilB} and \textit{cop/korB} genes and the \textit{sti} determinant.

\textbf{5.3 Cop/KorB binding on \textit{cop/korB} and \textit{kilB} genes.}

The Cop/KorB fusion protein was found to be functionally active and was able to specifically recognize and bind to a 0.65kb \textit{SacII}(16)-\textit{SalGI}(19) fragment containing the \textit{kilB} gene in a gel retardation assay (Section 3.7.1, Fig. 3.13). The Cop/KorB operator site for \textit{kilB} was determined, from deletion analysis, to lie somewhere along a 187bp region delineated by the \textit{SalGI}(19) and \textit{BsrEII}(17) sites which also contains the \textit{kilB} promoter (Section 3.7.3, Fig. 3.22). Stein et al. (1989) have also identified a similar region needed for KorB regulation \textit{in vivo}, although they report that a smaller 76bp fragment extending from the \textit{BsrEII} site to the first T
of the putative -35 region (incorporating the FspI site) was sufficient to confer promoter activity in vivo.

The Cop/KorB fusion protein also recognized and bound to a 0.5kb SpeI(53)-BclI(49) fragment containing the cop/korB gene and a 0.5kb Sau3A fragment containing the cop/korB promoter in gel retardation assays (Section 3.7.1, Fig. 3.13 and 3.15). The Cop/KorB operator site for cop/korB was located by deletion analysis to a 61bp region between the SpeI(53) and XhoII sites containing the cop/korB promoter (Section 3.7.3, Fig. 3.22). Stein et al. (1989) have also defined this region to be necessary for KorB regulation in vivo.

Binding of Cop/KorB protein to the cop/korB fragments was not as efficient as binding to the kilB fragment, since on most occasions there were residual amounts of unbound cop/korB probe remaining in the assay (Section 3.7.1, Fig. 3.16). This suggests that the Cop/KorB protein may have a greater affinity for the kilB promoter than for the cop/korB promoter. This was also demonstrated in the competition assays (Section 3.7.2, Fig. 3.19), where Cop/KorB binding to labelled cop/korB and kilB probes was competed-out more effectively with the addition of cold cop/korB DNA than with the addition of cold kilB DNA respectively. These differences in affinities may be important in the control of Cop/KorB and KilB protein concentrations in the cell (see Section 5.4). It should be noted that Stein et al. (1989) have reported that KorB in vivo was more efficient in repressing β-galactosidase expression controlled by the korB promoter than by the kilB promoter. However, this may be an inaccurate conclusion since Stein et al. (1989) were using a crude and subjective method of assessing expression (that is, by visually scoring colonies after overlaying with soft agar containing X-gal or MUG) since quantitation of β-galactosidase production from cells grown in liquid cultures was not reproducible.

Results from DNaseI protection analysis (Section 3.8) revealed that the Cop/KorB protein binds to complementary 33b regions on the cop/korB coding
(Fig. 3.24) and non-coding (Fig. 3.25) strands, which coincide with the location of the putative \textit{cop/korB} promoter. A DNaseI hypersensitive site divides these 33b regions into two equal 16b-blocks, and suggests that the Cop/KorB protein causes some torsional stress when binding to the \textit{cop/korB} operator thereby exposing the DNA to DNaseI activity at this site. The Cop/KorB protein binds to a relatively large operator site on the \textit{kilB} gene which spans 52b on the non-coding (Fig. 3.24) strand and 60b on the coding (Fig. 3.25) strand, coinciding with the location of the putative \textit{kilB} promoter. Furthermore, the Cop/KorB operator site on \textit{kilB} substantially overlaps the defined RNA polymerae binding site for this gene (Buttner and Brown 1985). The \textit{S. coelicolor} RNA polymerase covers a 61b region on the \textit{kilB} coding strand from nucleotides (nt) 2444 to 2504 and a 62b region on the non-coding strand from nt 2444 to 2506 (Buttner and Brown 1985) compared to nt 2456-2515 and nt 2462-2513 respectively for Cop/KorB (Appendix 1). Thus, Cop/KorB-mediated repression of the \textit{cop/korB} and \textit{kilB} promoters most probably occurs at the transcriptional level. By binding to the \textit{kilB} and \textit{cop/korB} genes at operator sites that overlap the -10 and -35 of the two promoters, Cop/KorB protein would physically prevent RNA polymerase from recognizing and gaining access to \textit{pkilB} and \textit{pcop/korB}. This principle of mutual exclusion underlies all cases of negative control where the repressor protein and RNA polymerase compete for promoter binding.

The defined \textit{kilB} and \textit{cop/korB} operator sites share considerable sequence similarity (as would be expected if both operator sites are being regulated by the same protein), including an inverted repeat structure spanning the -10 and -35 regions of both promoters (Fig. 3.26). Inverted repeats located within the promoters of genes are often indicative of operator sites for repressor proteins. The inverted repeat of the \textit{kilB} operator is stronger than that of the \textit{cop/korB} operator, since the former has a higher degree of matched intra-strand base-pairs (10/12 compared to 8/12, respectively). Furthermore, there is an additional inverted repeat present upstream of the \textit{kilB} promoter which is partially protected by Cop/KorB binding.
The combination of the larger kilB operator site for Cop/KorB binding and the stronger inverted repeat lying within this operator site may be further evidence, together with the gel retardation results, that the kilB promoter is more efficient at binding Cop/KorB than the coplkorB promoter. The relatively stronger affinity of Cop/KorB protein for the kilB promoter compared to its own promoter seems logical, otherwise the lethal affects of the kilB gene would be exerted (see Section 5.4).

It is interesting to note that the kilB protected region is almost twice the size of the coplkorB protected region (Fig. 3.26), and this raises the possibility that double the number of active Cop/KorB repressor units may be binding to pkilB compared to pcoplkorB. However, if this were true, one would expect to see two pairs of inverted repeats with similar sequence lying side-by-side. This would be analogous to the λcl right operator site which consists of three pairs of virtually identical inverted repeats (OR1, OR2, and OR3 overlapping the pRM and pR promoters) that can be bound by three units of λ CI repressor protein (Ptashne 1986). Although there is a second partially-protected inverted repeat located upstream of the main kilB protected inverted repeat, there is no sequence homology between them.

A more likely explanation for the larger protected region in kilB compared to coplkorB may be due to a slightly different physical interaction of Cop/KorB protein with these DNA sequences. Thus, although there may be the same number of Cop/KorB protein units binding to both kilB and coplkorB operator sites, the interaction between Cop/KorB and kilB may be relatively more extensive leading to a larger protected area for this promoter compared to the coplkorB promoter. One reason for this may be because of the existence of a weaker secondary transcription start site located upstream of the main kilB transcription start site (Stein et al. 1989); Cop/KorB may block transcription from the upstream promoter by physically covering this weaker transcription start site plus the main kilB promoter and transcription start site (Fig. 3.26). Cop/KorB would regulate expression from the upstream promoter in this manner (that is, by extensive binding to the
downstream promoter to block transcription from both the downstream and upstream promoters), as Cop/KorB is unable to bind directly to the -10 and -35 regions of the upstream promoter (Stein et al. 1989). This is based on a report by Stein et al. (1989) which states that insertions of ClaI linkers into the FspI site of the smaller 76bp BstEII(17)-FspI(18) fragment resulted in complete loss of promoter activity whereas linker-insertion into the corresponding site in the larger 187bp BstEII(17)-SalGI(19) fragment resulted in unregulated weak transcriptional activity.

The kilB and cop/korB operator sequences (averaging 56bp and 33bp respectively) are relatively long compared to the defined 17bp cro/cl operator sites of λ (Ptashne 1986). This is surprising since the Cop/KorB protein (80 amino acids) is similar in size to Cro (66 amino acids) and the amino-terminus of CI repressor (92 amino acids). Both Cro and CI proteins bind to individual operator sites on λ as dimers (Ptashne 1986), and Cop/KorB may also dimerise when binding to the kilB and cop/korB operators.

The data from the defined physical mapping of the kilB operator indicates that the sequence immediately after the last T of the -35 region is important in Cop/KorB binding (Fig. 3.26). This contradicts the indirect genetic data obtained by Stein et al. (1989) for the smaller 76bp BstEII(17)-FspI(18) fragment containing the kilB promoter which they claim is capable of being regulated by KorB in vivo. This discrepancy may reflect the inability of DNaseI to cleave bases located immediately adjacent to the protected DNA sequence because of size exclusion. However, it should be noted that there is at least a 12bp difference between the last T of the FspI site and the upper limit of the kilB operator as defined by footprinting analysis; this difference may be greater than the intrinsic error caused by using DNaseI.

There are several alternative methods of DNA footprinting (such as the location of contact points by dimethylsulphate resistance and the generation of hydroxyl radicals) which would give additional information about DNA-protein interactions,
and may clarify whether the "extra" base-pairs of the DNaseI-determined kilB operator sequence are specifically involved in Cop/KorB binding. Furthermore, to delineate the operator sequences of both promoters more exactly, one could carry out quantitative binding experiments and in vivo studies using operator sites with in vitro generated mutations in the putative binding sequence.

It should be noted that from the results of the DNA footprinting experiments, pQR206 contains a fully functional operator upstream of the cop/korB gene to which Cop/KorB protein can actively bind and negatively regulate cop/korB expression. Since pQR206-generated Cop/KorB protein was detected in E. coli crude extracts and accounted for at least 2% of total protein even with the presence of a functional operator on the plasmid, this indicates that the affinity of Cop/KorB for its own promoter may not be very high and thus allows for some cop/korB expression in vivo. However, it should be noted that the Cop/KorB protein is being unnaturally over-expressed from a strong lacZ promoter and that this may influence its ability to tightly autoregulate itself. Thus, the conditions in S. lividans may differ from E. coli in that Cop/KorB would be expressed from its own promoter and in probably relatively low amounts (as is usually the case with repressor proteins), resulting in a more rigorous control of pcop/korB transcription by Cop/KorB in S. lividans. Ideally, another plasmid construct should be made in which the cop/korB operator sequence is entirely removed (to allow for maximum over-expression) and in which native Cop/KorB can be expressed in large amounts. These two factors would be important if the Cop/KorB protein is to be purified for crystallographic studies.

5.4 Model for KorB regulation of korB and kilB genes.

Based on the data obtained from the mobility shift and footprinting experiments discussed above, a model for KorB binding and control of kilB and korB expression can be proposed (Fig. 5.1). This model assumes that the kilB promoter is more efficient than the korB promoter, and that the affinity of KorB protein for
### Model for KorB regulation of korB and kilB genes

#### A)

<table>
<thead>
<tr>
<th>Copy Number</th>
<th>korB</th>
<th>kilB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>expressed from weak promoter; no RBS (low conc.)</td>
<td>expressed from strong promoter and RBS; plasmid spread within mycelia (pock formation)</td>
</tr>
<tr>
<td>Medium</td>
<td>expressed (high conc.)</td>
<td>repressed</td>
</tr>
<tr>
<td>High (normal)</td>
<td>repressed</td>
<td>repressed</td>
</tr>
</tbody>
</table>

#### B)

- KilB
- KorB

<table>
<thead>
<tr>
<th>Time</th>
<th>KilB protein concentration</th>
<th>KorB protein concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the \textit{kilB} operator is stronger than for the \textit{korB} operator. Scenarios I, II, and III explained below are diagrammatically represented in Fig5.1(A+B), whereas scenario IV is only represented in Fig. 5.1B.

I) Upon transfer of pJU101 to a new host, the plasmid is initially present only at low copy numbers. At first, only limited amounts of KorB protein are synthesized in the cell. This is due to the less efficient transcription of \textit{korB} message from the relatively weaker \textit{korB} promoter and more importantly, to the inefficient translation of this \textit{korB} message as it lacks a leader sequence with a conventional ribosome binding site. The paucity of KorB protein in the cell results in the derepression of the \textit{kilB} gene, leading to a burst in KilB protein production. The latter is achieved by the efficient transcription of \textit{kilB} message from the strong \textit{kilB} promoter and the efficient translation from the strong ribosomal binding site present on the \textit{kilB} message. Thus, although the \textit{kilB} promoter may be more efficient than the \textit{korB} promoter, it is mainly the differential translation of the messages initiated at these two promoters which leads to the rapid synthesis of KilB protein. KilB may act to retard cell growth until an appropriate copy-number is reached in the cell before allowing the plasmid to spread throughout the recipient mycelium (which results in the formation of normal-sized pocks).

II) However, as plasmid copy-number increases to the desired level, the concentration of KorB protein in the cell also increases due to the combined gene dosage effect and subsequent transcription/translation of \textit{korB} message. There is now sufficient KorB present to actively repress \textit{kilB} transcription by binding to the \textit{kilB} operator. In this system, the KorB protein must have a higher affinity for the \textit{kilB} promoter than for its own promoter. Thus, given a choice between these two promoters, KorB protein will preferentially bind to the \textit{kilB} promoter and specifically inhibit \textit{kilB} transcription.

III) However, once KorB repressor concentration increases beyond a certain limit and all the stronger \textit{kilB} operator sites have been occupied, then KorB is able
to bind to its own weaker operator and negatively regulate production of excess KorB protein.

IV) Since KorB has a relatively low affinity for its own promoter, this allows for some read-through expression of korB to occur which ensures that there is always a basal level of KorB present in the cell. However, if KorB protein concentration drops below this basal level (for example, during cell division), then KorB dissociates from its own operator since it has a higher affinity for the kilB operator. This frees the korB promoter and allows it to resume functioning to provide more KorB repressor. Thus, a virtually constant level of KorB protein is maintained in the cell which in turn controls KilB protein concentration under a critical level. The model assumes that at any given instant, the concentration of KorB protein is high enough for both kilB and korB operator sites to be occupied.

The translational control of korB proposed in the model is similar to that involved in the control of CI repressor protein from bacteriophage lambda. Transcription of cl initiated at pRE produces more repressor than does transcription from pRM, since message synthesized from pRE has a strong ribosome binding site and is translated more efficiently than message synthesized from pRM which has no ribosomal binding site (Ptashne et al. 1976; Waltz et al. 1976). Translational control is probably important in several other genes that lack conventional ribosome binding sites, such as the transposon Tn1721-encoded tetracycline repressor in E. coli (Klock and Hillen 1986), and the aminoglycoside phosphotransferase (aph) and erythromycin-resistance (ermE) genes from Streptomyces (Bibb et al. 1985b). Thus, the sequence located 3' to the translational start codon may be important in the initiation of protein synthesis in these systems (Scherer et al. 1980, Gold et al. 1981).

The korB and kilB genes may form part of a regulatory network involved in the transfer and maintenance of pIJ101 in Streptomyces, since pIJ101 has no obvious selective advantage to its host (apart from fertility). This is similar to the IncP plasmid RK2 (56.4kb), where the kil-kor determinants form a complex regulatory
network involved in the control of plasmid maintenance or host range (for a review, see Kues and Stahl 1989), since either altering or deleting components of the regulatory network leads to the production of unstable plasmid derivatives with a decreased host range (Schmidthauser and Helinski 1985; Thomas et al. 1983).

RK2 encodes for several *kil* determinants (*kilA*, *kilB*, *kilC*, *kilD*, and *kilE*) which are potentially lethal to *E. coli* unless negatively regulated at the transcriptional level by various combinations of *kor* functions (*korA*, *korB*, *korC*, *korE*, and *korF*) (Fig. 5.2). Expression of both the *trfA* operon (encoding for the replication initiator protein and *kilD* determinant) and the *kilA* operon is negatively regulated by KorA in conjunction with KorB (or KorE) (Shingler and Thomas 1984; Schreiner et al. 1985; Young et al. 1985, 1987; Goncharoff et al. 1991). Interestingly, *kilD* is able to counteract the *korA-korB* repression of the *trfA* operon, although the mechanism for this is unknown (Schreiner et al. 1985; Theophilus et al. 1985). The *kilC* promoter and the dual *kilE* promoters are repressed by KorC; KorC is in turn positively regulated by KorA (probably by antitermination of the *korC* transcript) (Young et al. 1984; Thomas et al. 1988; Kornacki et al. 1990). *kfrA* (whose function is unknown) is regulated by KorA and KorF (Thomas et al. 1990a; Jagura-Burdzy et al. 1991). *kilB* is regulated by only one function, KorB (Bechhofer et al. 1986). *korA*, *korB*, and *korF* (which form part of the *trfB* operon) are autoregulated by KorA and KorB at the *trfB* promoter (Smith et al. 1984; Theophilus et al. 1985).

The *kil-kor* regulon of RK2 is directly involved in the control of plasmid replication, since the replication (*trfA*) gene is negatively repressed at the transcriptional level by the *korA*, *korB*, and *korE* gene products (Shingler and Thomas 1984; Schreiner et al. 1985). However, because the pIJ101 *kilA* and *kilB* determinants are involved in conjugal transfer, the regulation of these loci is directly linked with plasmid transfer rather than plasmid replication. Thus, the pIJ101 *kil-kor* system may instead be analogous to the IncN plasmid pKM101 (35.4kb) which has been shown to contain potentially lethal genes that are needed for conjugal transfer (Winans and Walker 1985).
Figure 5.2  The kil-kor regulon of the IncP plasmid RK2

Horizontal open arrows indicate direction of transcription of operons. Vertical arrows indicate negative regulation by kor gene products. Curved arrow 1 shows positive interaction of trfA replication initiator with oriV. Curved arrow 2 shows positive regulation of KorA on korC. TcR: tetracycline resistance; ApR: Tn1-encoded ampicillin resistance; TeR: tellurite resistance. [Adapted from Kues and Stahl, 1989]

Figure 5.3  The kil-kor regulon of the IncN plasmid pKM101

Filled boxes indicate location of tra, kil, and kor genes. Vertical arrows indicate negative regulation by kor gene products. [Adapted from Winans and Walker, 1985]
In pKM101, two \textit{kil} genes (\textit{kilA} and \textit{kilB}) and two corresponding \textit{kor} genes (\textit{korA} and \textit{korB}) have been identified; both \textit{kor} genes are required to control the lethality of each \textit{kil} gene (Winans and Walker 1985). \textit{korA}, \textit{korB}, and \textit{kilA} are located adjacent to the transfer genes whereas \textit{kilB} is located within the \textit{traE} gene (Fig. 5.3). \textit{kilB} is required for conjugal transfer of pKM101 whereas \textit{kilA} is necessary for the small-colony morphology on defined media that is characteristic of pKM101-containing strains (Winans and Walker 1985). The observation that genetic loci involved in pKM101 and plJ101 plasmid transfer are lethal to the host when unregulated by \textit{kor} gene products suggests that finely tuned genetic switches modulate the transfer process (Winans and Walker 1985; Kendall and Cohen 1987, Stein \textit{et al.} 1989).

5.5 Protein structure analysis of Cop/KorB.

Many regulatory proteins use a helix-turn-helix structural motif to recognize and bind to specific DNA sequences; one helix assisting to position the second (recognition) helix correctly in the major groove of the target DNA. Certain amino acids within the recognition helix make contact with bases exposed in the major groove of the DNA. The bihelical units of these DNA-binding proteins have a pattern of conserved amino acid residues similar to the helix-turn-helix motif of \textit{\lambda} CI repressor and Cro (Pabo and Sauer 1984).

The Cop/KorB amino acid sequence was scanned for homology to the helix-turn-helix domains of twenty-one such DNA-binding proteins (Pabo and Sauer 1984) and also to the repressor gene of the \textit{Streptomyces} temperate phage \textit{\phi}C31 (Sinclair and Bibb 1988). Two potential helix-turn-helix domains were found spanning from residues 11 to 30 (KorB1) and from residues 41 to 60 (KorB2) (Fig. 5.4a). Both regions have the conserved Ala and Ile residues needed for interaction between the two helices (Pabo and Sauer 1984), and in both cases the "turn" sequence linking the helices contains a single Pro residue to aid in helix termination. Furthermore, the presence of the Gly residue in the "turn" of the first predicted
a) Putative DNA-binding domains of KorB

<table>
<thead>
<tr>
<th></th>
<th>Helix A</th>
<th>Turn</th>
<th>Helix B</th>
</tr>
</thead>
<tbody>
<tr>
<td>434cro</td>
<td>Gln — Thr — Glu — Leu — Ala — Thr — Lys</td>
<td>Ala — Gly — Val — Lys</td>
<td>Gln — Gln — Ser — Ile — Gln — Leu — Ile — Glu — Ala</td>
</tr>
<tr>
<td>korB1</td>
<td>Ala — Glu — Ala — Glu — Ala — Ala — Leu</td>
<td>Lys — Pro — Leu — Gly</td>
<td>Gln — Gln — Arg — Ile — Lys — Leu — Leu — Ala — Glu</td>
</tr>
</tbody>
</table>

b) Helical wheel representation of the α-helix turn α-helix region of KorB
bihelical unit gives the elbow conformational flexibility, making this region a stronger candidate for the DNA-binding domain.

The three-dimensional structures of α-helices can be represented by two-dimensional figures called helical wheels (Schiffer and Edmundson 1967), in which possible side-chain interactions and the general characteristics of helices are more easily visualised. These helical wheels depict projections of the amino-acid side-chains onto a plane perpendicular to the long axis of the α-helix. The perimeter of each wheel corresponds to the backbone of the polypeptide chain and the external spokes to the side-chains. For an α-helix with 3.6 residues per turn, adjacent side-chains in the linear sequence are separated by 100° of arc on the wheel.

Fig. 5.4b shows a helical wheel representation of the Cop/KorB α-helices of domain 1. In helix A, there are two clusters of hydrophobic residues making up stabilisation arcs which are separated by Glu residues. Helix B (or the recognition helix) can be divided into a non-polar (hydrophobic) and polar (hydrophilic) face; the Ile in the non-polar face could interact with the Ala residue of helix A while the residues making up the polar face may be involved in specific DNA base-pair contact. The oppositely charged Glu and Lys residues of helix B (spaced i, i+4 apart) could assist in stabilising helix formation via electrostatic interactions as in the case of the troponin-C chain (Sundaralingam et al. 1985), and thus forms a favourable ion pair along the polar face of the helix (Marqusee and Baldwin 1987). The use of site-directed mutagenesis to create variants of the Cop/KorB protein, in conjunction with crystallographic studies, would confirm the location of the recognition helix and determine which amino acid residues are essential for folding and function.

Kendall and Cohen (1988) predicted that the α-helix-turn-α-helix region spans from residues 12 to 32 (KorB1 domain) or residues 56 to 75 (KorB2 domain). However, neither of these regions seem probable because in the KorB1 domain, a Pro residue is incorporated in helix A which would substantially destabilize any
helix formation. Furthermore, they suggested that helix B would start at residue 24, an Arg residue which has virtually no ability to initiate helix formation (Scheraga 1978). Kendall and Cohen (1988) claim that the next possible bihelical unit starts at residue 56 and 67 but again, Asn and Arg are both extremely poor initiators of helix formation (Scheraga 1978) and it is thus highly unlikely that the helix-turn-helix motif would form in this part of the protein sequence. Subsequent work by Stein and Cohen (1990) involving mutagenesis of the KorB protein suggests that their proposed KorB2 domain does not contain the DNA-binding site involved in the control of the \textit{kilB} promoter, since insertions of 8bp linkers into the \textit{XmaI} site (lying within the KorB2 domain) did not substantially affect KorB repressor activity.

5.6 Cop/KorB binding on \textit{sti}.

Deng \textit{et al.} (1988) have reported that the \textit{cop} gene product negatively regulates the copy number of Sti+ plasmids via the \textit{sti} determinant. Cop could act to decrease copy number in two possible ways:

a) firstly, by binding to the double-stranded form of \textit{sti} and inhibiting the \textit{production} of ssDNA by physically blocking host Rep protein from synthesizing any ssDNA intermediates. This was shown not to occur since active Cop/KorB protein does not recognize or bind to a double-stranded 0.7kb \textit{BclI(57)-SpeI(53)} fragment containing the \textit{sti} determinant in gel-retardation assays (Section 3.7.1, Fig. 3.18);

b) secondly, by binding to the single-stranded form of \textit{sti} and inhibiting the \textit{conversion} of ssDNA to double-stranded plasmid DNA (which would lead to a pool of ssDNA in the cell). This was also shown not to occur since ssDNA does not accumulate in \textit{S. lividans} harbouring pIJ702::\textit{cop/sti+} (pQR411a/b) constructs compared to \textit{S. lividans} harbouring pIJ702::\textit{sti+} (pQR421a/b) constructs (Section 4.2, Fig. 4.2). If Cop was able to bind to ssDNA at \textit{sti} as the means of controlling copy number, then ssDNA would accumulate in the cell since it would not be converted to the double-stranded form of the plasmid. This situation is not likely to
occur, since the accumulation of ssDNA is not normally observed for pIJ101. Furthermore, Pigac et al. (1988) have suggested that ssDNA is highly recombinogenic and that the undesired accumulation of vast amounts of ssDNA in the cell causes plasmid structural instability and rearrangements involving insertions and deletions. Thus, it seems unlikely that Cop would control copy number by inhibiting the conversion of ssDNA to the double-stranded form as the accumulation of ssDNA would lead to plasmid structural instability.

From basic analysis of plasmid yields obtained for Sti+ plasmids possessing and lacking cop/korB, it would seem that Cop/KorB does not significantly alter copy number. However, it should be noted that it is difficult to accurately assess plasmid copy number in *S. lividans* since copy number can vary greatly depending on when the sample was taken during growth (C. Wrigley-Jones, personal communication). Hence, the initial observations reported by Deng et al. (1988) concerning the copy number of Sti+ and Cop+Sti+ plasmids may be unreliable.

*sti* was confirmed to be the site for second-strand synthesis since insertion of *sti* in the correct orientation with respect to the basic replicon resulted in the decrease of ssDNA production from pIJ702 (regardless of the presence or absence of Cop/KorB) (Section 4.2, Fig. 4.2). *sti* reinserted in the opposite orientation with respect to the basic replicon had no effect on pIJ702 which continued to accumulate ssDNA (Section 4.2, Fig 4.2). Thus, *sti* is only active as the site for second-strand synthesis in its correct orientation with respect to the basic replicon as would be expected since synthesis of ssDNA is strand-specific.

The pIJ101 *sti* determinant was further defined to lie on a 0.53kb *SpeI*(53)-*SstII*(55) fragment (Section 4.3). Deng et al. (1988) have located *sti* to a 200bp fragment (delineated at one end by the *FspI*(52) site) which is almost completely contained within this 0.53kb *SpeI*(53)-*SstII*(55) fragment. However, Deng et al. (1988) claim that the *SpeI*(53) site lies within the *sti* determinant; this is clearly not the case since *sti* cloned on the *SpeI*(53)-*SstII*(55) fragment is still active as the site for second-strand synthesis (Section 4.3). Furthermore, the *SpeI*(53) site is situated...
immediately adjacent to the FspI(S2) site that defines one end of the sti fragment according to Deng et al. (1988), and thus there are no bases separating these two restriction enzyme sites.

The insertion of sti into pIJ702-pUC8 shuttle vectors, which were previously shown to be structurally unstable in S. lividans (R. Barallon, personal communication), seemed to stabilize these plasmid constructs. This observation may be important for the further development of efficient and stable cloning vectors based on pIJ702 for use in S. lividans. However, it should be noted that sti cloned in pUC8 was inserted into pIJ702 at the PstI site (Section 4.1, Fig. 4.1) whereas the shuttle vectors were constructed by inserting BamHI-cut pUC8 into pIJ702 cut with BglII and although unlikely, the instability of the shuttle-vectors could be due to the positional effect of pUC8 in pIJ702 rather than to the lack of a site for second-strand synthesis.

Thus, functional ssi sites may be crucial for maintaining plasmid structural stability by preventing the formation of large amounts of highly recombinogenic ssDNA, since the insertion of sti into the unstable pIJ702-pUC8 vectors prevented them from rearranging or accumulating ssDNA. This correlation between structural stability and the lack of detectable ssDNA has been reported for other organisms. For instance, in Staphylococcus aureus pT181-like plasmids, palA is needed for normal replication and stability since rearrangements affecting palA results in ssDNA accumulation and plasmid structural instabilities (Gruss et al. 1987; Novick 1989). Similarly, in the streptococcal plasmid pLS1, deletions affecting the ssi region causes the accumulation of ssDNA intermediates and plasmid instability (del Solar et al. 1987).

Almost all ssi sequences contain potential stem and loop structures, and both the nucleotide sequence and secondary structure may play critical roles in the functional activity of these signals (Marians et al. 1982; van der Ende et al. 1983; Stuitje et al. 1984; del Solar et al. 1987; Gruss. et al 1987; Bahk et al. 1988; Boe et al. 1989). Several potential stem-loop structures were identified within the 0.53kb
Figure 5.5

The 0.53kb SpeI(53)-SstII(55) sti region was scanned for stable inverted repeats. The most stable stem-loop structure found within this defined sti region lies between nucleotides 7310 and 7385 on the pJ101 sequence (Appendix I). The identified stem-loop structure has a stability of $\Delta G = -64.00\text{kcal/mol}$ as defined by Tinoco et al. (1973). Filled circles indicate bases that are homologous with the TAGCGT consensus sequence found in ssi regions of several plasmids and phages (refer to Fig. 5.6).
Figure 5.6  Nucleotide sequence comparison of DNA regions containing *ssi* sites of phages and plasmids

<table>
<thead>
<tr>
<th>Plasmid/Phage</th>
<th><em>ssi</em> nucleotide sequence</th>
<th>Host</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>pEl94</td>
<td>GCGAAAGTAGTAGCGACAGCTATTAA</td>
<td>Staphylococcus</td>
<td>(1)</td>
</tr>
<tr>
<td>pC221</td>
<td>AGCGGTCAAGTAGGTACAGCTATTAA</td>
<td>Staphylococcus</td>
<td>(1)</td>
</tr>
<tr>
<td>pC194</td>
<td>AATGTCGGCATAGCGTGAAGCTATTAA</td>
<td>Staphylococcus</td>
<td>(1)</td>
</tr>
<tr>
<td>pLS1</td>
<td>CGAAGGGCTTTAGGCTTCCGAGGA</td>
<td>Streptococcus</td>
<td>(1)</td>
</tr>
<tr>
<td>pLS1 (att)</td>
<td>TTCTTATGTAGCCGTGCTGTCATTTC</td>
<td>Streptococcus</td>
<td>(1)</td>
</tr>
<tr>
<td>ØX174</td>
<td>GGGCAAAATT1TTAATTTTTGCGCTGA</td>
<td>E. coli</td>
<td>(1)</td>
</tr>
<tr>
<td>M13</td>
<td>GGGCTTCCTTTAGGTTCCGAGTTTAG</td>
<td>E. coli</td>
<td>(1)</td>
</tr>
<tr>
<td>pACYC184</td>
<td>TCGAAGACCAGGACGTACAGTACAG</td>
<td>E. coli</td>
<td>(2)</td>
</tr>
<tr>
<td>p15A</td>
<td>TCGAAGACCAGGACGTCAGGATCAG</td>
<td>E. coli</td>
<td>(2)</td>
</tr>
<tr>
<td>pBR322</td>
<td>CGGAACGACCAGCGGACGCGAGTCAG</td>
<td>E. coli</td>
<td>(2)</td>
</tr>
<tr>
<td>pIJ101</td>
<td>GCGAAGACAGGACGTCGGCCGCGGCC</td>
<td>S. lividans</td>
<td>this work</td>
</tr>
<tr>
<td>Consensus</td>
<td>TAGCGGT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spel(53)-SstII(55) sti sequence, the most stable having a ΔG of -64kcal/mol which could form the stem-loop structure in vivo shown in Fig. 5.5 (L. Radnedge, personal communication). This stem-loop structure contains five out of six bases which appear in a hexanucleotide consensus sequence (TAGCGT) found in ssi sites of staphylococcal (Gruss et al. 1987), streptococcal (del Solar et al. 1987), and E. coli (Bahk et al. 1988) plasmids (L. Radnedge, personal communication). A comparison of the hairpin loop and the sites of second-strand synthesis of these plasmids is shown in Fig. 5.6 (L. Radnedge, personal communication).

5.7 Definition of the pIJ101 minimal replicon.

Kieser et al. (1982) have previously defined the limit of the pIJ101 basic replicon capable of autonomous replication in S. lividans to a 2.2kb SstII(16)-SstII(63) fragment. This 2.2kb region was confirmed to be viable in S. lividans when cloned into the E.coli vecor pBGS19- together with the thiosstrepton gene (pQR434) (Section 4.4, Fig. 4.4). This 2.2kb essential region contains the rep ORF, a second smaller ORF (orf56) and a non-coding region of DNA (Kendall and Cohen 1988). Results from deletion analysis indicate that the basic replicon is a 2.0kb Ball(15)-SstII(63) fragment (pQR435) containing only the rep ORF and the non-coding region of DNA (Section 4.4, Fig. 4.4). Thus, the orf56 gene product is not necessary for maintenance and replication of the plasmid. As yet, no function has been assigned to the orf56 region and it is not known whether this ORF does actually encode a functional protein product. Kendall and Cohen (1988) have speculated that since orf56 is located only 70bp away from the kilB gene, orf56 and kilB may belong to the same operon.

The non-coding region of DNA is important for survival of the plasmid in S. lividans since either removal of half of it by cleaving at the Apal(14) site (pQR446) or removal of the entire non-coding region by cleaving at the NotI(12) site (pQR436) resulted in non-viable plasmids (section 4.4, Fig. 4.4 and 4.5). Thus, this 280bp Apal(14)-Ball(15) region of DNA may contain either the rep promoter or the
origin of replication, or both. This 280bp ApaI(14)-BalI(15) region showed no homology to the plus origin consensus sequences of the staphylococcal pE194-like or pC194-like plasmids (Gruss and Ehrlich 1989). Furthermore, attempts to map the rep transcription start site were not successful - probably because the rep mRNA was expressed at low levels and at only specific times during cell growth. An alternative approach to investigate whether this 280bp ApaI(14)-BalI(15) fragment contains either the origin of replication or the rep promoter would be to subclone this fragment from pQR435 and place it into a promoter-probe vector.

It should be noted that both pQR434 and pQR435 were unstable in *S. lividans* and accumulated ssDNA (Section 4.4). Rearrangements in general seemed to affect the pBGS19- part of the construct, which is the *E. coli* cloning vector. Since the introduction of *sti* into pIJ702-pUC8 shuttle vectors prevented these plasmids from rearranging in *S. lividans*, *sti* was inserted into pQR435 in order to try to stabilize this construct. However, pQR435 with *sti* inserted in the correct orientation with respect to the basic replicon was still structurally unstable and accumulated ssDNA. This may be due to several reasons. Firstly, the alignment between the *sti* site and the rep gene in pBGS19- may not be optimised, and the intervening thiostrepton resistance gene may interfere with the replication process since it is an efficiently transcribed and translated gene. Secondly, there may be an additional factor located on pU702 that is needed for efficient replication which has been removed during the construction of the minimal replicon. Thirdly, the existence of the f1 origin of replication on pBGS19- may cause structural problems during production of ssDNA since it is also capable of forming a strong stem-loop *in vivo*. Thus, replication may have been terminated prematurely at the f1 site or other secondary sites present on pBGS19- having sequence or structural resemblance to the normal termination signal, leading to deletions and rearrangements of the plasmid. A similar problem was reported where hybrid plasmids composed of pC194 (or pE194), pBR322, and phage f1 sequences frequently underwent deletions in *E. coli* (Michel and Ehrlich 1986; Ballester *et al.* 1989; Sozhamannan *et al.* 1990), and it was proposed that the
pC194 Rep protein initiates replication at its normal origin but terminates prematurely at false termination sites in the vicinity of the f1 replication origin during rolling circle replication (Michel and Ehrlich 1986; Ballester et al. 1989).

A 1.4kb NotI(12)-SstII(63) fragment containing the rep ORF (pQR431a) produced a 50kDa protein in vitro when expressed from the lacZ promoter, confirming that the rep ORF does encode for a protein (Section 4.5, Fig. 4.8). However, no Rep protein was detected in vivo from E. coli cells harbouring pQR431a. Thus, the Rep protein may have been unstable when produced in E. coli, and may have been prone to proteolysis in vivo because of its relatively large size. An alternative explanation for the lack of Rep protein production in vivo in E. coli may have been due to the insufficient synthesis of rep message from the lacZ promoter.

5.8 Conclusion.

The main objectives set out at the start of the project were achieved. The kill-override function of the KorB protein was strongly established from results of gel retardation and footprinting assays using the kilB gene. However, no evidence was found for the alleged copy-number control function of the KorB protein proposed by Deng et al. (1988). The next stages of this work would be to over-express native KorB protein and to purify it using affinity chromatography since the operator sequence of the KorB protein is now known. Furthermore, fractions from a column could be easily analysed for presence of KorB protein by using the gel retardation assay developed in this study. Purified KorB protein would allow more quantitative studies to be done on KorB binding to the kilB and korB operator sites and thus to determine the relative difference in affinities between these two operators. Purified protein can also be crystallized to determine the three-dimensional structure of KorB and the location of the DNA-binding domain of the protein can be confirmed if co-crystals with its operator are generated. Site-directed mutagenesis would provide an
in-sight into which amino-acids are directly involved with base-pair contacts in the major groove of the DNA operator site.

The preliminary results obtained from the replication analysis of pIJ101 provides the basis for further work. The problem of adequate Rep protein expression in vivo in E. coli may be solved by subcloning the rep gene under the control of a stronger, more efficient promoter in order to over-produce sufficient Rep protein for studying replication control in pIJ101. Purified Rep protein can be used to test for DNA binding activity, sequence specific endonuclease activity and topoisomerase-like activity on the plus origin. This, in turn, would allow the precise definition of the sequence required for the nicking-closing and covalent attachment of Rep protein to ori. Long term goals would involve crystallization and mutagenesis of Rep protein to define its active site.
REFERENCES.


Scherer GFE, Walkinshaw MD, Arnott S, and Morre DJ (1980). The ribosome binding sites recognized by *Escherichia coli* ribosomes have regions with signal character in both the leader and protein coding segments. *Nucleic Acids Res* **8**:3895-3907.


APPENDIX I. Nucleotide sequence of pIJ101.

The sequence is numbered starting at the first base of the recognition sequence of the unique BamHI site. The translated amino acid sequence of the predicted proteins encoded by the plasmid are shown. (Taken from Kendall and Cohen, 1988). Mapped -10 and -35 promoter regions are boxed, and transcription start sites are arrowed (Buttner and Brown 1985, 1987; Stein et al. 1989).