NUCLEOTIDE ANALYSIS OF TWO ACTINOMYCETE AMINOGLYCOSIDE RESISTANCE DETERMINANTS

By

David John Holmes

A dissertation submitted for the degree of

Doctor of Philosophy

at the

University of Leicester

April 1989

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

To my parents

Abstract

NUCLEOTIDE ANALYSIS OF TWO ACTINOMYCETE AMINOGLYCOSIDE RESISTANCE DETERMINANTS

David J. Holmes

Resistance to aminoglycosides in the organisms that produce them is often ascribed to the well characterised and clinically important antibiotic modifying enzymes. However, at least three aminoglycoside producing actinomycetes, namely *Micromonospora purpurea*, *Streptomyces tenjimariensis*, and *Streptomyces tenebrarius* possess ribosomes that are refractory to some members of this class of drugs. In these cases, resistance is due to methylation of rRNA of the small ribosomal subunit. This study supports the possibility that this mechanism might be more widespread than hitherto suspected.

Two of the methylase genes have been analysed at the nucleotide level and their transcripts mapped. The gentamicin resistance methylase gene (*kgmA*) from *M. purpurea* codes for a 36 kDa protein consisting of 249 amino acids. Like most actinomycete genes, *kgmA* is not expressed in *E. coli* from its own promoter, although the determinant was expressed in this Gram-negative host as a result of DNA rearrangement. Sequence analysis of the mutated plasmid suggested that the methylase was expressed as a translational fusion with the *lacZ'* gene of pUC18, a view that was later confirmed. Transcript mapping revealed that *kgmA* is probably read from a single promoter but that it might be part of a polycistron.

The second gene examined confers resistance to kanamycin and apramycin, and originated in *S. tenjimariensis*. This determinant (*kamA*) was shown to encode a predicted protein of 155 amino acids with a molecular weight of 19 kDa. Unlike *kgmA*, this gene could not be expressed as either a transcriptional or translational fusion in *E. coli*. Transcription of *kamA* is directed by tandem promoters and is a monocistron since the the transcript terminates only 160 bp downstream of the stop codon.

CONTENTS

TABLES AND FIGURES

ABBREVIATIONS

A_x	Absorbance at a wavelength of x nm in a 1 cm light path
Am	Ampicillin
ATCC	American Type Culture Collection
ATP	Adenosine 5'-triphosphate
bp	Base pairs
BPB	Bromophenol blue
BSA	Bovine serum albumin
CIAP	Calf intestinal alkaline phosphatase
Cm	Chloramphenicol
cpm	Counts per minute
CTAB	Hexadecyltrimethylammonium bromide
CTP	Cytosine 5'-triphosphate
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytosine 5'-triphosphate
ddATP	2',3' -dideoxyadenosine 5'-triphosphate
ddCTP	2',3' -dideoxycytosine 5'-triphosphate
ddGTP	2',3' -dideoxyguanosine 5'-triphosphate
ddTTP	2',3' -dideoxythymidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside 5'-triphosphate
DMSO	Dimethyl sulphoxide
DNase	Deoxyribonuclease

DTT	Dithiothreitol
EDTA	Diaminoethane tetra-acetic acid
EGTA	Ethylene glycol-bis (b-aminoethyl ether) N,N,N',N' -tetra acetic acid
GTP	Guanosine 5'-triphosphate
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethane sulphonic acid
IMS	Industrial Methylated Spirits
IPTG	Isopropyl β-D-thiogalactopyranoside
kb	Kilobase pairs
kDa	Kilodalton
MPa	Megapascals
mRNA	Messenger ribonucleic acid
N.C.I.B.	National Collection of Industrial Bacteria
O.D.	Optical density
PEG	Polyethylene glycol
PEP	Phospho(enol) pyruvate
PIPES	Piperazine-N,N'-bis [2-ethanesulphonic acid]
psi	pounds per square inch
rev min $^{-1}$	Revolutions per minute
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
tRNA	Transfer ribonucleic acid
S	Svedberg unit (10^{-13} seconds)
SDS	Sodium dodecylsulphate
TCA	Trichloroacetic acid

TES	N-tris (hydroxymethyl) methyl-2-aminoethane sulphonic acid
TEMED	N,N,N',N',-tetramethylethylenediamine
Tris	Tris(hydroxymethyl) aminomethane
TTP	Thymidine 5'-triphosphate
UTP	Uridine 5'-triphosphate
UWGCG	University of Wisconsin Genetics Computer Group
(v/v)	Ratio of volume to volume
(w/v)	Ratio of weight to volume
X-gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

ACKNOWLEDGEMENTS

The list of people deserving acknowledgement is endless and I begin by apologising for any accidental omissions from the following list.

Firstly, I would like to thank Professor Eric Cundliffe, not only for allowing me to work in his laboratory and for initiating the process by which I became registered for this degree, but in addition for sound advice, encouragement and copious amounts of beer.

Many other colleagues also deserve my gratitude, particularly Dr. Jill Thompson and Dr. Alan Beauclerk who showed great patience during my early times in the lab, but also others from Leicester including Dr. Tricia Skeggs, Dr. Amrik Thiara, Dr. Magdalena Zalacain and Miss Gail Jenkins. I should bring attention as well, to Mr. Derek Gamble whose efficiency as storeman and good humour made working at Leicester so much easier.

Many thanks are due to others who have been important to my academic career such as Mr. Roger Higton and particularly Dr. Michael Cannon whose enjoyment of life infuses all those who meet him (especially me).

The support of friends is always important and those of note who have remained so after many years include Nick, Keith, James, and Carl. In this context I would particularly like to thank Magda, of course.

I express my appreciation to Mr. Ian Riddell and the Central Photographic Unit for photographs and the audio-visual services for their patient assistance with the final production. I also thank Magda for typing this thesis.

Finally, I am very grateful to my family; to Craig, my mother and father, for their constant support and encouragement throughout my education. Special thanks are owed to Mr. David McAlpine who has been a considerable influence and who shares in any success I may achieve.

CHAPTER 1

INTRODUCTION

Introduction

The study of actinomycetes has been inspired by several motives, mostly concerning the biology of a differentiating prokaryote and the fact that this collection of organisms is the single largest source of antimicrobial agents. Work on the order *Actinomycetales* has focussed upon the genus *Streptomyces* and techniques have rapidly developed to optimise the manipulation of specific members of this genus, particularly *S. lividans* and *S. coelicolor*. The study of *Streptomyces* is an attractive proposition due to a number of novel features exhibited by these bacteria. Of importance to the researcher is the fact that *Streptomyces* are non-pathogenic. At least, the only adverse effect attributed to a streptomycete is the occurrence of scabies in cattle, otherwise these micro-organisms can be handled with impunity. More scientific interest has focussed on the observation that actinomycetes differentiate, first growing as branching basal mycelia, eventually putting up aerial hyphae which coil, septate and form spores (Fig. 1.1), and it is during this morphological development that the switch from primary to secondary metabolism occurs. The appearances of the various actinomycetes are strikingly different, from the coloured compounds they elaborate when grown on solid media and the colour of their spores, to the morphology of those spores when viewed under the electron microscope. Over the years, they have been classified according to various criteria, but many members of the order *Actinomycetales* have recently been reclassified, and in this dissertation will be referred to according to the latest nomenclature (Dietz, 1988) as shown in Table 2.1.

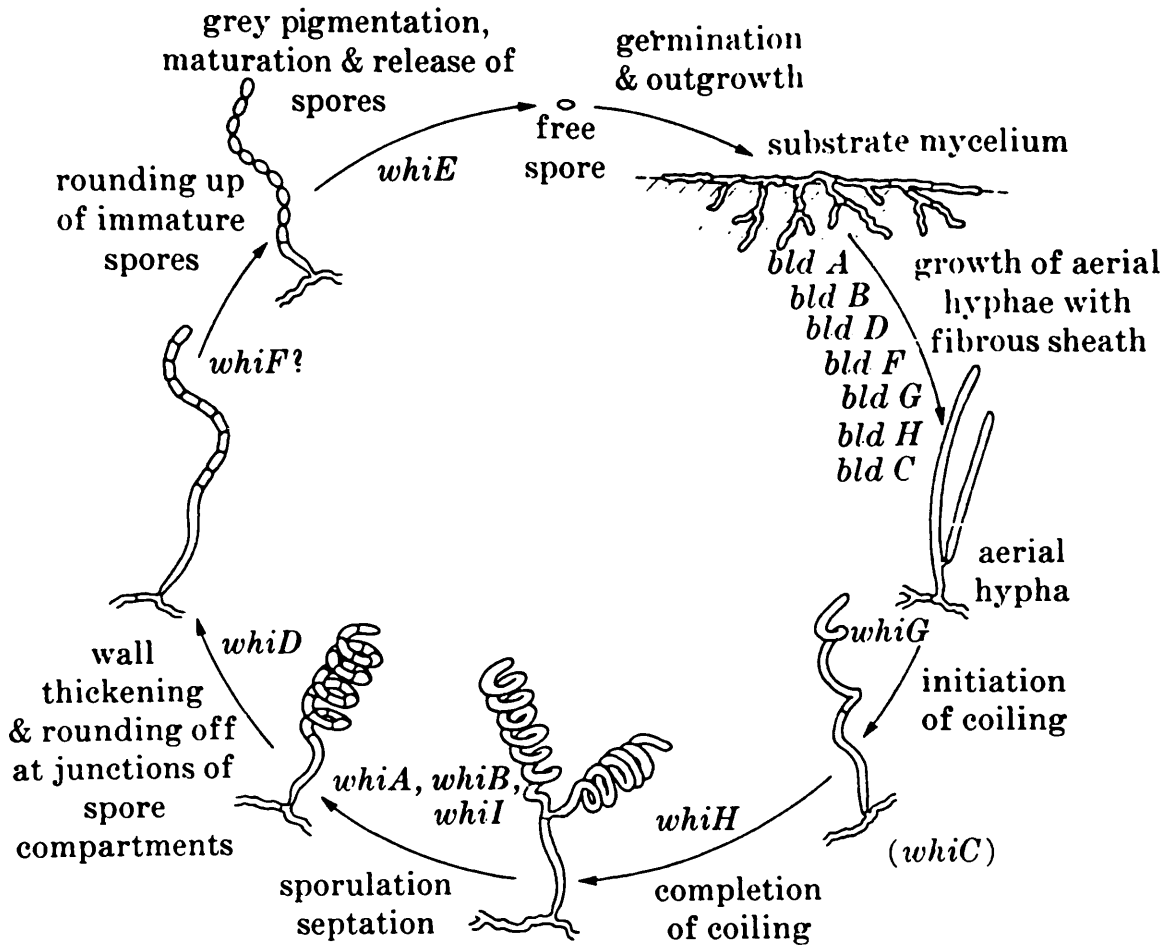


Fig. 1.1 Schematic representation of the developmental cycle of *Streptomyces coelicolor*.

Those genes associated with morphological differentiation are indicated at the points in the cycle which are interrupted by mutations in the various *bld* and *whi* genes. From Hopwood 1988.

Apart from their involvement in secondary metabolism, interest in antibiotics produced by actinomycetes has been stimulated by their obvious commercial value. The production of increased yields of antibiotic is a premium which until recently has been achieved by spontaneous and induced mutagenesis. The spontaneous mutations reflect a certain genomic instability observed in *Streptomyces*, and the role of this phenomenon in the occurrence of chloramphenicol sensitivity and arginine auxotrophy is well documented (Dyson and Schrempf, 1987; Altenbuchner and Cullum, 1985), although the range of spontaneous mutations includes deletions, duplications, transpositions, base pair substitutions and reading frame shifts. This genome instability has been used to explain the anomalously large genome size of *Streptomyces* since, renaturation studies on the DNA of *S. coelicolor* A3(2) led to an estimate of the genome size of about 10^4 kilobases, more than twice the amount of unique DNA possessed by *Escherichia coli* (Antonov *et al.*, 1978) although a recent estimate suggests a genome size of 6×10^3. It is unclear why *S. coelicolor* should contain so much more potential genetic information than *E. coli.* Although morphological development and secondary metabolism requires genes that do not occur in the non-differentiating bacterium, this probably does not account for all of the extra DNA. Moreover, another differentiating Gram-positive prokaryote, *Bacillus subtilis,* has a genome size roughly equal to that of *E. coli.* The direct mutagenic agents employed by commercial concerns include ultraviolet light (Drake and Baltz, 1976; Witkin, 1969, 1976) as well as DNA modifying compounds such as NTG and MMS. Indeed, even neutron bombardment has been used successfully to improve productivity of virginiamycin.

The employment of mutagens in an attempt to increase antibiotic yield was used extensively prior to the development of technical procedures for genetic recombination and gene cloning in *Streptomyces* and is still favoured by some groups. The view was that, since most antibiotic biosynthetic pathways are complex multistep processes which are poorly understood, general mutagenesis was preferable to the isolation of specific genes. However, the advent of procedures which allow the isolation of entire biosynthetic pathways along with the corresponding resistance gene(s) may lead to a more rationalised programme of strain improvement. In some cases the isolation of these biosynthetic gene clusters from cosmid libraries was made possible by the prior isolation of the appropriate resistance gene which is commonly clustered with them. This is a widespread feature of genes which encode the enzymes of biochemical pathways in *Streptomyces*, in contrast to those involved in morphological development (at least in *S. coelicolor* A3(2)) which are scattered around the chromosome. Using an appropiate antibiotic resistance gene then, to probe a cosmid library has allowed the isolation of complete biosynthetic pathways which can subsequently be studied in detail, identifying rate limiting steps. This approach has been successfully used to isolate genes involved in the biosynthesis of oxytetracycline (Rhodes *et al.*, 1984), streptomycin (Ohnuki *et al.*, 1985) and erythromycin (Stanzak *et al.*, 1986). Analysis of these gene clusters should resolve some of the questions concerning the regulatory signals involved in the switch from primary to secondary metabolism as well as allowing a logical programme of strain improvement by the use of genetic engineering to identify and alleviate the rate limiting step in a given pathway. Obviously this would generate a new limiting

reaction which in turn could be modified in an ever growing spiral of increased antibiotic production.

1. Gene cloning in *Streptomyces*.

Gene cloning in actinomycetes has been limited almost exclusively to the genus *Streptomyces* and it was the discovery that they harboured plasmids that made possible the construction of vectors for cloning. The earliest suggestions of the presence of plasmids in *Streptomyces* were based on the fact that strains could be "cured" of certain phenotypes, for example, *Streptomyces scabies* could be cured of melanin production by treatment with acridine dyes (Gregory and Huang, 1964). However, it now seems that this phenomenon was due to the gross deletions and rearrangements of DNA alluded to earlier (Schrempf, 1983). More accurately therefore, SCP1 was the first extrachromosomal element identified (Vivian, 1971) and the proof of its involvement in fertility and antibiotic production in *S. coelicolor* were the first examples of plasmid borne functions, although SCP1 could not be physically isolated probably due to its large size (>200 kb), its low copy number and, more to the point, the fact that is linear (Chater and Bruton, 1983; Kinashi *et al.*, 1987). The first *Streptomyces* plasmid to be characterised was SCP2* a fertility determinant from *S. coelicolor* (Bibb *et al.*, 1977), but since then a variety of plasmids have been identified in a great many strains (Table 1.1) and recent technical advances have aided their characterisation, allowing them to be used to construct cloning vectors.

Table 1.1. Naturally occurring *Streptomyces* plasmids.

Strain	Plasmid	Size (kb)	Copy Nº	Comments	References
S. ambofaciens KA-1028	pSA1	8.0	2	High proportion of plasmid associated with chromosome or membrane	Omura *et al.*, 1981 Ikeda *et al.*, 1982a,b
S. bikiniensis KA-421	pSB1	29	-		Omura *et al.*, 1981
	pSB2	57	-		Omura *et al.*, 1987
S. coelicolor A3(2)	SCP1	410-590	low	Linear plasmid encoding biosynthesis of and resistance to methylenomycin	Vivian, 1971 Kinashi *et al.*, 1987
	SCP2	31	1-2	Fertility factor	Schrempf *et al.*, 1975
	SLP1	9.4-14.5	1	Autonomous plasmids derived from the chromosome	Bibb *et al.*, 1981
	SLP4	-	-	Detected by pock formation	Hopwood *et al.*, 1983
	circle	2.6	0.2	Integrates into the chromosome No evidence for autonomous replication	Lydiate *et al.*, 1986
S. fradiae	-	420			
S. griseus	-	-	-	Plasmid DNA isolated from wild type is absent from *Streptomyces* non-producers	Xue *et al.*, 1981
S. lasaeiensis NRRL 3382R	-	520	-	Involved in lasalocid resistance	Kinashi *et al.*, 1987
S. lividans ISP5434	pIJ101	8.9	40-300	Replication, transfer and pock-forming functions have been mapped	Kieser *et al.*, 1982
	pIJ102	4.0	40-300		
	pIJ103	3.9	40-300		
	pIJ104	4.9	40-300		
S. ribosidificus ATCC 21294	pSR1	79	-	Present also in ribostamycin non-producing mutants	Okanishi *et al.*, 1980 Noriji *et al.*, 1980
S. venezuelae 3022a	pUC3	32	-	Does not code for chloramphenicol biosynthesis	Malik *et al.*, 1977 Ahmed and Vining, 1983

The development of methods for the isolation of DNA and a means by which this DNA could be introduced into *Streptomyces* protoplasts at high frequency (Bibb *et al.*, 1978) has led to the rapid evolution of DNA cloning techniques involving *Streptomyces* plasmids. The vectors routinely used have mostly been developed from the three plasmids SLP1.2, SCP2* and pIJ101, those based on the former two being low copy number, those on the latter of high copy number. Plasmid SLP1.2 (Fig. 1.2) was one of several episomes obtained after genetic crosses between *S. coelicolor* and *S. lividans*. Although forming an integral part of the chromosome in *S. coelicolor*, SLP1.2 is a autonomous replicon with a copy number of 4-5 in *S. lividans* (Bibb *et al.*, 1981). The presence of unique restriction sites which were evidently not in regions required for maintenance, transfer or replication, allowed plasmid SLP1.2 to be used to clone several genes, including, the thiostrepton resistance gene (*tsr*) from *Streptomyces azureus*, the gene which codes for a viomycin phosphotransferase (*vph*) from *S. vinaceus* and the neomycin phosphotransferase (*aph*) and acetyltransferase (*aac*) encoding genes from *S. fradiae* (Thompson *et al.*, 1980; 1982a; 1982b). These genes have subsequently been used in the construction of useful cloning vectors of which the most commonly used is pIJ702. The replicon of pIJ702 is derived from the previously mentioned pIJ101 (Fig. 1.3) which was isolated from *S. lividans*. In addition to the minimal replicon of pIJ101, plasmid pIJ702 contains the thiostrepton resistance methylase (*tsr*) and a tyrosinase (*mel*) from *S. antibioticus* (Katz *et al.*, 1982) (Fig. 1.4). When grown on media supplemented with tyrosine, *S. lividans* harbouring pIJ702 produces a

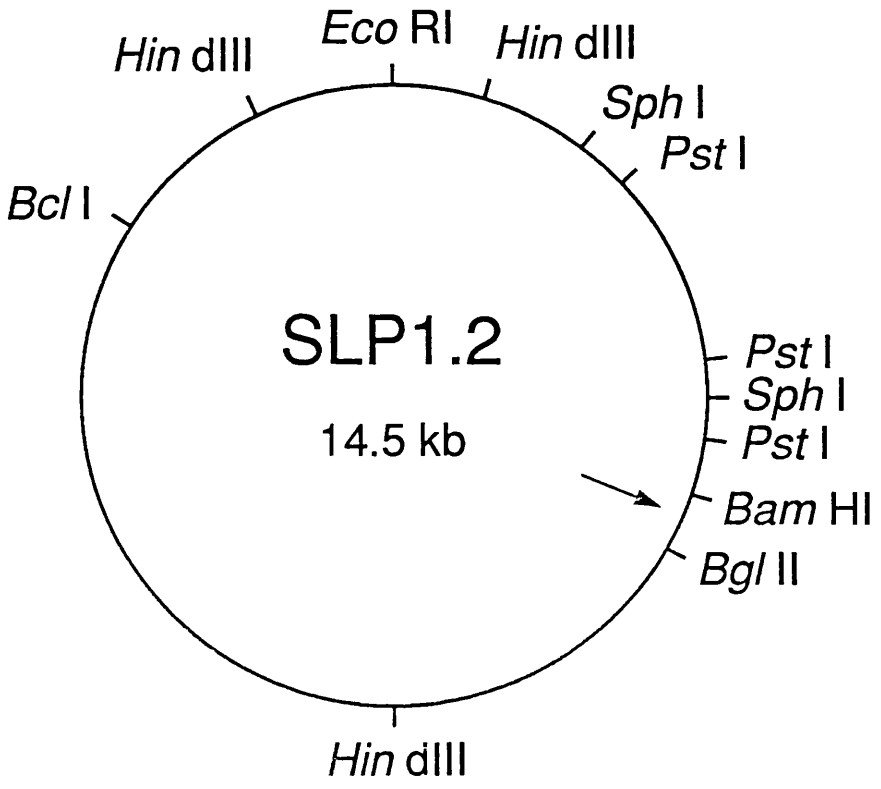


Fig.1.2 Restriction endonuclease cleavage map of plasmid SLP1.2

Unique restriction sites exist for *Bam*HI, *Bcl*I, *Bgl*II, *Cla*I, *Eco*RI, *Kpn*I, *Xho*I. The arrow indicates where the plasmid is integrated into the *S. coelicolor* A3(2) chromosome.

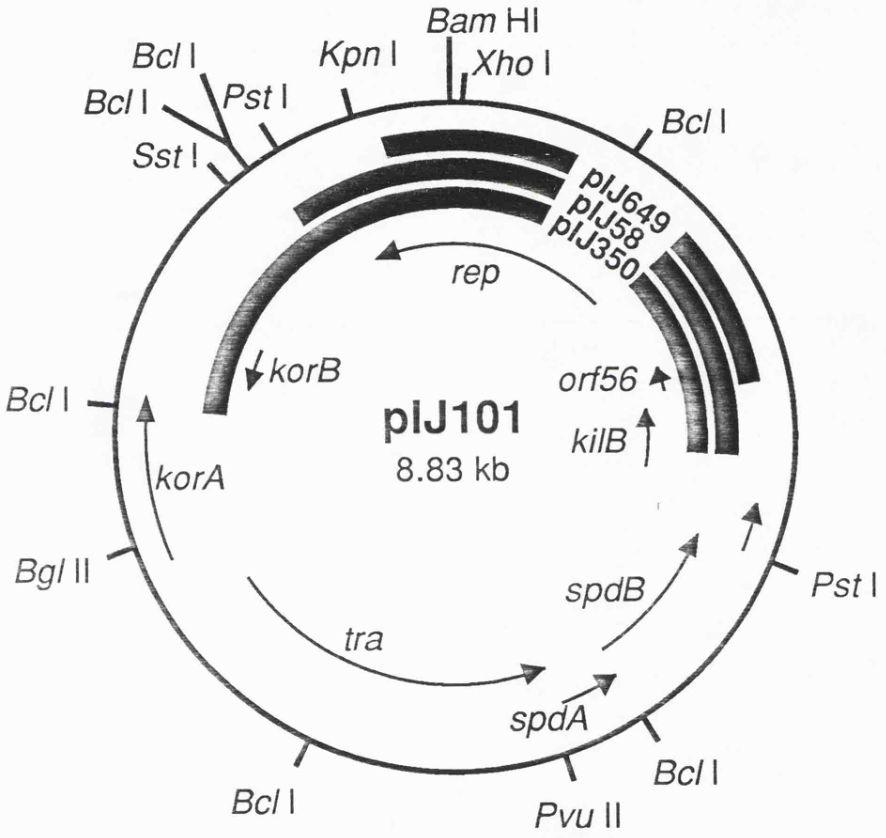


Fig.1.3 Restriction endonuclease cleavage and functional map of the high copy number plasmid pIJ101.

The precise map as well as the location of ORFs were made available following the publication of the nucleotide sequence of the plasmid.

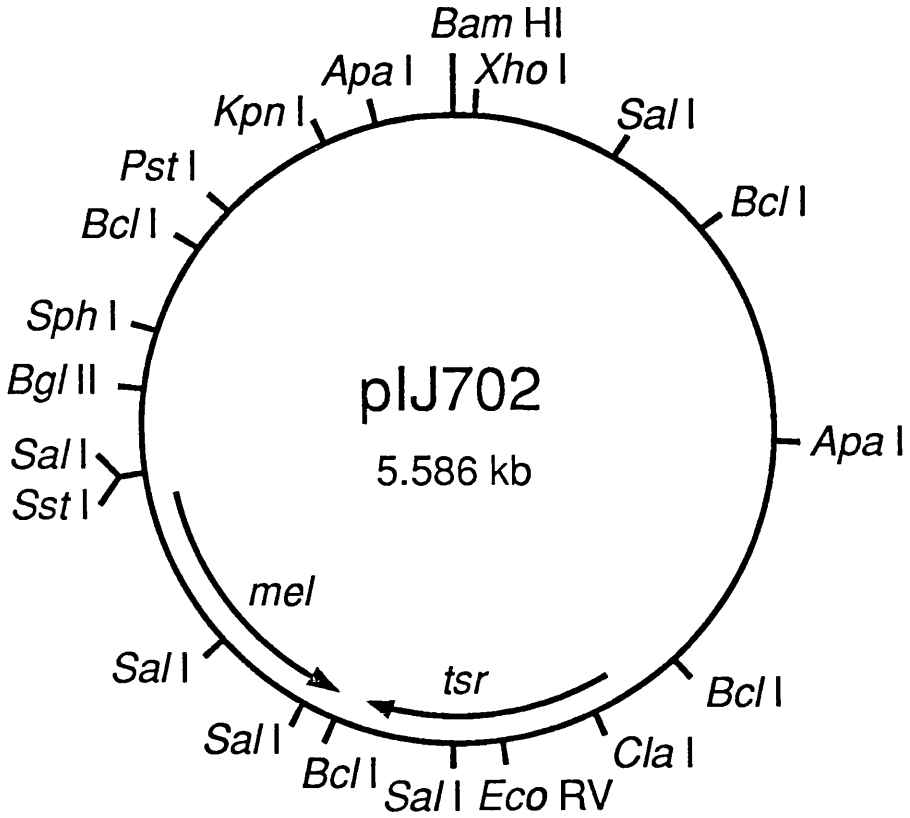


Fig.1.4 Restriction endonuclease cleavage map of the high copy number plasmid pIJ702.

The cloning vector pIJ702 is a derivative of the high copy number plasmid pIJ101 and contains the thiostrepton resistance gene (*tsr*) from *S. azureus* as a selectable marker. It also incorporates DNA (*mel*) from *S. antibioticus* which directs the synthesis of the black pigment melanin. Insertion of DNA into the *Sph*I, *Bgl*II, or *Sst*I restriction sites within *mel* usually prevents pigment production.

black pigment due to the action of the tyrosinase. Cloning pieces of DNA into any of the three unique restriction sites within the *mel* gene leads to insertional inactivation and is used to indicate the number of recombinants while thiostrepton is employed as selection for primary transformants. Plasmids derived from pIJ101 have a broader host range than those from SLP1.2, in addition to a much higher copy number of between 40 and 300 depending on the growth phase. Although it may be useful to have cloning vectors with different copy numbers since multiple copies of certain genes might be undesirable, there is insufficient evidence to indicate whether cloning certain DNA using high copy number vectors is potentially deleterious to the host. Certainly, deletions or rearrangements have occurred when DNA has been introduced into pIJ702 (Feitelson and Hopwood, 1983; Feitelson *et al.*, 1985; Skeggs *et al.*, 1987; Calcutt and Cundliffe, manuscript in preparation); however, this has not been directly linked to copy number effects. On the other hand, attempts to clone actinorhodin genes of *S. coelicolor* A3(2) failed when using the same vector, but succeeded when the low copy number SCP2* derivative, pIJ922 (Fig. 1.5), was used (Malpartida and Hopwood, 1983). Despite such problems, pIJ702 remains the most popular cloning vector for *Streptomyces* .

As well as the range of plasmids, a number of bacteriophage vectors have been developed and are available for DNA manipulation in *Streptomyces*. Derivatives based on øC31 (Harris *et al.*,1983; Rodicio *et al.*, 1985) contain the *tsr* gene and/or the *vph* gene for selection and are available with or without an *att* site. This DNA sequence enables

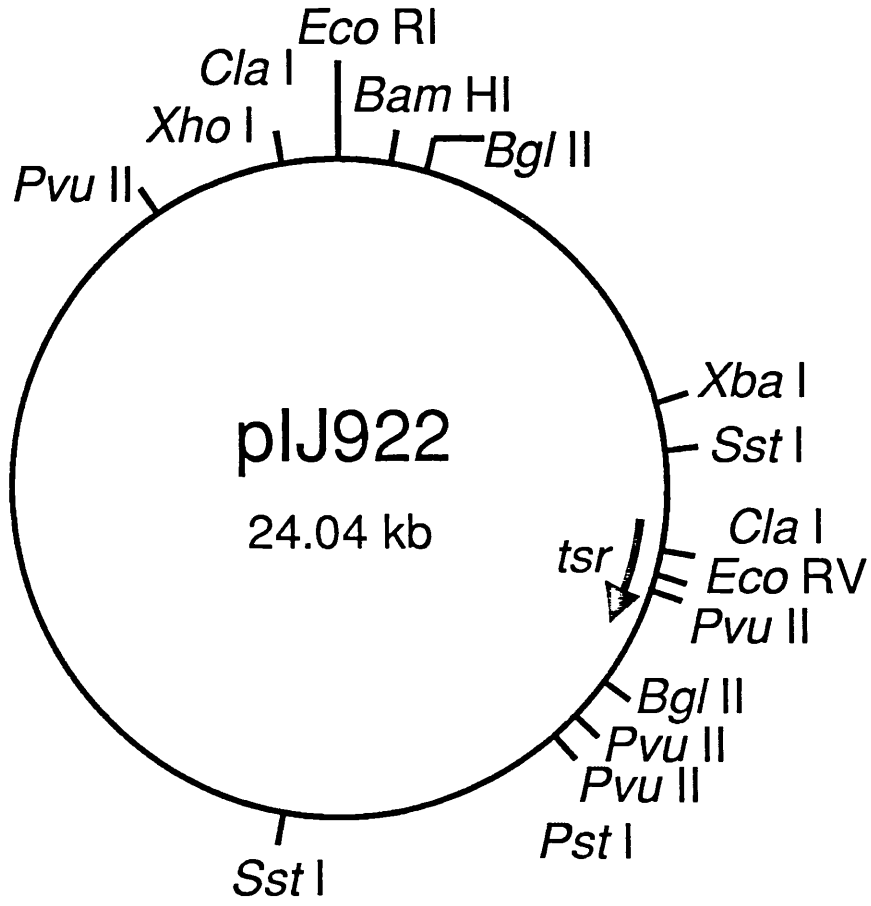


Fig. 1.5 Restriction endonuclease map of plasmid pIJ992.

This is a low copy number plasmid derived from SCP2* and was used to clone the actinorhodin biosynthetic gene cluster.

the phage to be integrated into the chromosome at a complementary sequence to become a lysogen. Bacteriophage lacking the *att* site can still lysogenise a host, but only if they contain inserted DNA which is homologous to a sequence in the chromosome. However, if the insert does not contain a complete transcriptional unit, a mutation will arise after recombination. This is the basis of "mutational cloning" which allows genes to be identified without the prior isolation of specific mutants and enables the transcriptional patterns within cloned fragments to be analysed (Chater and Bruton, 1983; 1985; Malpartida and Hopwood, 1986).

Several so-called shuttle vectors have been described which can replicate in different genera, e.g. in *E. coli* and *Streptomyces*. Due to certain difficulties these have not been used extensively for cloning experiments. The major problems include instability in *Streptomyces*, sometimes resulting in complete loss of the plasmids or deletions which prevent subsequent stable replication in *E. coli*. Nevertheless, the advantages of an efficient, stable shuttle vector are obvious, since plasmid construction especially where multiple cloning steps are anticipated, can be more conveniently performed in *E. coli* because of its faster growth rate.

The use of cloning methods to isolate genes has given researchers unquestionable advantages. For example, the expression in *S. griseofuscus* of a gene, *tlrA*, from *S. fradiae* which gives resistance to tylosin has facilitated the discovery of the mode of action of its product (Zalacain and Cundliffe, manuscript submitted). This methylase which

modifies 23S rRNA has been cloned in a high copy number plasmid and thereby has circumvented the need for extensive purification of the enzyme from *S. fradiae*. More to the point, if crude extracts from *S. fradiae* had been used as the source of methylase activity in that study, the possibility would have remained that this strain might have possessed several rRNA methylases not present in a control strain. In any case, it has now been shown that, while *S. fradiae* is constitutively resistant to tylosin, *tlrA* is only expressed under growth conditions in which tylosin is produced, despite the fact that the antibiotic is not the inducer. So, transplanting *tlrA* into *S. griseofuscus* greatly facilitated the detailed study of this gene and its product. Similarly, removing a gene from its background in *S. kanamyceticus* revealed its involvement in kanamycin resistance in that organism. Resistance to the antibiotic in the producer had been previously attributed to acetyltransferase activity since the ribosomes from *S. kanamyceticus* were sensitive to kanamycin *in vitro*. As a result of genetic manipulation, a ribosomal resistance gene was transplanted from an environment in which its expression was formerly controlled and was now constitutively expressed (Nakano *et al.*, 1984).

Increasing the number of copies of a gene has proved useful in the identification of the corresponding promoters. Attempts to detect in *S. azureus* mRNA complementary to the cloned *tsr* gene have failed, whereas they have been successful with thiostrepton resistant clones in *S. lividans*. This suggests that transcription of the *tsr* gene in *S. azureus* is either very weak or tightly controlled. Accordingly, when promoter sequences from a number of *Streptomyces* genes were

compared, those preceding the *tsr* gene were found to be weakest when tested for their ability to direct transcription of a promoterless indicator gene (Ward *et al.*, 1986).

The development of these efficient systems has allowed the investigation and isolation of a number of genes. Typically, these have been antibiotic resistance genes although genes involved in primary metabolism and differentiation have now been isolated as well. In addition it is now possible to pursue the nucleotide sequences that interact with RNA polymerase, ribosomes and other nucleic acid binding moieties, the study of which has already answered some of the questions concerning the switches which control the development of actinomycetes.

2. Gene expression in *Streptomyces*.

The production of a protein from the genetic information can be essentially divided into two parts; transcription, in which the DNA is copied as mRNA and translation in which the ribosome decodes the message as an amino acid sequence. Although the two processes are inextricably linked in the cell, they will be considered separately in this section for clarity.

2.1. Transcription.

Although more is known about *E. coli* than any other prokaryote, it would be a mistake to believe that we know everything about this bacterium. Only about 25% of the approximately 4000 genes

of *E. coli* have been identified (in any way at all) and models of some of the better characterised biochemical reactions are soon revealed as being flawed when probed more deeply. The process of transcription falls into this latter class. The way in which a gene is transcribed is well documented. An enzyme (RNA polymerase) binds to DNA at a defined site (promoter) upstream of a coding sequence. The RNA polymerase synthesises a complementary copy of one strand of the DNA (elongation) and eventually stops at a defined site with (rho dependent) or without (rho independent) extraneous protein factors.

More is probably known about the initiation step of transcription than any other possibly because a number of genes are understood to have their expression controlled at this point. In the case of *E. coli* the RNA polymerase binding site is apparently well defined and the promoter consists of two regions considered to be important in RNA polymerase recognition. These two regions are hexameric sequences centred approximately 10 and 35 bases upstream of the transcription initiation point. The identification of a number of transcription start points has allowed the definition of consensus sequences for the "-10" and "-35" regions and the so-called "spacer" between them. They are, for the former TATAAT and TTGACA for the latter, separated by a distance of 17 bp (Hawley and McClure, 1983). Experiments using the *ampC* promoter, which differs from the consensus in 1 position in each of these three regions (Fig. 1.6), supported this hypothesis, since mutations in this promoter towards the consensus resulted in an increase in transcriptional activity (Jaurin

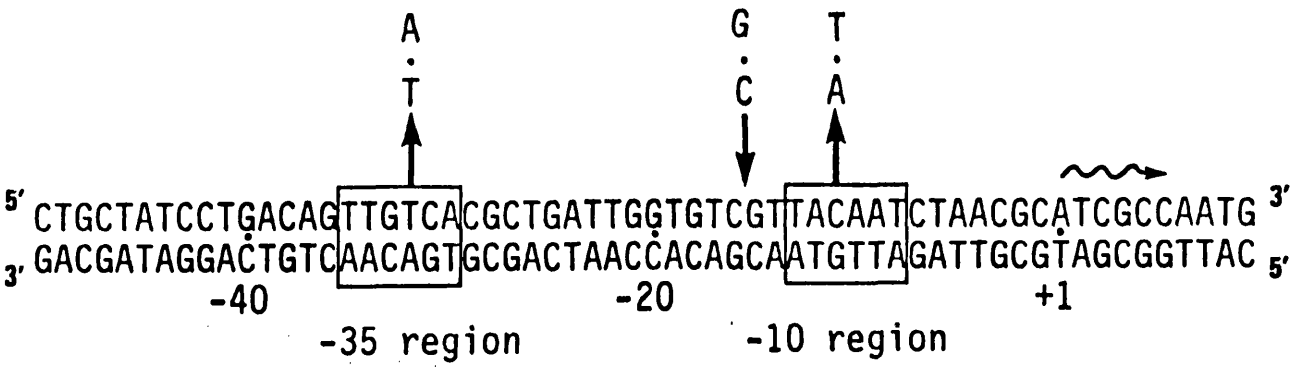


Fig. 1.6 The nucleotide sequence of the *E. coli ampC* promoter

This promoter differs from the consensus by a single nucleotide in each of the three positions considered to comprise a promoter. Mutations in this sequence, towards that of the consensus for an *E. coli* promoter, are indicated and each result in an increase in transcriptional activity in *S. lividans* as well as *E. coli*.

et al., 1982). To reduce the RNA polymerase binding site to these regions, although useful, is an oversimplification, because when this holoenzyme binds to DNA, it does so over approximately 70 nucleotides, ranging from 50 nucleotides upstream of the initiation site to 20 bases downstream (Schmitz and Galas, 1979; Siebenlist *et al.*, 1980). In various studies, the merits of the consensus sequence have been tested and in one case the upstream promoter of the *E. coli gal* operon was mutated. The -35 region was removed and replaced with various non-canonical sequences and, although reducing the promoter activity to a half of the wild-type level, transcription still initiated at the same nucleotide, implying that the -35 sequence was not essential for RNA polymerase recognition (Ponnambalam *et al.*, 1986). Moreover, when a series of promoters were graded by their *in vivo* strength, a synthetic "consensus" promoter was found to be one of the less efficient promoters (Deuschle *et al.*, 1986). In fact, none of the stronger promoters contained both of the consensus hexameric sequences. So, although there is little doubt that these regions of DNA are important for RNA polymerase recognition, these various lines of evidence suggest that sequence information flanking this area may be important for promoter function as well. A prototype consensus sequence was determined for a region from +1 to +20 (+1 being the initiation nucleotide) based on some coliphage T5 promoters such as P_{N25} (Gentz and Bujard, 1985). When a synthetic consensus prokaryotic promoter employed the +1 to +20 sequence of P_{N25}, a fourteen fold increase in promoter activity was observed. Similarly, when the same promoter had its +1 to +20 sequence replaced with that of $P_{D/E20}$, another T5 promoter, the strength was again increased *in vitro* and *in vivo*

(Kammerer *et al.*, 1986). These results were supported when "downstream regions" from the *lac* and *trp* promoters were used to replace that of the synthetic consensus sequence and found to have an intermediate effect (Kammerer *et al.*, 1986). Although this region is not involved in promoter recognition by RNA polymerase, it does demonstrate the importance of sequences outside those normally considered to comprise a promoter.

RNA polymerase apo, or core, moiety consists of four polypeptides ($\beta'\beta\alpha_2$) which is then supplemented by another subunit, known as a sigma factor, to form the polymerase holoenzyme and it is the sigma factor that is thought to interact with, and recognise the -35 and -10 sequences. Until recently it was believed that the primary RNA polymerase holoenzyme, which incorporated a sigma factor termed σ^{70}, was responsible for all the transcription in *E. coli*. This view was changed with the report that the *htpR* gene, required for the increased expression of a class of genes in response to heat shock, functioned in combination with the core enzyme to direct transcription from heat shock promoters (Grossman *et al.*, 1984). The product of the *htpR* gene, which has been renamed *rpoH*, co-purifies with and binds quite tightly to the core enzyme and is assumed to be a sigma factor. It has been reported that the *ntrA* gene (also called *glnF*) which is required for the expression of a set of genes involved in nitrogen fixation and assimilation, is also a novel σ factor (Hirschman *et al.*, 1985; Hunt and Magasanik, 1985). Furthermore, the *E. coli* phage T4 has also been shown to utilise a new σ subunit for transcription of phage late genes. The gene 55 protein, in combination with core enzyme, can direct

transcription *in vitro* from late promoters (but not the holoenzyme $E\sigma^{70}$) (Kassevetis and Geiduschek, 1984). It has been observed that both the nitrogen-regulated and the heat shock promoters contain very little homology to the promoter consensus sequence around the -10 region (Ausubel, 1984; Cowing *et al.*, 1985) (Table 1.2) and it is presumably the ability of minor sigma factors to recognise these sequences that allows the specificity of expression.

Table 1.2 Promoter regions of genes of the *E. coli* heat shock system

Gene or consensus	-35 region	13-15 bp region	-10 region	+1
groE	TTTCCCCCTTGAA	GGGGCGAAGCCAT	CCCCATTTCTCTGGTCAC	
dnaK P1	TCTCCCCCTTGAT	GACGTGGTTTACGA	CCCCATTTAGTAGTCAA	
dnaK P2	TTGGGCAGTTGAA	ACCAGACGTTTCG	CCCCTATTACAGACTCAC	
htpG P1	GCTCTCGCTTGAA	ATTATTCTCCCTTGT	CCCCATCTCTCCCACATC	
rpoD P_{hs}	TGCCACCCTTGAA	AAACTGTCGATGTGG	GACGATATAGCAGATAA	
Consensus	T tC CcCTTGAA	13-15 bp	CCCCATtTA	
E. coli	tcTTGACa	16-18 bp	TAtAaT	cat

The observation that sigma factors mediate temporal gene expression (in the case of T4 late genes) or response to environmental stress (heat shock and nitrogen fixation) in enteric bacteria suggests a method for regulating gene expression in more complex prokaryotes, whose responses are more dramatic.

The first bacterium to be examined for such phenomena was *Bacillus subtilis*. Like *E. coli*, the principal form of the complete holoenzyme in *B. subtilis* recognises the typical prokaryotic promoter. The sigma factor from *Bacillus* however, has a molecular weight of 55 kDa (unlike the 70 kDa protein of *E. coli*) and was called σ^{55}. Alternative sigma factors have been found in *B. subtilis* and no less

than nine different polypeptides are thought to interact with the core enzyme to confer a variety of promoter recognition specificities. The appearance of these factors is associated with the onset of endospore or of phage development.

As with the coliphages, the *Bacillus* phages utilise different sigma factors to control development and the orderly expression of genes that leads to the formation of virions by the lytic phage SPO1 can be divided into those of early, middle and late genes. The early genes are recognised by $E\sigma^{55}$, including gene 28, the product of which turns on the middle genes. In turn, two products of middle genes (33 and 34) turn on the late genes. It was subsequently shown that SPO1 genes 28, 33 and 34 encode sigma factors required to allow RNA polymerase to recognise a different class of promoters (Talkington and Pero, 1979). The gene 28 protein (σ^{gp28}) acts alone to recognise the promoters of the middle class genes, whereas the gene 33 and 34 proteins ($\sigma^{gp33\text{-}34}$) act synergistically to direct transcription of late genes.

Unlike *E. coli*, *B. subtilis* exhibits programmed gene expression which results in endospore formation. *B. subtilis* does this under conditions of environmental stress which triggers a developmental response. The pattern of gene expression in this bacterium changes in response to nitrogen, phosphorus and carbon in a transition from vegetative growth to stationary phase. In addition to sporulation, these changes include the synthesis of flagella and the development of motility, the acquisition of the state of competence for DNA uptake, as well as the secretion of proteases and antibiotics. Studies *in vitro* show

that cloned sporulation genes are transcribed by a variety of holoenzymes containing minor sigma factors. Two of these genes, *spoVG* and *spoVC*, are recognised by two holoenzymes $E\sigma^{37}$ and $E\sigma^{32}$ but not by the polymerase associated with vegetative growth $E\sigma^{55}$. About an hour into development σ^{55} and σ^{37} are replaced on the RNA polymerase by a further transcriptional modifier known as σ^{29}. Expression of σ^{29} is itself a sporulation specific event controlled by the *spoO* loci, a set of vegetatively expressed genes which induce the expression of sporulation early genes. The protein σ^{29} continues to direct transcription from *spoVG* and *spoVC* as well as at least one other gene whose promoter is not recognised by σ^{55} or σ^{37}. If the temporal programme of sporulation gene expression is controlled largely by an SPO1-like cascade of sigma factors, then cloned genes that are turned on late in development should provide templates for detecting additional species of sigma factor. Indeed, the use of specific templates *in vivo* and *in vitro* has revealed the more subtle aspects consistent with such a model. A gene, *ctc*, shown recently to be discrete from but adjacent to *spoVC* (Igo and Losick, 1986) is transcribed by $E\sigma^{37}$ and $E\sigma^{32}$ and the sites on the *ctc* promoter to which these proteins bind have been well characterised (Moran *et al.*, 1981; 1982; Tatti and Moran, 1984; Hay and Moran, 1985).

Regulation of the *ctc* gene differs markedly from another gene that is also recognised by the $E\sigma^{37}$ and $E\sigma^{32}$, *spoVG*. Like *spoVG*, transcription of *ctc* is induced at the end of exponential phase of growth. However, the nutritional conditions that support *spoVG* induction are different to those that stimulate *ctc* transcription.

Expression of *ctc* is strongly induced in rich medium that contains high concentrations of glucose and glutamine and is relatively poorly transcribed in sporulation medium, whereas transcription of *spoVG* is induced under sporulation conditions and is poorly expressed in rich media containing glucose. The genetic requirements necessary for expression of the two genes also differ. Transcription of *spoVG* is dependent upon some of the genes from the *spoO* loci. In contrast, mutations in the *spoO* loci which abolish *spoVG* expression actually enhance transcription of the *ctc* gene (Zuber and Losick, 1983). Thus, in spite of the fact that both of these genes are transcribed by the same minor form of RNA polymerase holoenzyme *in vitro*, and are both induced at the end of exponential growth, their modes of regulation are different. Most models proposed to account for those phenomena involve the interaction of further auxiliary regulatory proteins and several candidates have already been earmarked. Whether any of these proteins turn out to be yet more sigma factors remains to be seen.

The control of gene expression exerted via alterations in the specific response of RNA polymerase is obviously employed in *Bacillus subtilis* but in addition a further level of control is involved which makes the whole cascade of events very complex.

When *Streptomyces* RNA polymerase was analysed, it was shown to resemble those of other prokaryotes in having a core structure $\alpha_2\beta\beta'$(Chater and Cooper, 1975; Jones, 1979; Buttner and Brown, 1985). It was also shown to recognise and utilise promoters

from *E. coli*, *Serratia marcesans* and *Bacillus subtilis* which conformed to the consensus *in vivo* in *S. lividans* (Bibb and Cohen, 1982). In more detail, mutations in the *ampC* promoter referred to earlier that alter the level of transcription in *E. coli* have a comparable (although lesser) effect in *S. lividans* (Jaurin and Cohen, 1984). It was not surprising then, when an RNA polymerase fraction from *S. coelicolor* was isolated and shown to utilise the *veg* promoter (a *Bacillus* promoter that conforms with the consensus) in transcription assays *in vitro* (Westpheling *et al.*, 1985). Evidence to support this came from the observation that *E. coli* plasmids were able to direct coupled transcription-translation in cell-free extracts of *S. lividans* (Thompson *et al.*, 1984). The picture was confused however when the *rpoD* gene from *E. coli*, coding for the major sigma factor σ^{70}, was used to design a probe in hybridisation studies using *Streptomyces* DNA (Takashaki *et al.*, 1988). More specifically the probe was an oligonucleotide whose sequence reflects the amino acid sequence which is thought to be involved in promoter recognition (Gitt *et al.*, 1985; Landlick *et al.*, 1984; Stragler *et al.*, 1985). In Southern analysis this probe found single signals in the genomes of *E. coli*, *B. subtilis* and *Staphylococcus aureus* but multiple signals with *S. coelicolor* A3(2). Furthermore, the corresponding fragments were cloned and sequenced and are thought to be homologues of *rpoD*. This raises the possibility that *S. coelicolor* has more than one holoenzyme that recognises the consensus promoter. This observation is supported by similar experiments using a probe derived from *Myxococcus xanthus*, whose genome has a 65% G+C content and therefore more compatible with that from *Streptomyces*. The probe consisted of DNA coding for a putative sigma factor, itself

isolated by virtue of its ability to hybridise with *rpoD*. Using the *Myxococcus* probe then, hybridising clones were isolated from a gene library of *S. coelicolor* A3(2). One such clone has been partially sequenced and current data suggest that the DNA codes for an *rpoD*-like protein since the degree of similarity between its amino acid sequences and the corresponding genes from *M. xanthus, E. coli* and *B. subtilis* is so high. Thus there is strong evidence for the existance, in *Streptomyces*, of an RNA polymerase holoenzyme (or multiple holoenzymes) that recognises the consensus prokaryotic promoter.

When fragments of *S. lividans* chromosomal DNA were cloned in an *E. coli* expression vector, some were found to have the ability to direct transcription (Jaurin and Cohen, 1985). Genomic DNA was digested with *Sau*3A and small fragments (120-300 bp) were cloned in an expression vector. This promoter probe vector utilised the *ampC* gene of *E. coli* as an indicator of the activity of the *S. lividans* derived promoters that had been introduced upstream. Production of the β-lactamase encoded by the *ampC* gene can be detected by colorimetric assay *in vitro* or by ampicillin resistance *in vivo*. There is a linear correlation between β-lactamase production and ampicillin resistance and this marker had proven useful in the isolation and characterisation of the "up-mutations" in transcriptional regulatory signals referred to earlier (Jaurin *et al.*, 1981; 1982; Jaurin and Normark, 1983). When the promoter sequences were identified they were, not surprisingly, similar to the *E. coli* canonical sequence. However, the DNA sequences revealed a number of direct repeats of unknown function the like of which, although not normally part of *E. coli* sequences, have been

associated with regulation, for example in bacteriophage λ, where they are recognised by the activator protein cII. Significantly, these A+T rich *Streptomyces-E. coli*-like promoters (SEP) were less active in *Streptomyces* than in their heterospecific host. For example, one SEP, which was five times as active as the wild type *ampC* promoter in *E. coli*, registered an activity of only 2.5% of that associated with the wild-type *ampC* in *S. lividans*. It is also interesting to note that, if the number of promoters isolated in that study is representative, these SEP's comprise only about 10% of the promoters in *S. lividans*. Furthermore, this calculation was based on the assumption that the total number of promoters in *S. lividans* was equal to that in *E. coli*, a figure which is almost certainly an underestimate. If however, *Streptomyces* contain multiple sigma factors in order to recognise promoters of vegetative genes which perhaps are subtly distinct from their *E. coli* counterparts (having different spacer lengths for example) then more such SEP may exist which although apparently conforming to the consensus sequence, are not utilised by *E. coli*. Examination of some of the promoters isolated from *Streptomyces* reveal that they do resemble the *E. coli* promoters in having relatively A+T rich -35 and -10 regions, but consistent with the previous argument, many are not utilised in *E. coli*. On the other hand, many of the *Streptomyces* promoters do not possess these characteristic sequences. Some, like the promoters for the thiostrepton resistance methylase gene from *S. azureus* (*tsr*) and those for the *S. fradiae* aminoglycoside phosphotransferase gene (*aph*), have a sequence at -10 that approximates the consensus but no corresponding sequence at -35 (Table 1.3), while other gene promoters, like that of *endo H* from

Table 1.3. Promoter regions of some streptomycete genes.

Promoter	-45 region to -35 region		-10 region		mRNA start
E. coli [a]	TTGACa	-17 bp-	TAtAaT		
*gyl*P1	TTGACG	-17 bp-	GAGACT	-8 bp-	U
*gyl*P2	TCGAAC	-19 bp-	TAGAGT	-6 bp-	A
ermE P1	TGGACA	-14 bp-	TAGGAT	-6 bp-	G
ermE P2	TTGACG	-18 bp-	GAGGAT	-6 bp-	C
Streptomyces [a]	TTGaca	-18 bp-	tAGgaT		
aph P1	CGAAAGGCGCGGAACGGCGT	-12 bp-	CATGAT	-8 bp-	A
aph P2	CGGTGGGGGATTCCGGCCGA	-12 bp-	CCATGT	-8 bp-	U
tsr P1	TTGCCGGTCAGGGCAGCCAT	-14 bp-	TAGGGT	-6 bp-	A
tsr P2	GCTCGACGCAGCCCAGAAAT	-14 bp-	AATACT	-6 bp-	U
hyg	AACGTCCCCGACGTGGCCGA	- 6 bp-	CATCGT	-12 bp-	G
pac	CGCGCCCCCGCGGCCGCCAC	- 8 bp-	CACCAT	-9 bp-	C
endoH	ATTGCATGATTGA	- 7 bp-	CCGGCGGGCAGGG	-7 bp-	G
vph P1	GCAGCGCCGTGTGCGGCCTG	- 7 bp-	CCGGCGGGAGCGA	-6 bp-	U

[a] consensus sequences

S. plicatus (a gene encoding an endoglycosidase) possess neither a -35 or -10 canonical sequence (Table 1.3). Moreover, most *Streptomyces* genes appear to have multiple promoters, a feature that occurs only rarely in other prokaryotes. It has already been shown in *B. subtilis* that "alternative" sequences exist which are recognised by minor forms of RNA polymerase and it was research along these lines, that began to resolve some of the questions about how these genes might be transcribed and ultimately, how their expression could be regulated.

It has already been noted that an RNA polymerase fraction, isolated from *S. coelicolor* was able to utilise the *veg* promoter of *B. subtilis*. In addition, in the same study, a second polymerase fraction able to utilise the *ctc* promoter but not the *veg* from *B. subtilis* was detected (Westpheling *et al.*, 1985). This minor form of holoenzyme also could recognise the endoglycosidase H gene promoter isolated from *S. plicatus*, and this was the first indication that *S. coelicolor* might have numerous sigma factors with a possible role in regulation of differentiation and secondary metabolism. It was the detailed study of another gene which really showed how complex this system might become. *S. coelicolor* produces an extracellular agarase (Hodgson and Chater, 1981), the first enzyme in a catabolic pathway, that permits the organism to utilise agar as a sole carbon source. The expression of the gene encoding this enzyme (*dagA*) is induced by the products of agar degradation and is glucose repressible. The *dagA* gene has been cloned (Kendall and Cullum, 1984; Bibb *et al.*, 1987) and the transcription pattern defined (Buttner *et al.*, 1987). High resolution S1 mapping revealed that no less than four promoters were responsible for

directing transcription of that gene. Furthermore, these promoters showed marked heterogeneity and were eventually shown to be recognised by at least three different holoenzymes (Buttner *et al.*, 1988). One of the promoters, *dagA*P4 (Fig. 1.7) resembles the prokaryotic consensus, and is recognised by a sigma factor, (designated σ^{35} by virtue of its molecular weight) that had been previously described as having recognised the *Bacillus veg* promoter. Another holoenzyme that was purified, was shown to direct transcription not only from one of the *dagA* promoters (*dagA*P3 in fact) but also from the *ctc* promoter. This sigma factor, designated σ^{52}, is thought to be equivalent or identical to that called σ^{49} by other workers. The holoenzyme that utilised *dagA*P2 incorporated a sigma factor (σ^{28}) which did not obviously recognise any of the other *dagA* promoters. Transcribing activity corresponding to the last promoter (*dagA*P1) was labile and was not isolated but could well be due to yet another minor holoenzyme form. These numerous promoters then could be used to explain the response of *S. coelicolor* to induction and repression of the agarase gene. However, other *Streptomyces* determinants, namely the α-amylases of *Streptomyces limosus* (Long *et al.*, 1987; Virolle and Bibb, 1988) and *Streptomyces venezuelae* (Virolle *et al.*, 1988), that are also subject to glucose repression, are transcribed from only single promoters with similar sequences to *dagA*P4 (a consensus-like sequence). However, differential transcription of *dagA* was observed when grown in different media, revealing that promoters *dagA*P1 and *dagA*P4 were induced in the presence of agar and repressed by glucose. The levels of the *dagA*P2 and *dagA*P3 transcripts in exponential growing cells was low under either set of conditions. In

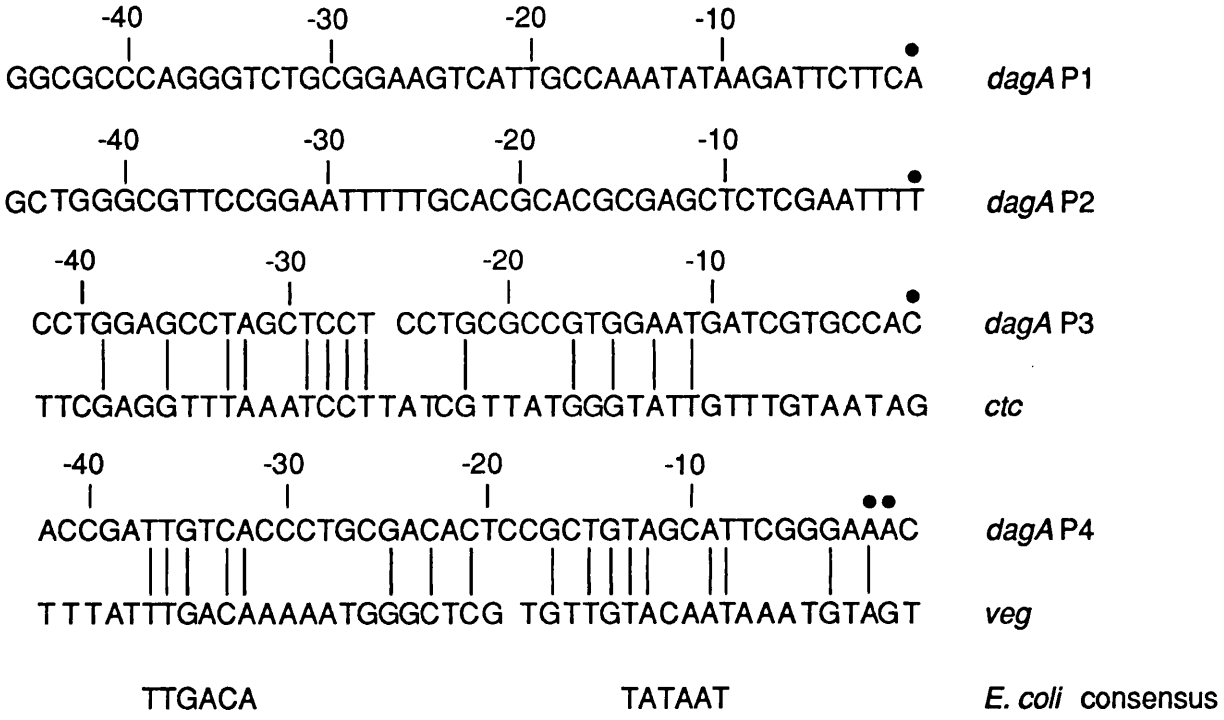


Fig. 1.7 Nucleotide sequences of the four promoters directing transcription of *dagA*.

No less than four promoters direct transcription of the extracellular agarase gene of *S. coelicolor*. The relevent sequences are compared with the prokaryotic consensus or the *Bacillus ctc* promoter

contrast, when cells were grown to stationary phase in a fermenter, enhanced levels of transcription were observed from *dagA*P2 and *dagA*P3.

However, expression of the agarase gene from *S. coelicolor* is not the only example of catabolic repression or temporal regulation exerted at the level of transcription. The small *gal* operon from *S. lividans* contains 3 genes in the order *galT* (galactose-1-phosphate uridyl transferase), *galE* (UDP-4-epimerase) and *galK* (galactokinase) (Fig.1.8) (Fornwald *et al.*, 1987). The operon is transcribed from two promoters, one of which (*gal*P1) is upstream of *galT*, the other (*gal*P2) being in the intergenic space between *galT* and *galE*. The latter promoter is constitutively expressed and resembles a "standard" *E. coli*-like signal whereas the former is glucose repressed and induced in the presence of galactose. The hypothetical reason for constitutive expression of *galE* and *galK* is that they are required for biochemical processes other than galactose utilisation. In *E. coli*, at least, UDP epimerase has a role in cell wall biosynthesis by generating UDP galactose in the absence of galactose. A second role of galactose kinase is less easily explained. It is possible that the true inducer of the *gal* operon is not galactose but galactose phosphate, in which case *galK* expression would be a prerequisite for galactose induction and evidence exists to support this hypothesis (Brawner *et al.*, 1988). The more interesting of the two *gal* promoters, that which is inducible by galactose and repressed by glucose, has been analysed in some depth. Inspection of the sequence immediately identifies -10 and -35 sequences, but these are separated by 21 base pairs. Precedents exist for such promoter structure in that the tightly regulated temporally

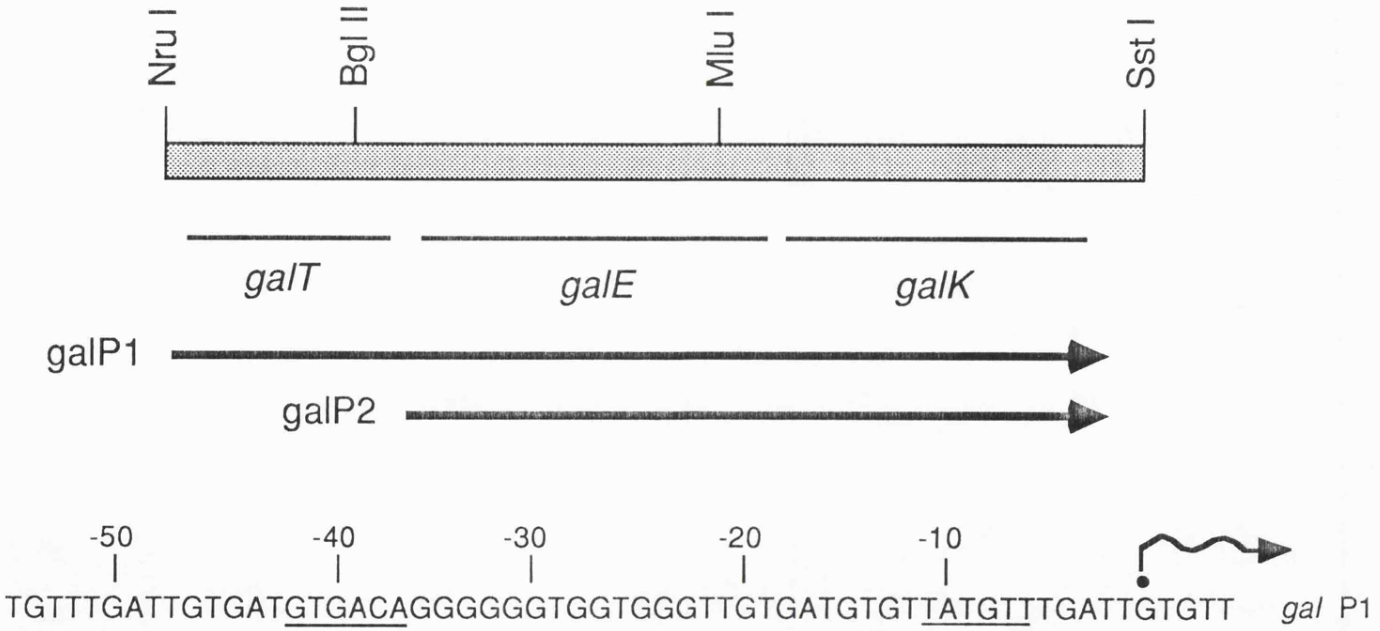


Fig. 1.8 Organisation of the *S.lividans gal* operon.

The direction and extent of transcripts is shown along with the positions of the three genes that comprise the operon. The sequence of *gal* P1 is shown and the candidates for the -10 and -35 sequences are underlined. They are separated by a larger than usual distance of 21 or 24 bp.

expressed *spoIIE* (Guzman *et al.*, 1988) and *spoIIG* (Kenney and Moran, 1987) possess sequences with near perfect identity with the consensus -35 and -10 regions but have an uncharacteristic spacing of 21 bp and 22 bp respectively. Although *gal*P2 seems to be recognised by polymerase holoenzyme $E\sigma^{28}$, it is not clear which holoenzyme may recognise the *gal*P1 promoter (Westpheling and Brawner, 1989). The sequence from +1 to -55 contains a number of direct repeats and, while the function of these repeats is unknown, they have been investigated using site directed mutagenesis and although not surprisingly most mutations have some effect on the level of transcription, some of these appear to affect the induction/repression of the system by galactose or glucose.

It is clear, therefore, that catabolite control can be mediated via transcriptional signals and, as we have seen in the case of *dagA*, the transcription pattern can change during development.

Morphological development and the onset of secondary metabolism, the period during which antibiotics are made, is of particular interest, and the isolation of antibiotic biosynthetic pathways has allowed detailed investigation of control mechanisms, for example in the case of streptomycin biosynthetic pathway. Streptomycin is an aminocyclitol antibiotic produced by, among others, *Streptomyces griseus*. To date, 12 *str* genes have been identified, clustered around a streptomycin-6-phosphotransferase designated *aphD* or *strA*. *S. griseus* contains a further phosphotransferase which modifies streptomycin in the 3" position (*aphE*) but otherwise the gene cluster is almost identical

to that in *S. glaucescens*, another streptomycin producer. It was thought, until recently that all the functionally repetitious steps in the pathway were performed by the same enzyme (Piepersberg *et al.*, 1988). Since this assumption is probably no longer true, as many as 25 to 30 genes may be involved in this one biosynthetic pathway. Even as it stands, the protein products encoded by the presumed *str* genes are mostly questionable or completely unknown. However, a mechanism of regulation of part of the pathway has been proposed. Although general control is thought to be exerted by nutritional factors such as phosphate (Martin and Demain, 1980) and by A-factor (Khokhlov, 1982), regulatory phenomena specific for the *str* genes were identified using mutants and expression of genes in *S. lividans*. Studies show that *strB1* (a putative amidinotransferase) is only expressed in the presence of a functional *strR* gene (Distler *et al.*, 1987). However, transcription from the *strB1* promoter occurs even in the absence of the *strR* protein, implying that *strR* does not code for a sigma factor. The action of *strR* suggests that the protein may be an antitermination factor, and a model has been proposed in which regulation of *strB1* expression is controlled by antitermination by the product of *strR* (Fig. 1.9).

Another operon in which antitermination in thought to play a part is that involved in glycerol utilisation. The *S. coelicolor gyl* operon has been cloned and the transcripts mapped (Seno *et al.*, 1984; Smith and Chater, 1987; 1988). The operon consists of two genes, *gylA* which codes for a presumptive glycerol kinase and *gylB* encoding an sn-glycerol-3-phosphate dehydrogenase, although one transcript goes beyond *gylB* by over 1 kb and may contain another non-essential

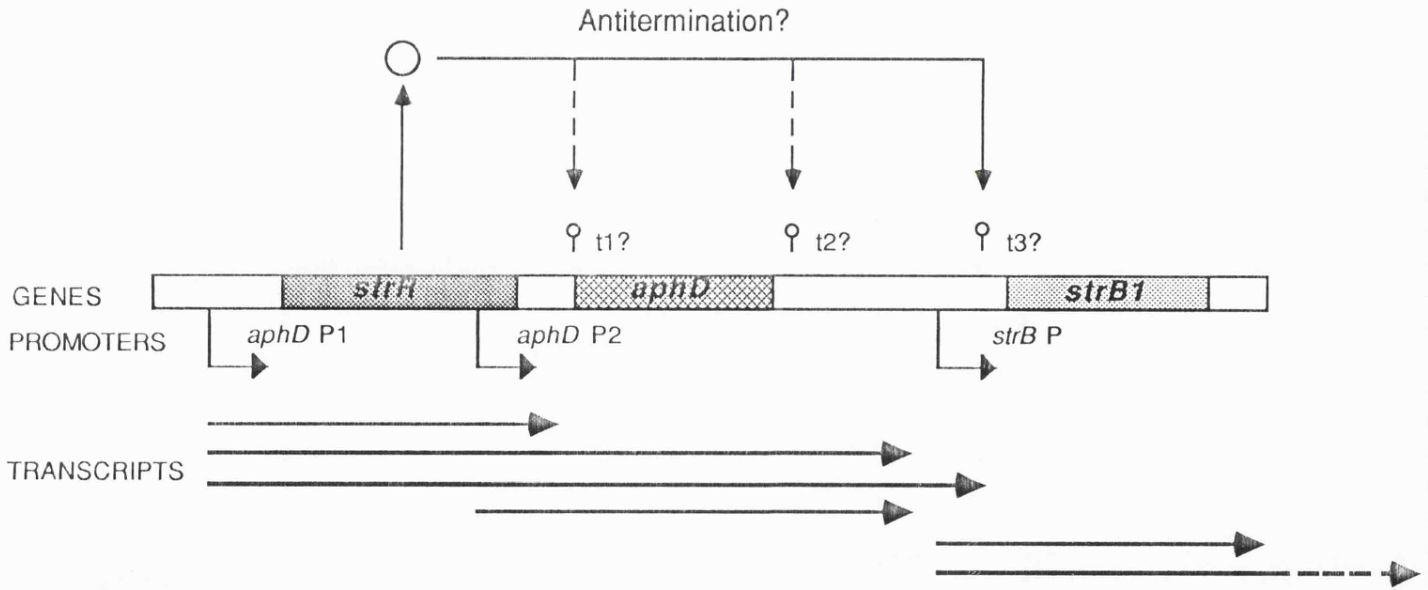


Fig. 1.9 A model for the mechanism regulation of part of the streptomycin biosynthetic gene cluster.

The promoters indicated have been identified by S1 mapping, but the mechanism of regulation, by antitermination, shown here is hypothetical and based on sequence data.

gene (*gylX*). Another 0.9 kb transcription unit is located immediately upstream of *gylABX* and is thought to code for a regulating factor (Fig. 1.10). The operon is inducible by glycerol and glucose-repressible, whereas the regulatory gene is only glycerol inducible. Transcript analysis revealed that while the operon was transcribed from a single promoter, the transcripts terminated at 2 or 3 different positions. The product of the 0.9 kb gene then is thought to act as an antiterminator, thereby activating genes downstream in the operon. In the examples discussed thus far, regulation of gene expression in primary and secondary metabolism has been controlled by recognition of a promoter by a minor sigma factor as in *Bacillus* or by the antitermination mechanism described here.

Regulation of nitrogen metabolism in *Streptomyces* on the other hand seems to resemble more the Gram-negative bacteria like *E. coli* than the Gram-positives like *Bacillus*. Apparently, regulation in *Streptomyces* is mediated via glutamine synthase and glutamate dehydrogenase and, like *E. coli*, metabolic imbalance during sudden changes in nitrogen availability is avoided by inhibition of glutamine synthase by adenylylation (Reitzer and Magasanik, 1987; Streicher and Tyler, 1981).

The metabolism of *Streptomyces* is necessarily complex due to the change, in response to nutrient limitation, from primary to secondary metabolism which runs concurrently with morphological development. Some examples of possible control and co-ordination of secondary metabolism have been discussed, but the question of how

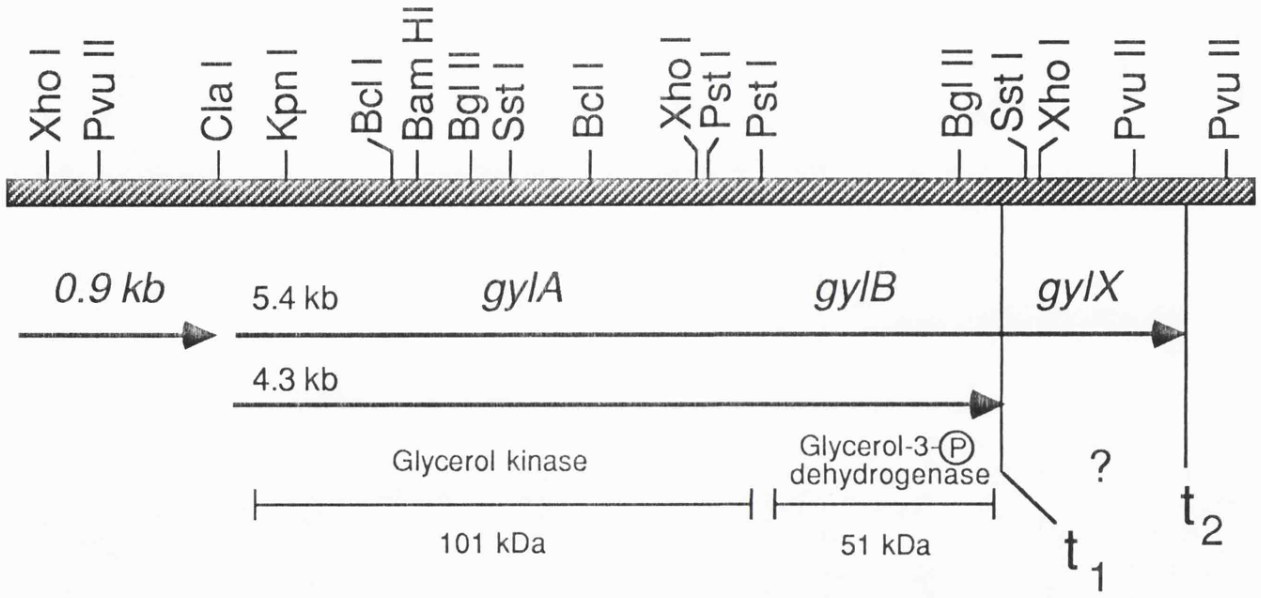


Fig. 1.10 Organisation of the *S. coelicolor gyl* operon

The transcription pattern is indicated which suggests that although the operon is transcribed from a single promoter these transcripts terminate at more than one site.

this is linked to morphological development remains. The genes involved in aerial hyphae and spore formation appear to be scattered around the chromosome, but nevertheless many have been identified and mutations have been mapped in genes responsible for aerial hyphae formation (*bld* for "bald" mutants) as well as those for sporulation (*whi* for white).

Mutations in at least seven *bld* genes have been identified in *S. coelicolor* A3(2) (Merrick, 1976; Puglia *et al.*, 1986; Champness, 1988), but the most well defined is *bldA*. These mutants grow vigorously as substrate mycelium but instead of producing aerial hyphae the surface layer becomes fragmented and produces aberrant cell walls. However, growing on mannitol as carbon source, this mutant apparently develops normally, going on to form spores. Despite this, and irrespective of growth conditions, *S. coelicolor* carrying a mutation in *bldA* produces none of the antibiotics normally synthesised by this organism. The DNA sequence of *bldA* was at first confusing as it did not appear to code for a protein. This was indeed the case as *bldA* almost certainly codes for the leucyl-tRNA that recognises the rare codon (at least in *Streptomyces*) TTA (Lawlor *et al.*, 1987). Apart from possessing all of the characteristic features of a tRNA, *bldA* has 73% identity with $tRNA^{Leu}$ (TTA/G) from *E. coli* (Yamaizumi *et al.*, 1980) and is 66% similar to the equivalent tRNA from *B. subtilis* (Wawrousek *et al.*, 1984). It must mean that few, if any genes involved in vegetative growth contain TTA codons although this might not be surprising in an organism whose genome is 73% G+C. In fact only four DNA sequences of *Streptomyces* genes available so far contain TTA codons and all of

these are involved in secondary metabolism. The *carB* gene of *S. thermotolerans* (the carbomycin producer) contains two TTA codons (Epp *et al.*, 1987) and is not expressed at all in a *bldA*$^-$ strain. The partial expression of other genes in a *bldA* mutant has resulted in the proposition of a model by which such an anomalous effect could exert a regulatory function. Two of these "partially expressed genes" are aminoglycoside modifying enzymes, namely the hygromycin B phosphotransferase (*hyg*) from *S. hygroscopicus* (Zalacain *et al.*, 1986) and the streptomycin phosphotransferase (*sph*) from *S. glaucesans* (Vögtli and Hütter, 1987). Since the TTA codons are close to the N-terminal end of these genes, the model proposes that, in *bldA*$^-$ strains translation initiates at another site, further into the coding sequence (Fig. 1.11). This requires the existence of a ribosomal binding site and initiation codon downstream of the TTA codon, a feature which both these genes appear to possess. In addition, the proteins would need to be active as phosphotransferases in the absence of the N-terminal portion and this is supported by the observation that the first 39 amino acid residues of the *sph* protein are not included in an extensive sequence similarity to the amino-terminus of the Tn5 encoded streptomycin phosphotransferase. The hypothesis is an attractive one as it allows expression of the inactivating enzymes before aerial hyphae formation (i.e. *bldA* expression) and antibiotic biosynthesis, thereby rendering the cells resistant to their potentially lethal product, and goes further to suggest that the "mature" product following *bldA* expression could, with now additional properties, activate antibiotic biosynthesis. The model is appealing in that it links possible primary and secondary metabolism but has one caveat

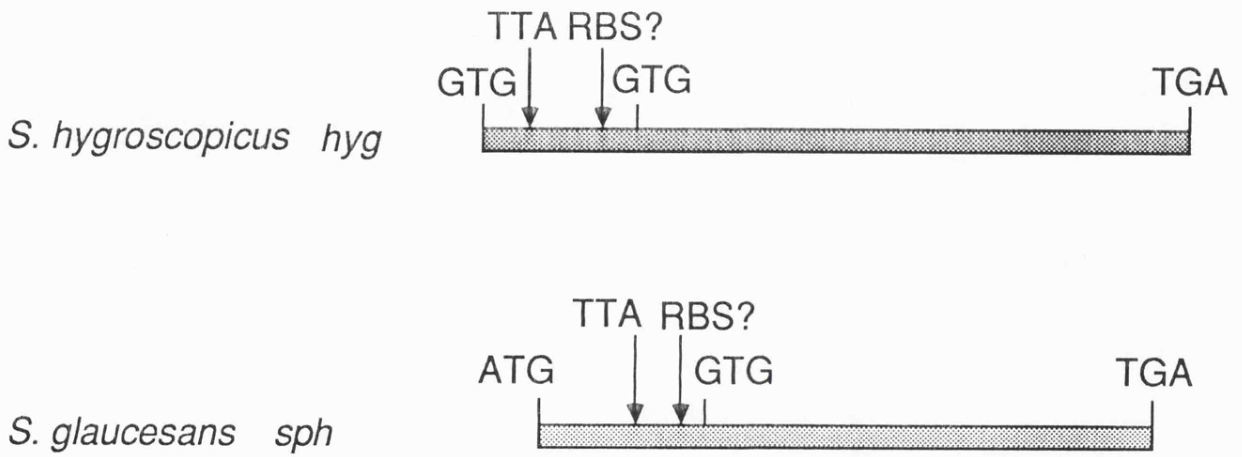


Fig. 1.11 The structure of two aminoglycoside resistance determinants

It has been proposed that, since alternative translational start sites are available in these two genes, that they are utilised during early growth prior to the expression of *bldA*. Clearly, the two genes cannot be expressed from the first iniation codon before *bldA* as it encodes the tRNA that recognises TTA.

concerning the nature of the two genes involved. The proposal is based on expression of the cloned *hyg* or *sph* genes from alternative start codons in a *bldA* mutant of *S. coelicolor* A3(2) resulting in an enzyme whose catalytic activity is not as great as that of the mature protein. This is assayed by the ability of such a strain to grow in the presence of exogenous hygromycin or streptomycin. While these antibiotics inhibit protein synthesis they also cause mistranslation of messenger RNA in prokaryotic systems (Davies *et al.*, 1965). Expression of these genes in a *bldA* mutant under such conditions could be simply the result of misreading of the TTA codon. This possibility could be tested by growing a *bldA* mutant containing *carB* in the presence of non lethal concentrations of aminoglycoside antibiotics and assaying for an increase in carbomycin resistance.

Genes involved in sporulation have also been isolated subsequent to the generation of mutants. These so-called "white" mutants put up aerial hyphae which fail to septate and form spores. So far, nine different genes have been implicated in sporulation (Chater, 1972) and of these only three have been cloned (Chater *et al.*, 1988). The *whiG* gene has been sequenced and encodes a putative polypeptide of about 31 kDa. A search of computer databases suggested that *whiG* might code for a sigma factor. In fact expression of this gene on a high copy number plasmid results in earlier sporulation suggesting that the gene product is a rate limiting step in spore formation (Méndez and Chater, 1987).

The mechanisms discussed so far, have rather global

consequences as far as the cell is concerned. However, a novel mechanism yet to be discussed in detail is the possible involvement of antibiotics in regulating gene expression. The discovery of thiostrepton-inducible promoters has opened up a whole new role for antibiotics. Thiostrepton, among other inhibitory actions, affects the ribosome associated (p)ppGpp synthase and thus interferes with stringent response. The stringent response is an integral part of the switches that initiate secondary metabolism since induction of a relaxed phenotype is usually concomitant with loss of antibiotic production (Ochi, 1988). Of perhaps more significance, was the observation that *S. lividans* containing pIJ702 produced additional proteins in the presence of thiostrepton. Subsequently, a gene encoding one of these proteins was cloned by "reverse genetics". The promoter of this gene (*tipA*) was shown to be truly thiostrepton inducible by subcloning it in the promoter probe vector pIJ486 (Murakami *et al.*, 1989). This vector contains a promoterless aminoglycoside phosphotransferase gene derived from Tn5 as the indicator. It was then possible to illustrate thiostrepton-inducible neomycin resistance. The antibiotic thiostrepton has already been shown to bind to a defined site on an oligoribonucleotide; more specifically the antibiotic binds to 23S rRNA at or around residue A-1067 (Thompson *et al.*, 1982; 1988; Beauclerk *et al.*, 1985; E. Cundliffe, personal communication). By analogy this raises the possibility that thiostrepton could interact directly with the DNA of the thiostrepton inducible promoter. The question remains, however, why *S. lividans*, an organism that should never encounter thiostrepton (other than by hand of a molecular biologist) should have genes whose expression is influenced by that

antibiotic, but having made this observation such experiments might now be more meaningfully performed using the thiostrepton producer *S. azureus*. It would then be interesting to identify the nature of the proteins whose expression is induced by thiostrepton in this organism.

Another antibiotic that may regulate gene expression at the level of transcription is novobiocin. This antibiotic inhibits DNA gyrase by binding to the B subunit of the tetrameric (A_2B_2) enzyme, and as a consequence alters the level of supercoiling of DNA in the cell. The role of DNA gyrase in the cell is to maintain a certain level of supercoiling in conjunction with another enzyme, topoisomerase I. The antagonistic effect of these two enzymes maintains a homeostatic control over the DNA topology, vital for the expression of some genes. Indeed, the effect of supercoiling on gene expression is well documented in *E. coli* and *Bacillus* (Gellert, 1981; Wang, 1985). Resistance to novobiocin in the producer, *S. sphaeroides*, is due to an altered *gyrB* gene product (Thiara and Cundliffe, 1988), which presumably fails to bind the antibiotic. Interestingly, the *gyrB* gene of *S. sphaeroides* is inducible by novobiocin, and subsequent isolation of the promoter indicated that regulation was at the level of transcription. A 650 bp fragment of DNA from the *gyrB* gene was cloned in pIJ487 to give novobiocin-inducible neomycin resistance (A. Thiara, personal communication). It remained to be established whether novobiocin bound directly to the DNA, to some activator protein or if some other mechanism was operating, questions which were answered, in part, by the use of another drug, ciprofloxacin. While novobiocin binds to, and inhibits energy transduction by the B subunit of gyrase, the fluoroquinolone group of

antibiotics (of which ciprofloxacin is a member) act by binding to the A subunit. It turns out that ciprofloxacin (a chemically very different compound to novobiocin) is also an inducer of the *gyrB* promoter(s) and it would appear that the effect is mediated via DNA topology (i.e. it is the level of supercoiling that affects the level of transcription directed by the *gyrB* promoter(s)). This observation, along with the fact that *S. sphaeroides* possesses two *gyrB* genes, one of which encodes a novobiocin-sensitive product (Thiara and Cundliffe, manuscript in press), suggests an exquisite role for novobiocin in regulating gene expression in that organism. One might envisage a set of conditions in which a transient change in the level of supercoiling could switch on and off a whole subset of genes. During vegetative growth, the superhelical density might be maintained by the novobiocin-sensitive gyrase and topoisomerase I. However, when novobiocin biosynthesis began, the sensitive gyrase could be inhibited and the superhelical density would drop and we have already seen that this could cause some gene promoters to become active. At some point the *gyrB* encoding the novobiocin-resistant gyrase would be turned on, increasing the superhelical density of the genomic DNA and thereby switching off the genes previously activated by the inhibitory action of novobiocin. Is regulation of gene expression a hitherto unrecognised "function" of some antibiotics?

Given that actinomycetes are complex Gram-positive bacteria which exhibit temporal and spatial gene expression as well as elaborating a number of rather strange compounds, it is not surprising that the control of this expression not only reflects some of the

precedents set by other prokaryotes but has some facets that are apparently unique to this order.

2.2. Translation.

The influence of the high G+C content in the DNA of actinomycetes is immediately obvious when examining the translated sequence of a transcript. As a result of the degeneracy of the genetic code, nearly all amino acid codons are of the form XY(A/G) or XY(U/C), and therefore, although certain limitations are set upon base selection in the first and second positions, the limits for the third or "wobble" base are somewhat less restrictive. As a consequence, *Streptomyces* coding sequences reveal a codon bias which results in the G+C content in the third position of the codon being as high as 98%. The development of computer analyses which detect the non-random distribution of bases within coding sequences has allowed the prediction of genuine open reading frames within DNA sequences (Bibb *et al.*, 1984; Gribskov et al., 1984). The limited (but now quickly expanding) number of *Streptomyces* sequences has thrown out one or two anomalies with regard to the recognition and translation into protein of messenger RNA. For example, a number of streptomycete genes initiate not at AUG but at GUG as observed in the cases of the agarase gene (*dagA*) from *S. coelicolor* (Buttner *et al.*, 1987), the *S. hygroscopicus* hygromycin phosphotransferase gene (*hyg*) (Zalacain *et al.*, 1986) and the erythromycin resistance determinant (*ermE*) from *Saccharospolyspora erythraea* (formerly *Streptomyces erythraeus*) (Bibb *et al.*, 1985) as well as others (Adams *et al.*, 1988; Neal and Chater, 1987). However, there are examples of the use of alternatives

to AUG as an initiation signal in *E. coli* in which approximately 2.5% of its genes initiate using GUG. In *Bacillus* a number of cases are known where TTG is utilised as the first codon and a case has even been documented (in the eukaryotic Sendai virus) where ACG is used as the translation initiation signal (Curran and Kolakofsky,1988).

The level of expression of a gene can be increased or decreased by altering the rate of initiation of translation. The degree of complementarity between the sequence just upstream of the initiating codon and the 3' end of 16S rRNA of the small ribosomal subunit can affect the level of expression of a gene (Shine and Dalgarno, 1974; Warburton *et al.*, 1983; Min *et al.*, 1988). Indeed, some *E. coli* genes are not expressed in *B. subtilis* apparently because *Bacillus* requires a high level of complementarity between the mRNA and 16S rRNA. Like those of *E. coli*, ribosomes from *Streptomyces* do not demand a strong "Shine-Dalgarno" interaction allowing a number of heterospecific genes to be expressed in this host. However, some *Streptomyces* genes have been shown to have no apparent ribosome binding site at all and in more than one instance, the transcript of a *Streptomyces* gene has been shown to initiate at, or next to, the first base of the initiation codon (Janssen *et al.*, 1985). What it is that the ribosome recognises in these transcripts (since it must surely recognise something) remains unresolved.

Although few examples exist of regulation of gene expression in *Streptomyces* at the level of translation, one of note is that of *tlrA* (*ermSF*). This gene confers resistance to the MLS antibiotics

(macrolides, lincosamides and streptogramin B) and was isolated from *Streptomyces fradiae* (Birmingham *et al.*, 1986). The gene product is an rRNA methylase which modifies residue A-2058 of 23S rRNA in the same way as the *ermE* gene product (Zalacain and Cundliffe, manuscript submitted; Skinner *et al.*, 1983). This resistance gene belongs to a family of inducible MLS resistance determinants which include, *ermA* and *ermC* from *Staphylococcus aureus* (Murphy, 1985), *ermAM* which originates in *Streptococcus sanguis* (Horinuochi *et al.*, 1983), the *Bacillus licheniformis ermD* (Gryzcan *et al.*, 1984), and a resistance gene from *Bacillus sphaericus*, *ermG* (Monod *et al.*, 1987). Not only do all the products of these genes apparently have the same mode of action, they are all inducible, and evidence exists which has allowed a model of translational attenuation to be proposed. This model is based on structural evidence from sequence studies in which a complex system of leader mRNA secondary structures are involved. The *tlrA* mRNA leader sequence, like those for the other MLS resistance genes, appears to be capable of internally pairing to form stable secondary structures (Kamimiya and Weisblum, 1988) (Fig. 1.12). The model, as applied to *tlrA*, postulates that the leader sequence functions as a translational attenuator in which ribosomes translating the putative leader peptide, would stall in response to inducing antibiotics. This would result in an alternative structure being adopted in which the ribosome binding site for the methylase gene was no longer sequestered allowing synthesis of the enzyme. Thus, in the mechanism proposed, the expression of the *tlrA* gene is regulated by the availability of its ribosome binding site (Fig. 1.12).

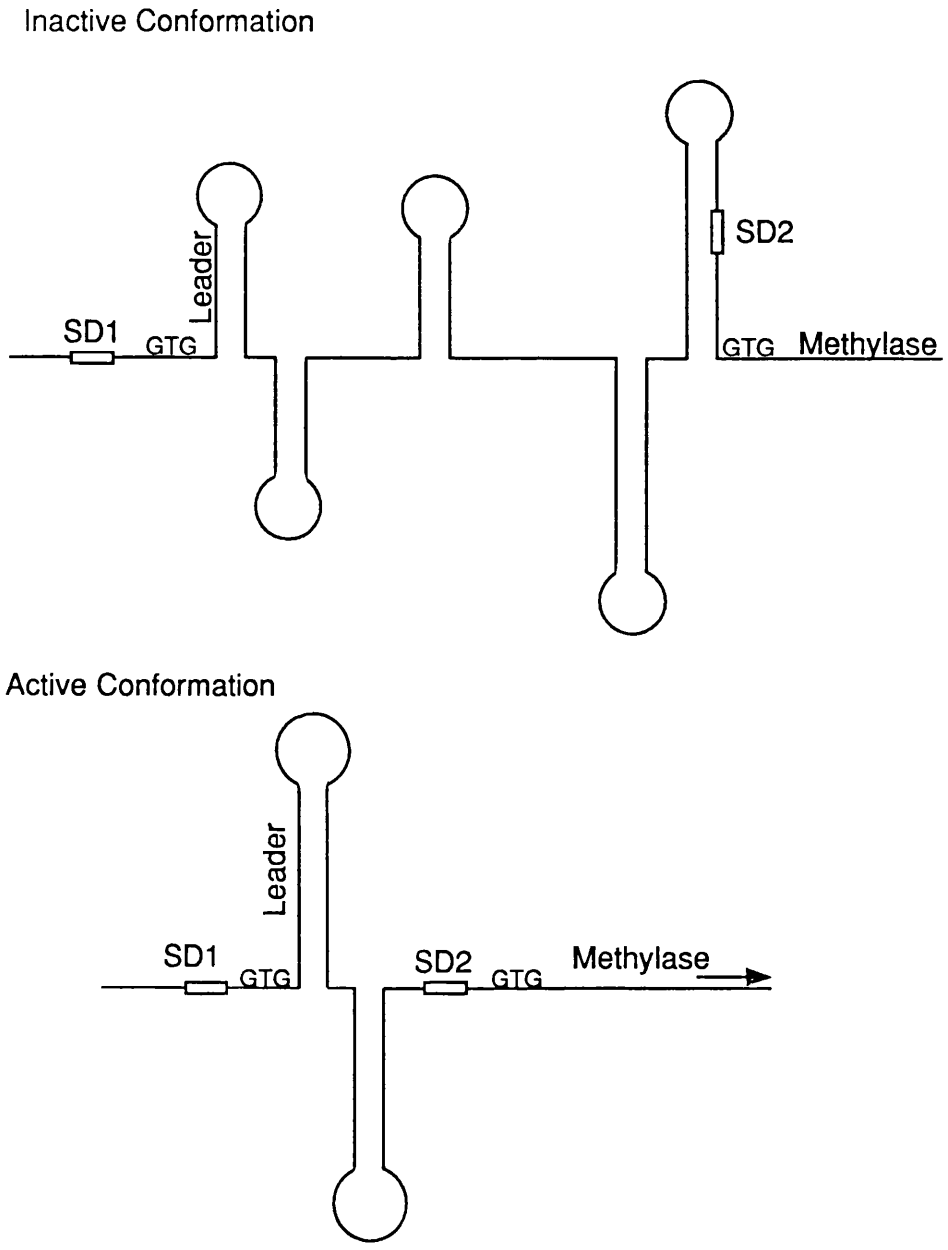


Fig 1.12 Model proposed for translational attenuation of *tlrA*

In the inactive conformation, the ribosome binding site of the methylase gene (SD2) is sequestered in a hairpin loop thereby preventing translation of that determinant. However, ribosomes binding to SD1 and translating the leader peptide can be stalled by erythromycin, allowing the mRNA to adopt an alternative structure (referred to as the active conformation) which releases SD2 such that ribosomes may then translate *tlrA*.

Genetic studies of *Streptomyces* have focussed on the secondary metabolism of these organisms and control of the expression of those genes appears to be (mostly) at the level of transcription. Whether examination of some more genes and operons involved in primary metabolism will produce further examples of translational regulation of expression remains to be seen.

3. Heterospecific gene expression.

The expression of genes in a foreign background is not always straightforward due to the many obstacles already discussed. Despite this, heterologous expression of a gene is often a desirable goal, whether because the natural host is difficult to exploit (e.g. slow growing or pathogenic) or because only a suitable heterospecific host is available for the purposes (e.g. an auxotrophic mutant). Examples of the former include two bacterial strains of clinical interest, *Mycobacterium tuberculosis* and *Mycobacterium leprae*, which have *in vivo* doubling times of 15 to 20 hours and 12 to 20 days respectively (indeed the latter species will not grow *in vitro*). Recently, considerable progress has been made in the expression of mycobacterial genes in *E. coli* using cosmid vectors, and the results suggested that these genes could, at best, be expressed only weakly from their own signals in that host (Clark-Curtiss *et al.*, 1985), a conclusion which was reinforced when *M. bovis* BCG DNA was cloned in a lambda vector (Thole *et al.*, 1985). By the use of expression vectors, clones were isolated which complemented *E. coli* auxotrophic mutations (Clark-Curtiss *et al.*, 1985) or which produced polypeptides which would react with monoclonal antibodies specific for *M. leprae* and *M. tuberculosis* (Thole *et al.*, 1985;

Young *et al.*, 1985a; 1985b). Obviously, the use of an alternative (non-pathogenic) host for expressing mycobacterial genes is desirable, and although non-pathogenic mycobacteria exist, cloning procedures have not been developed for this genus. Since mycobacteria belong to the order Actinomycetales, a suitable heterospecific host might be *Streptomyces*. Thus far, the ability of *S. lividans* to use the expression signals of genes from *Mycobacterium bovis* BCG has been tested *in vivo* by using gene fusions by inserting random DNA fragments from *M. bovis* BCG into promoter-probe plasmids in *E. coli* and in *S. lividans* (Kieser *et al.*, 1986). The results indicated that not only did *S. lividans* utilise a high proportion of mycobacterial promoters, but that the *Streptomyces* host recognised *M. bovis* BCG translational signals almost as well as its own. Moreover, several hybrid proteins with amino termini derived from *M. bovis* BCG seemed to be stable in *S. lividans*.

This, of course, leads to another reason for expressing various genes in suitable heterospecific hosts. The ability to produce large quantities of a cloned gene product using expression vectors is often not achieved simply by inserting the gene of interest behind suitable signals on a plasmid. When insulin is produced in *E. coli*, the reducing environment inside the cells prevents the formation of important disulphide bonds (Chan *et al.*, 1981). This can be circumvented by getting the protein exported by fusing it to an appropriate signal peptide (Talmadge and Gilbert, 1982). Under such conditions, the protein is secreted into the periplasmic space where the signal peptide is cleaved off and the authentic mature protein can be recovered by an osmotic shock procedure. However, this is not always successful in that

production of specific proteins can be substantially reduced; for example, when human growth factor or human necrosis factor were fused to the signal peptide of alkaline phosphatase, synthesis of the hybrid was significantly lower than when the signal sequence was absent. *Streptomyces*, which secrete a number of enzymes directly into the growth media, are now being used as possible hosts for foreign genes. The commercial production of proteins such as insulin in *Streptomyces* is a viable proposition since methods for growing this organism in industrial (150,000 litres) fermenters is already established.

On the other hand, expressing *Streptomyces* genes in *E. coli* presents similar problems to those of *Bacillus*. The numerous sigma factors present in *Streptomyces* which have no counterpart in *E. coli* could quite conceivably prevent promoter recognition. Nevertheless, it would facilitate cloning experiments (particularly where multiple steps are involved) if at least some actinomycete genes could be expressed in *E. coli*, since the doubling time of the latter bacterium is approximately a quarter that of actinomycetes. Most of the genes transferred into *E. coli* not surprisingly fail to be expressed and those that are, often result from transcription from *E. coli* signals. When the PABA synthase gene from *S. griseus* was subcloned in *E. coli* using pBR322 as vector, expression of the gene occurred as a result of a deletion of *S. griseus* DNA (Gil and Hopwood, 1983). Likewise, the *S. acrimycini* chloramphenicol acetyltransferase was expressed in the same vector only as a consequence of a deletion event. When the *cat* gene was cloned in the *Bam*HI site of pBR322 as a 1.7 kb *Bcl*I fragment, all of the

CHAPTER 2

MATERIALS AND METHODS

chloramphenicol resistant *E. coli* contained a recombinant plasmid in which the *S. acrimycini* DNA was in the same orientation and which had suffered a 0.7 kb deletion. Subsequent examination of recombinant plasmid from the ampicillin resistant and chloramphenicol/tetracycline sensitive transformants revealed that the gene was not expressed from the "intact" plasmids regardless of the orientation of the insert (Gil *et al.*, 1985). The quite plausible explanation of these effects proposed by the authors was that a barrier to transcriptional readthrough into the *cat* gene from the pBR322 *tet* promoter had been removed. However, rearrangements of DNA is not a prerequisite for the expression of *Streptomyces* genes in *E. coli.* The hygromycin phosphotransferase (*hyg*) gene originally isolated from *S. hygroscopicus* was expressed in *E. coli* when cloned in the *Pst*I site of pBR322 (Zalacain *et al.*, 1987) and *pac*, a gene encoding the puromycin acetyltransferase from *S. alboniger* was expressed in *E. coli* using pUC19 as vector (Vara *et al.*, 1985). Despite the fact that no deletions occurred in these cases, expression was still the result of transcription from *E. coli* promoters. Indeed, expression of *pac* in pUC19 was shown to be subject to inducibility by IPTG. Similarly, the *S. erythraeus ermE* gene was transcribed from the *lac* promoter in pUC18, and the gene was expressed only when cloned in the same orientation as the direction of transcription of *lacZ* (Katz *et al.*, 1987). So although the two promoters that direct transcription of *ermE* in *Streptomyces* have sequences that resemble the canonical prokaryotic sequences, they do not allow transcription in *E. coli* (Bibb *et al.*, 1985).

The use of mutants for complementation is a common method

for cloning genes. Indeed, the *E. coli* genes that enable this bacterium to utilise 3,4, dihydroxyphenylacetate (HPA) were cloned by precisely this method, i.e. their ability to complement mutations which did not allow *E. coli* to grow on HPA as sole carbon source (Jenkins and Cooper, 1988). In addition, the same *E. coli* mutants were used to isolate the corresponding genes from *Klebsiella peumoniae* (A. Fawcett, personnal communication).

In a similar set of experiments a gene (*leu*) was isolated from *S. rochei* which complemented the *leu2* mutation in *S. lividans*, as well as the *leuA* mutation in *E. coli*, indicating that *S. rochei leu* and *S. lividans leu2* genes encode α-isopropylmalate synthases (Hercomb *et al.*, 1987). This opens the possibility of isolation of various *Streptomyces* genes without the prior generation of suitable mutants if an appropriate *E. coli* strain exists.

So, there are a number of reasons for expressing genes in a foreign host. It may allow commercial production of important proteins; alternatively it may simply be more convenient to express a gene in a surrogate organism e.g. if the natural host is either slow growing or pathogenic.

The quantity of data concerning gene expression in prokaryotes is large but the database is by no means complete. As far as actinomycetes are concerned, it is currently an exciting field in which to study since novel discoveries are frequently being made.

By examination of genes from *Streptomyces* and other actinomycetes we may be able to clarify the mechanisms underlying primary metabolism, secondary metabolism and differentiation in these bacteria.

METHODS

1. Origin, maintenance and growth of organisms.

1.1. Bacterial strains.

The bacteria used in this work and their relevant characteristics are indicated in Table 2.1.

1.2. Growth and preservation of actinomycetes.

Spores from all *S. lividans* strains were obtained by growth on NE agar which contained 1% (w/v) glucose, 0.2% (w/v) yeast extract, 0.1% (w/v) beef extract, 0.2% (w/v) casamino acids and 2% (w/v) agar, adjusted to pH 7.0 with KOH.

All *Streptomyces* were preserved as spore suspensions in glycerol. These were prepared by removing spores from the surface of an NE agar plate using 5-10 ml of sterile water and filtering them through a cotton wool plug to remove mycelium. Spores were then collected by centrifugation in a Hereaus-Christ centrifuge at 3000 rev min^{-1} for 10 minutes at room temperature and resuspended in 20% (v/v) glycerol and stored at -20°C.

Confluent plates of *Streptomyces* spores were prepared by spreading a loop of spore suspension over the surface of an NE agar plate followed by incubation at 30°C for 4-5 days.

The various *Micromonospora* species were obtained by growth

Table 2.1. Bacterial strains used.

Organisms	Relevant characteristics	References
S. lividans TK21	Plasmid free strain	Hopwood *et al.*, 1983
S. lividans TSK1	TK21 containing pIJ702	Skeggs*et al.*, 1985
S. lividans JR14	TK21 containing pLST14	Thompson *et al.*, 1985
S. lividans TSK31	TK21 containing pLST31	Skeggs*et al.*, 1987
S. lividans TSK41	TK21 containing pLST41	Skeggs *et al.*, 1985
S. lividans TSK51	TK21 containing pLST51	Skeggs*et al.*, 1987
S. tenjimariensis ATCC 31603	Istamycin producer	Inst. Microbial Chem. SS939 Patent strain
S. tenebrarius NCIB 11028	Nebramycin producer	Higgens and Kastner, 1968
S. vinaceus NCIB 8852	Viomycin producer	Caltrider, 1967
S. kanamyceticus ISP 5500	Kanamycin producer	Ser, 1957
M. purpurea DSM 43036	Gentamicin producer	Leudemann and Brodsky, 1964
M. melanosporea DSM 43141	Non producer	Baldacci and Locci, 1961
M. inyoensis ATCC 77600	Sisomicin producer	Wagman and Weinstein, 1980
M. rhodorangea NRRL 5326	G418 producer	Wagman and Weinstein, 1980
E. coli NM522	Δ(*lac-pro* AB), r_k^- m_k^+ *thi*, *sup* E [F'*pro* AB,*lac* $I^Q Z\Delta$ M15]	Gough and Murray, 1983
B. subtilis BD170 ATCC 33608	*trp* C2,*thr* 5	Gryzcan *et al.*, 1978
E. coli A19	RNase I^-	Gesteland, 1966

on MGM agar which contained 0.3% (w/v) beef extract, 0.5% (w/v) tryptone, 0.1% (w/v) glucose, 2.4% (w/v) starch, 0.5% (w/v) yeast extract, 0.6% (w/v) calcium carbonate, 0.4% (w/v) potassium chloride and 2% agar.

All *Micromonospora* were preserved as mycelial scrapings in glycerol. They were prepared by removing the mycelia from an MGM agar plate using 5-10 ml sterile water. These were collected by centrifugation in a Hereaus-Christ centrifuge at 3000 rev min^{-1} for 10 minutes at room temperature and resuspended in 20% (v/v) glycerol.

1.3. Growth and preservation of *E. coli*.

All *E. coli* strains were grown on LB agar at 37°C for 16-24 hours. LB agar contained 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) sodium chloride and 2% agar. Cells from an LB agar plate which contained 1000-5000 colonies were removed in 5 ml minimal salts medium [0.2% (w/v) glucose, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 90 mM $Na_2HPO_4.12H_2O$, 22 mM KH_2PO_4, 9 mM NaCl, 19 mM NH_4Cl] and centrifuged in a Heraeus-Christ centrifuge at 3000 rev min^{-1} for 10 minutes at room temperature. The cells were then resuspended in 0.5 ml minimal salts medium containing 20% (v/v) glycerol, and stored at -20°C.

2. Cell-free protein synthesis: Conditions for assay and preparation of components.

The following methods were used to prepare subcellular fractions for *in vitro* protein synthesis. Except were stated, all

manipulations were carried out at 0-4°C and the final products were divided into small portions, rapidly frozen in an IMS-CO_2 bath and stored at -70°C.

2.1. Preparation of coupled transcription-translation systems from *S. lividans* TK21.

This method is essentially as published (Thompson *et al.*, 1984).

Spores from *S. lividans* TK21 were removed from a confluent NE agar plate by agitation in 10 ml 0.1% (v/v) Triton X-100, and used to inoculate 2 x 1 litre YEME medium supplemented with 0.5% polyethylene glycol 6000 in a 2 l baffled flask. YEME contains 1% (w/v) glucose, 0.3% (w/v) yeast extract, 0.5% (w/v) peptone, 0.3% (w/v) malt extract and 0.1% (w/v) $MgCl_2.6H_2O$.

Mycelium from 6 x 1 litre cultures, incubated for 14-16 hours at 30°C in a New Brunswick orbital shaker at 250 rev min^{-1}, were harvested by centrifugation at 9000 rev min^{-1} for 10 minutes in a Beckman JA10 rotor. The mycelium was resuspended in 400 ml buffer I [10 mM HEPES-KOH pH 7.6 at 20°C, 10 mM $MgCl_2$, 1 M KCl, 10 mM 2-mercaptoethanol and 10% (v/v) glycerol] and washed by centrifugation as above. The washing procedure was repeated twice with buffer I and twice with buffer II [50 mM HEPES-KOH pH 7.6 at 20°C, 10 mM $MgCl_2$, 60 mM NH_4Cl, 5 mM 2-mercaptoethanol and 10% (v/v) glycerol]. Following the final wash, the mycelium was resuspended in 150 ml buffer II, collected by filtration onto Whatman

Nº 1 paper, weighed and resuspended in 2.5 ml buffer II per gramme wet weight. The suspension was then passed through a chilled French pressure cell at 10000-12000 psi (70-90 Mpa). Unbroken mycelium and cell debris were cleared from the preparation by centrifugation at 15000 rev min^{-1} for 30 minutes in a Beckman SW27 rotor. The supernatant was removed and recentrifuged under identical conditions. The resulting 30000 x g supernatant was designated S30 and was typically 200-300 A_{260} units ml^{-1}. Then, while the bulk of the S30 was held at 0°C, a small portion of S30 was treated with micrococcal nuclease for increasing incubation times to determine that necessary to remove endogeneous DNA and RNA that would otherwise contribute to the plasmid-independent activity of the system. S30 (30 A_{260} units) was incubated with 1 ml nuclease (stock solution: 150 units ml^{-1} in 50 mM glycine-KOH pH 9.2 at 20°C, containing 5 mM $CaCl_2$) in 1 mM $CaCl_2$ at 30°C. Samples containing 5 A_{260} units were removed at 10 minutes intervals and EGTA-KOH (pH 7.0 at 20°C) was added to 2 mM final concentration in order to chelate calcium ions and render the calcium-dependent nuclease inactive. A portion (2 A_{260} units) of each nuclease-treated sample was then assayed for coupled transcription-translation in the presence and absence of exogeneous plasmid, to determine the minimum incubation time needed to remove plasmid-independent activity. The remainder of the S30 was then appropriately treated with nuclease prior storage at -70°C.

2.2. Conditions for coupled transcription-translation assays.

Assays were performed in 30 µl volumes and contained 26%

(v/v) synthesis mix, 2-3 μg plasmid DNA, 2 μl [^{35}S] methionine (14 μCi μl^{-1}, 800 Ci mmol^{-1}), magnesium acetate to give 12 mM final Mg^{2+} and 2 A_{260} units of S30.

Synthesis mix contained 200 mM HEPES-KOH pH 7.6 at 20°C; 140 mM ammonium acetate; 280 mM potassium acetate; 7 mM DTT; 5 mM ATP (sodium salt, pH 7.0 with Tris); 3.4 mM CTP, GTP and UTP (all sodium salts, pH 7.0 with Tris); 100 mM PEP (trisodium salt, pH 7.0 with Tris); 19 amino acids (minus methionine) each at 1.4 mM; 7.5% (w/v) PEG 6000; 260 mM calcium folinate and 100 units pyruvate kinase in 20% (v/v) glycerol.

Incubations were at 30°C for 25 minutes when 2 μl 0.22 mg ml^{-1} methionine were added and incubation continued for a further five minutes to allow completion of radiolabelled proteins. A 5 μl sample was removed into 0.1 M KOH and heated at 95°C for 7 minutes to hydrolyse methionyl-tRNA while the remainder was stored at -20°C. Following the addition of excess 10% (w/v) TCA, acid precipitable material was collected on Whatman GF/C, extensively washed with 5% (w/v) TCA and dried under infra-red light. The radioactivity retained on the filters was estimated by liquid scintillation spectrometry using a toluene-based scintillation fluid (Fisofluor Nº3).

2.3. Polyacrylamide gel analysis of the products of protein synthesis *in vitro*.

Samples (containing approximately 2 x 10^5 cpm) prepared as above, were mixed with one third volume of loading buffer (375 mM

Tris-HCl pH 8.8 at 20°C; 4% (w/v) SDS; 35% (v/v) glycerol; 2.7 M 2-mercaptoethanol; 0.01% BPB) heated for 10 minutes at 95°C prior loading onto a 12% (w/v) polyacrylamide gel (14.5 x 10.5 x 1.5 cm) containing 0.1% (w/v) SDS, prepared according to the standard procedure of Laemmli (1970). Electrophoresis was typically for 16 hours at 12 mA constant current. The gels were fixed in 7% (v/v) acetic acid, treated with AMPLIFY and dried onto Whatman Nº1 paper, prior to fluorography using Fuji RX film at -70°C.

2.4. Conditions for protein synthesis directed by polyuridylic acid.

Cell-free synthesis of polyphenylalanine was carried out in 50 µl reaction mixtures. Reactions were started by the addition of 25 µl of "poly U cocktail" to ribosomes (5 pmol) and S100* in a volume of 25 µl. The poly U cocktail contained 40 mM HEPES-KOH (pH 7.6 at 20°C), 20 mM $MgCl_2$, 200 mM KCl, 5 mM ATP (sodium salt), 0.75 mM GTP (sodium salt), 10 mM PEP (trisodium salt), 19 amino acids (minus phenylalanine) each at 75 µM, L-[U-^{14}C] phenylalanine (10 µCi ml^{-1}, 518 mCi $mmol^{-1}$), 2 mg ml^{-1} *E. coli* tRNA (phenylalanine specific), 1 mg ml^{-1} polyuridylate (potassium salt) and 200 units ml^{-1} pyruvate kinase in 20% (v/v) glycerol. The final concentrations of magnesium and monovalent cations were 15 mM and 150 mM respectively.

The reactions were performed at 30°C. 10 µl samples were removed into approximately 1 ml 0.1 M KOH at 10 minutes time intervals and processed as described in Section 2.2.

2.5. Preparation of cell-free extracts for *in vitro* protein synthesis directed by polyuridylic acid.

E. coli strains were grown overnight in 50 ml LB medium at 37°C with vigorous shaking. 10 ml of this culture were used to inoculate each 1 litre LB (in a 2 litre flask) and the cells grown until the absorbance of the culture at 600 nm was approximately 0.6. The cells were harvested by centrifugation at 7000 rev min^{-1} for 5 minutes at 4°C in a Beckman JA10 rotor, and washed by similar centrifugation twice through a buffer containing 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 1 M NH_4Cl and 5 mM 2-mercaptoethanol and once through a buffer containing 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 50 mM NH_4Cl and 5 mM 2-mercaptoethanol. The cells were broken by grinding with an equal weight of alumina and resuspended at a concentration of 1 mg ml^{-1} in 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 50 mM NH_4Cl and 5 mM 2-mercaptoethanol containing 5 μg ml^{-1} DNase. Unbroken cells and cell debris were cleared from the preparation by centrifugation at 18000 rev min^{-1} for 30 minutes at 4°C in a Beckman JA21 rotor. The resulting 30000 x g supernatant was designated S30.

The S30 was layered on an equal volume of 10 mM Tris-HCl pH 7.6 at 20°C, 30 mM $MgCl_2$, 1 M NH_4Cl, 5 mM 2-mercaptoethanol and 20% (w/v) sucrose and subject to centrifugation at 40000 rev min^{-1} for 14 hours at 4°C in a Beckman Ti75 rotor. The supernatant was

designated S100* and was dialysed against 4 x 1 litre 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 50 mM NH_4Cl and 5 mM 2-mercaptoethanol while the ribosomal pellet was resuspended in 3 ml of the same buffer and layered over an equal volume of that buffer supplemented with 40% (w/v) sucrose. After centrifugation at 40000 rev min^{-1} for 5 hours at 4°C in a Beckman Ti75 rotor, the ribosomal pellet was resuspended in 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 50 mM NH_4Cl and 5 mM 2-mercaptoethanol. Both ribosomes and S100* were aliquoted in small portions, rapidly frozen in an IMS/CO_2 bath and stored at -70°C.

2.6. Preparation of ribosomal subunits.

Ribosomes prepared as above were dialysed against 3 x 1 litre of a buffer containing 10 mM Tris-HCl pH 7.6 at 20°C, 1 mM $MgCl_2$, 50 mM NH_4Cl and 5 mM 2-mercaptoethanol in order to dissociate the particles into their component subunits. The dissociated ribosomes were then layered onto linear 10-30% (w/v) sucrose gradients in the above buffer and centrifuged for 1.5 hours at 40000 rev min^{-1} in a Sorvall TV850 rotor. The gradients were then pumped through an ISCO UA5 density gradient fractionator with 60% (v/v) glycerol and the A_{254} was monitored continuously. Fractions containing 50S and 30S subunits were separately pooled, the magnesium chloride concentration raised to 10 mM and the particles collected by centrifugation at 40000 rev min^{-1} for 16 hours at 4°C in a Beckman Ti70 rotor. The ribosomal subunits were finally resuspended in 10 mM Tris-HCl pH 7.6 at 20°C,

10 mM $MgCl_2$, 50 mM NH_4Cl, 5 mM 2-mercaptoethanol and stored in small aliquots at -70°C.

3. Quantitation of nucleic acid and ribosomal components.

The concentration of nucleic acid or ribosomal components was determined by measuring their absorbance at 260 nm. The following conversion factors were used: 1 A_{260} unit is equivalent to 50 $\mu g\ ml^{-1}$ double stranded DNA, 37 $\mu g\ ml^{-1}$ single stranded DNA, 45 $\mu g\ ml^{-1}$ RNA, 29.4 pmol ml^{-1} 70S ribosomes, 46 pmol ml^{-1} 50S ribosomal subunits or 87 pmol ml^{-1} 30S ribosomal subunits.

4. Methylase assays.

Methylase activity, specific to certain strains, was identified in S100* preparation described earlier. The assay for enzyme activity was performed in a buffer containing 50 mM HEPES-KOH pH 7.5 at 20°C, 10 mM Tris-HCl pH 7.6 at 20°C, 10 mM $MgCl_2$, 50 mM NH_4Cl, 5 mM 2-mercaptoethanol and 5 μCi [^{3}H-methyl] S-adenosyl-methionine (500 mCi $mmol^{-1}$). Assays were performed in a total volume of 0.2 ml containing 40 pmol 30S ribosomal subunits and 2% (v/v) S100*. The assay mix, in the absence of 30S subunits, was preincubated for 10 minutes at 35°C before initiating the reaction by adding those particles. Samples (20 μl) were removed at intervals into approximately 1 ml ice-cold 5% (w/v) TCA and incubated for a further 20 minutes at 0°C, before collecting acid precipitable material on Whatman GF/C filters. The filters were extensively washed with 5% (w/v) TCA, dried under infra-red light and the radioactivity retained on the filters was

estimated by liquid scintillation spectrometry using a toluene based scintillation fluid.

5. Isolation and manipulation of DNA.

5.1. Preparation of total genomic DNA.

The isolation of total DNA from various actinomycetes was carried out using the lytic procedure of Smith (in Hopwood *et al.*,1985).

In the case of *Streptomyces*, growth was in YEME medium supplemented with 34% sucrose and 0.5% glycine, whereas *Micromonospora* were grown in liquid MGM medium. Each were incubated at 30°C for 40-42 hours and harvested by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature in a Heraeus-Christ centrifuge. The mycelium was washed by centrifugation through 10.3% (w/v) sucrose. The pellet was resuspended in 4 ml "lysozyme solution" (2 mg ml^{-1} in Tris-HCl pH 8.0 at 20°C, 25 mM Na_2EDTA pH 8.0, 10.3% (w/v) sucrose) and incubated for 10-30 minutes at 37°C. After the addition of 4 ml Kirby mix [2% (w/v) sodium triisopropylnapthalene sulphonate, 12% (w/v) sodium 4 aminosalycilate, 0.1 M Tris-HCl pH 8.0 at 20°C, 6% (v/v) phenol equilibrated with 0.1 M Tris-HCl pH 8.0 at room temperature], the preparation was agitated on a vortex mixer for 1 minute. Phenol (saturated as above): chloroform (50:50) (8 ml) was added, the mixture agitated for a further 15 seconds and then centrifuged at 3000 rev min^{-1} for 10 minutes. The supernatant was removed, re-extracted with an equal volume of phenol:chloroform and centrifuged as above.

Following precipitation of DNA and RNA from the aqueous phase with an equal volume of isopropanol in the presence of 300 mM sodium acetate, nucleic acid was collected by centrifugation at 3500 rev min^{-1} for 20 minutes at room temperature in a Hereaus-Christ centrifuge. The pellet was washed with 80% (v/v) ethanol and dissolved in 5 ml TE buffer (10 mM Tris-HCl pH 8.0 at 20°C, 1 mM Na_2EDTA pH 8.0 at 20°C) containing 40 $\mu g\ ml^{-1}$ RNase (pre-heated to 95°C for 15 minutes to inactivate contaminating deoxyribonuclease according to Maniatis *et al.*, 1982). The preparation was incubated at 37°C for 30 minutes, extracted with phenol:chloroform and reprecipitated with isopropanol and salt as above. The DNA was finally resuspended in 1 ml TE buffer and stored at 4°C.

5.2. Preparation of plasmid DNA from *Streptomyces* .

Since all the *Streptomyces* plasmids used in this study contained the thiostrepton resistance gene as a primary selectable marker, cultures were routinely grown in media supplemented with thiostrepton (20 $\mu g\ ml^{-1}$). All plasmids were prepared by the alkaline lysis method of Kieser (1984), from 25 ml cultures grown in YEME supplemented with 34% sucrose and 0.5% glycine for 40-48 hours at 30°C.

Mycelium was harvested by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature, washed by centrifugation through 10.3% (w/v) sucrose and resuspended in 5 ml final volume of "lysozyme solution" (see above). Following incubation at 37°C for 20 minutes, 2.5 ml of freshly prepared alkaline SDS (0.3 M sodium

hydroxide, 2% SDS) was added an the mixture agitated immediately on a vortex mixer. The lysate was incubated at 70°C for 20 minutes and then slowly cooled to room temperature. The preparation was then extracted with 2 ml unbuffered phenol:chloroform (50:50) and centrifuged for 15 minutes at 3000 rev min^{-1}. The aqueous phase was removed, extracted twice with neutral phenol:chloroform and twice with chloroform. Nucleic acid precipitated by addition of an equal volume of isopropanol in the presence of 300 mM sodium acetate and incubated at room temperature for 20 minutes. The precipitate was then collected by centrifugation at 3500 rev min^{-1} for 20 minutes in a Hereaus-Christ centrifuge at room temperature, washed with 80% (v/v) ethanol, dried *in vacuo* for 7 minutes and resuspended in 1 ml TE buffer containing 40 $\mu g\ ml^{-1}$ RNase. After incubation at 37°C for 30 minutes, the preparation was extracted with phenol:chloroform and with chloroform alone and reprecipitated as above. The DNA was dried and finally resuspended in 1 ml TE buffer.

Occasionally the DNA was purified on a CsCl/EtBr gradient. Plasmid DNA was diluted to 4 ml with TE buffer and 4.2 g of CsCl added. When the CsCl had fully dissolved 0.2 ml ethidium bromide (10 mg ml^{-1}) was added and this solution was centrifuged at 50000 rev min^{-1} for 10-16 hr at 20°C in a Beckman VTi65.2 rotor. When visualised in ultra violet light, two bands could be seen. The DNA forming the lower band was collected, and the ethidium bromide removed by extraction with water saturated butanol. The supercoiled plasmid was precipitated with sodium acetate and isopropanol at -20°C. The DNA was collected by centrifugation for 10 minutes in a Burkard

microcentrifuge, washed with absolute ethanol, dried *in vacuo* and resuspended in 0.1 ml TE buffer.

5.3. Large scale preparation of *E. coli* plasmids.

Cultures (50 ml) were grown for 16 hours at 37°C in LB medium supplemented with 200 $\mu g\ ml^{-1}$ ampicillin. Cells were harvested in universals by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature in a Hereaus-Christ centrifuge, resuspended in 2 ml of lysozyme solution (25 mM Tris-HCl pH 8.0 at 20°C; 10 mM Na_2EDTA pH 8.0 at 20°C; 50 mM glucose; 5 mg ml^{-1} lysozyme) and incubated at room temperature for 15 minutes. Cell lysis was achieved by the addition of 4 ml of freshly prepared alkaline SDS solution (0.2 M NaOH; 1% SDS). The mixture was inverted several times and incubated at 0°C for 10 minutes. Chromosomal DNA was precipitated by the addition of 3 ml of 3 M potassium acetate (pH 4.8 at 20°C; see materials). The contents of the universal were mixed thoroughly and incubated on ice for 10 minutes. The precipitate was removed by centrifugation at 3000 rev min^{-1} for 10 min at room temperature in a Hereaus-Christ centrifuge. The aqueous phase was extracted with phenol:chloroform and chloroform before precipitating the nucleic acid by the addition of 0.6 volume of isopropanol and incubation at room temperature for 15 minutes. The nucleic acid was collected by centrifugation at 3500 rev min^{-1} for 20 minutes at room temperature in a Hereaus-Christ centrifuge. The pellet was washed with absolute ethanol, dried *in vacuo* and resuspended in 0.5 ml TE buffer containing 40 $\mu g\ ml^{-1}$ RNase. After incubation at 37°C for 30 minutes the preparation was extracted as above and precipitated with

2.5 volumes of ethanol and 300 mM sodium acetate at -20°C for 1 hour. The plasmid DNA was collected by centrifugation for 10 minutes in a Burkard microcentrifuge, dried *in vacuo* and resuspended in 0.1 ml TE buffer. When the plasmids were destined for use in coupled transcription-translation reactions, RNase was not used.

5.4. Small scale preparation of *E. coli* plasmids.

When a large number of plasmids were required for analysis either by restriction mapping or by sequencing, a modification of the procedure described by Birnboim and Doly (1979) was employed. Selected colonies were grown in 2 ml LB medium supplemented with 200 mg ml^{-1} of ampicillin in an 11.5 ml Sarstedt test tube, suspended horizontally and shaken vigorously at 37°C overnight. 1.5 ml were transferred to a 1.6 ml microcentrifuge tube and harvested by centrifugation for 2 minutes. The cells were resuspended in 0.1 ml of lysozyme solution (see above) and incubated at room temperature for five minutes. Lysis was achieved by the addition of 0.2 ml freshly prepared alkaline SDS (see above) to the tubes which were sharply inverted 2-3 times and incubate, on ice, for 5 minutes. Chromosomal DNA was then precipitated by adding 0.15 ml 3 M potassium acetate (see materials) pH 4.8 at 20°C and mixing the contents of the tube by inverting it and agitating briefly (1-2 seconds) on a vortex mixer. After incubation on ice for 5 minutes the precipitate was collected by centrifugation in a Burkard microcentrifuge. The aqueous phase was extracted with phenol:chloroform and chloroform before the nucleic acid was precipitated using 2 volumes of isopropanol and incubation at room temperature for 5 minutes and collected by centrifugation in a

Burkard microcentrifuge for 10 minutes. The pellet was washed with absolute ethanol (precooled at -20°C), dried *in vacuo*, and resuspended in 0.1 ml TE buffer containing 40 μg ml^{-1} RNase. The preparation was incubated at 37°C for 1 hour before extraction with phenol:chloroform and chloroform. The aqueous phase was taken and the plasmid DNA precipitated by the addition of 2.5 volumes of isopropanol and 300 mM sodium acetate. The DNA was collected, washed and dried as above. The pellet was resuspended in 40 μl TE buffer. For restriction analysis 1 μl of the preparation was adequate, and when required for sequencing, 10μl was used.

6. Agarose gel electrophoresis of DNA samples.

DNA preparations were routinely analysed in horizontal 0.7% (w/v) agarose gels cast in TEA buffer [40 mM Tris-acetate pH8.0 at 20°C, 2 mM Na_2EDTA]. DNA samples containing 10% (v/v) sample buffer [TEA buffer supplemented with 50% glycerol, 0.01% (w/v) xylene cyanol FF and 0.01% (w/v) BPB] were loaded onto gels (8 cm x 8 cm x 3 mm) which were submerged in TEA and electrophoresed at 80 V for 30-40 minutes. DNA was visualised by staining in 1 μg ml^{-1} ethidium bromide for 15 minutes and observation on a u.v. transilluminator. Lambda phage DNA fragments from digestion with restriction endonuclease were used as size markers for linear molecules (Daniels *et al.*, 1980).

7. Restriction, phosphatase treatment and ligation of DNA.

DNA was cleaved with various restriction endonucleases under the conditions specified by the manufacturers. Reactions were typically

in volumes less than 50 µl. When the DNA was required for further manipulation, it was subjected to additional treatment. After extraction with an equal volume of neutral phenol:chloroform, the DNA was precipitated from the aqueous phase with 3 volumes of ethanol and 300 mM sodium acetate using an IMS/CO_2 bath for 10 minutes. The precipitate was collected by centrifugation for 10 minutes in a Burkard microcentrifuge, washed in 80% ethanol, dried under vacuum for 7 minutes and finally resuspended in TE buffer.

Occasionally, vector DNA was also treated with calf intestinal alkaline phosphatase (CIAP) during restriction. This enzyme removes the terminal phosphates after cleavage and therefore abolishes the recircularisation of vector molecules in ligation mixtures. Typically, 0.5-1 unit of enzyme were used to treat several microgrammes of DNA. Before phenol:chloroform extraction and subsequent precipitation, the CIAP was inactivated by heating the sample at 75°C for 15 minutes in the presence of 0.1% (w/v) SDS, 10 mM Tris-HCl pH 8.0 at 20°C, 100 mM sodium chloride and 1 mM Na_2 EDTA.

In subcloning experiments, vector and donor DNA fragments were ligated at approximately 1:2 molar stoichiometry in 25-30 µl at a concentration of about 40 µg ml^{-1}. Ligation was carried out in a buffer containing 50 mM Tris-HCl (pH 7.6 at 20°C), 10 mM $MgCl_2$, 1mM ATP, 1 mM DTT, 5% (w/v) PEG-8000, 1 unit T4 DNA ligase was added and incubation was carried out at room temperature for 16-20 hours. The ligation mixture was the used directly for transformation experiments.

8. Isolation of specific DNA molecules from agarose gels.

In some experiments it was necessary to purify particular DNA fragments prior to ligation. The DNA species of interest were separated by electrophoresis in a 1% (w/v) low melting point agarose in TEA buffer under the conditions described previously. The DNA was visualised on a u.v. transilluminator.

The method employed was the CTAB method described by Landgridge *et al.*, (1980). Gel slices containing the required DNA were removed from the gel and incubated in 1.6 ml microcentrifuge tube at 70°C for 20 minutes. The following manipulations were performed at 37°C. An equal volume of water/CTAB (see materials) and an equal volume of butanol/CTAB (see materials) were added and the contents of the tube agitated for 30 seconds using a vortex mixer. The phases were separated by centrifugation for 1 minute in an MSE microcentaur microcentrifuge. The upper butanolic phase was retained and the aqueous phase re-extracted with half a volume of butanol/CTAB. After centrifugation as above, the butanol phases were pooled. The preparation could now be returned to room temperature. The DNA was extracted from the butanol phase by the addition of 1/4 volume of 0.3 M sodium acetate pH 7.0 at 20°C. The mixture was agitated by a vortex mixture for 30 seconds and the phases separated by centrifugation for 1 minute in a microcentrifuge. The lower, aqueous phase was removed and, as a consequence of an increase in volume, 3 M sodium acetate was added to return the concentration of those ions to 300 mM. The aqueous phase was extracted with chloroform before being precipitated with 3 volumes of ethanol and incubation in an IMS/CO_2 bath for 10 minutes. The DNA was collected by centrifugation for 10

minutes in a Burkard microcentrifuge, dried *in vacuo* and resuspended in 10 µl TE buffer.

9. Southern hybridisation of DNA.

Over the period in which this work was completed a second preferable method of Southern analysis was developed. However, some analysis described here were performed using the prior technique. Both methods are described, but will be listed separately.

9.1. The GeneScreen method.

9.1.1. Transfer of DNA.

DNA samples to be transferred to GeneScreen plus (New England Nuclear) were separated by electrophoresis in a 0.7% agarose gel in TEA buffer at 6 V cm^{-1} for 3 hours. To allow efficient transfer of large DNA molecules, the DNA in the gel was "acid-nicked" by incubation in 0.25 M HCl at room temperature. The gel was then incubated in a solution containing 0.4 M sodium hydroxide and 0.6 M sodium chloride for 30 minutes and then in 1.5 M sodium chloride plus 0.5 M Tris-HCl pH 7.5 at 20°C for 30 minutes.

To prepare the membrane for DNA transfer, it was cut to the size of the gel, washed in distilled water and then incubated in a solution containing 3 M sodium chloride and 0.3 M sodium citrate for 15 minutes.

The blotting apparatus was set up with a wick of Whatman

3MM paper supported on a glass plate. The ends of the wick were placed in a solution containing 3 M sodium chloride and 0.3 M sodium citrate and the gel placed on top of this, with the top surface uppermost. The membrane was placed onto the gel and the gel surrounded by spacers composed of double layered Saranwrap to prevent the buffer from bypassing the gel and membrane. Five pieces of dry filter paper were placed above the membrane and a 10 cm stack of absorbent paper towels were held in position above these, under a 1 kg weight. Transfer was allowed to continue for 16-24 hours at room temperature, after which the membrane was removed from the gel, incubated in 0.4 M sodium hydroxide for 1 minute to denatured the transferred DNA and then briefly in an excess of a solution containing 0.2 M Tris-HCl (pH 7.5 at 20°C), 0.3 M sodium chloride and 0.03 M sodium citrate. The membrane was dried at room temperature and then incubated at 42°C for 6 hours in 10 ml prehybridisation buffer in a sealable plastic bag. This contained 50% deionised formamide, 1% (w/v) SDS, 1 M sodium chloride and 10% dextran sulphate.

9.1.2. Hybridisation.

After 6 hours prehybridisation at 42°C, the denatured radiolabelled DNA probe (see section 11) and salmon sperm DNA were introduced into the plastic bag containing the membrane. This was then resealed and incubated for a further 16 hours at 42°C with constant agitation. After this period of hybridisation, the membrane was removed from the bag, washed twice in 100 ml of 0.3 M NaCl plus 0.03 M sodium citrate for 5 minutes at room temperature, twice in 200 ml of a solution containing 0.3 M NaCl, 0.03 M sodium citrate and 0.1%

(w/v) SDS at 65°C for 30 minutes and then twice in 100 ml of 15 mM sodium chloride, 1.5 mM sodium citrate at room temperature for 30 minutes. This washing procedure removed from the membrane probed DNA which had less than 85% similarity to the bound DNA. The filter was then dried and radioactive bands visualised by autoradiography with Fuji RX film.

9.2. The Hybond-N method.

The DNA samples were subject to electrophoresis in an agarose gel under the conditions described above.

The gel was washed in 0.25 M hydrochloric acid for 7 minutes before denaturing the DNA in a solution containing 0.5 M sodium hydroxide and 1.5 M sodium chloride for 30 minutes at room temperature. Finally the gel was neutralised by washing in 3 M sodium chloride and 0.5 M Tris-HCl (pH 7.4 at 20°C) for 30 minutes at room temperature. The blotting apparatus was set up as above with the following differences. With the 3MM wicks set on a glass plate the gel was topped with Hybond-N nitrocellulose membrane (Amersham), cut to the size of the gel and pre-soaked in a buffer containing 0.045 M sodium citrate and 0.45 M sodium chloride. A sheet of Whatman Nº1, presoaked as above, was placed on this and on top of this was a 10 cm stack of paper towels. The paper towels were under a 1 kg weight, and wet towels were replaced every 5 minutes for the first 30 minutes and every 15 minutes for the next 1.5 hours.

The membrane was then removed, washed briefly in 0.045 M

sodium citrate, 0.45 M sodium chloride and allowed to dry at room temperature. The membrane was wrapped in a single layer of Saranwrap and placed on a u.v. transilluminator for 2 minutes, with the side that was in contact with the gel facing downwards. The membrane was prehybridised in 25 ml buffer containing 0.27 M NaCl, 0.015 M NaH_2PO_4 (pH 7.4 at 20°C), 1.5 mM Na_2 EDTA (pH 7.4 at 20°C), 1% (w/v) SDS, 6% (w/v) PEG 6000 and 0.5% (w/v) dried milk in a perspex chamber for 1 hour at 65°C. The membrane was then transferred to a second chamber containing 25 ml of the same solution containing 25 ng of radiolabelled DNA probe and hybridisation was for 16 hours at 65°C with constant agitation. The membrane was then washed briefly in 0.045 M sodium citrate, 0.45 M sodium chloride at room temperature before being washed 3 times in the same solution containing 0.1% (w/v) SDS at 65°C for 10 minutes and finally 3 times in 200 ml of a solution containing 7.5 mM sodium citrate, 0.075 M sodium chloride and 0.1% (w/v) SDS at 65°C for 10 minutes. This washing procedure was as stringent as that for GeneScreen plus. The membrane was allowed to dry and the radioactive bands visualised by autoradiography.

10. Preparation of bacteria for transformation with plasmid DNA.

10.1. Preparation, transformation and regeneration of *S. lividans* protoplasts.

The method employed was that of Bibb *et al.*, (1978), modified by Thompson *et al.*, (1982a). Since transformation is via protoplasts

and these are highly susceptible to traces of detergent, all glassware was acid-washed and sterile plastic items were used.

Mycelium from a 25 ml culture grown in YEME medium supplemented with 34% sucrose and 0.5% glycine at 30°C for 38-40 hours, was harvested by centrifugation at 3000 rev min^{-1} for 10 minutes in a Hereaus-Christ centrifuge and washed by centrifugation through 10.3% sucrose at room temperature. Protoplasts were generated by incubation of the washed mycelium in 4 ml lysozyme solution (1 mg ml^{-1} lysozyme in L buffer) at 30°C for 30 minutes. L buffer contained 10.3% (w/v) sucrose, 2.5 mM magnesium chloride, 2.5 mM calcium chloride, 1.4 mM dipotassium sulphate, 0.4 mM potassium dihydrogen phosphate, 25 mM TES-NaOH (pH 7.2 at 20°C) and 0.2% (v/v) trace element solution. The trace element solution contained (per litre) 40 mg $ZnCl_2$; 200 mg $FeCl_3.6H_2O$; 10 mg $CuCl_2.2H_2O$; 10 mg $MnCl_2.4H_2O$; 10 mg $Na_2B_4O_7$ and 10 mg $(NH_4)_6Mo_7O_{24}.4H_2O$. The suspension was broken up and incubated for a further 15 minutes at 30°C and then diluted with 5 ml P buffer (as L buffer, but with 10 mM $MgCl_2$ and 25 mM $CaCl_2$). Protoplasts were separated from residual mycelium by passage through a cotton wool plug, and collected by centrifugation at 3000 rev min^{-1} for 10 min in an MSE bench top centrifuge. The protoplast pellet was resuspended in 0.5 ml P buffer and the concentration of protoplasts determined by measurement of the optical density at 600 nm (1 OD unit is equivalent to 1.5 x 10^9 protoplasts).

Protoplasts (4 x 10^9) were diluted with 5 ml P buffer and

centrifuged as above, immediately prior transformation. The pellet was resuspended in a minimal volume of P buffer (approximately 0.1 ml), to which was added 10-20 µl ligation mix. Then, 0.5 ml T buffer [25% (w/v) PEG 1000, 2% (w/v) sucrose, 1mM K_2SO_4, 75 mM $CaCl_2$, 35 mM Tris-maleic acid pH 8.0 at 20°C and 0.2% (v/v) trace element solution] was added, followed by 5 ml P buffer, not more than 30 seconds after T buffer. The protoplasts were harvested by centrifugation, resuspended in a minimal volume of P buffer and then diluted to 1 ml with the same buffer.

Protoplasts (4 x 10^8) were spread onto each regeneration plate (R2YE), since this number has been reported to be optimal for regeneration (Thompson *et al.*, 1982a). R2YE agar contained 10.3% sucrose, 1.4 mM K_2SO_4, 50 mM $MgCl_2$, 1% (w/v) glucose, 0.4 mM KH_2PO_4, 20 mM $CaCl_2$, 0.3% (w/v) L-proline, 0.5% (w/v) yeast extract, 0.01% (w/v) casamino acids (Difco), 25 mM TES-NaOH pH 7.2 at 20°C, 5 mM NaOH, 0.2% (v/v) trace element solution and 2.2% (w/v) Difco agar. The plates were incubated in a laminar flow hood for 3-4 hours prior to use. In this time, they had lost approximately 10% weight.

After 18-22 hours incubation at 30°C, primary transformants were selected by flooding the plates with 2 ml thiostrepton suspension (0.5 mg ml^{-1} in water). After further incubation for 4-5 days, the transformants were replica plated onto medium containing the appropriate selection.

10.2. Preparation and transformation of competent cells from *E. coli.*

Strain NM522 was used as the cloning host for all the manipulations in *E. coli*. This strain does not possess a functional ß-galactosidase since the α-peptide portion is absent. However, when plasmid vectors containing the α-peptide region are introduced into this strain, ß-galactosidase function is restored by intragenic complementation. Expression of this gene is increased in the presence of an inducer such as isopropyl-ß-D-thiogalacto-pyranoside (IPTG). This system is particularly useful when organisms are grown in a medium containing X-gal in addition to IPTG. X-gal is a chromogenic substrate for ß-galactosidase, the product of which is blue in colour. As a consequence, *E. coli* NM522 carrying a plasmid that contains a gene coding for α-peptide of ß-galactosidase are blue in colour when grown in the presence of IPTG and X-gal. However, if DNA is inserted in the α-peptide gene then no functional ß-galactosidase is produced and the colonies are white. Thus there is a convenient colour test to determine whether plasmid vectors contain inserts.

The method for obtaining competent cells is that of Mandel and Higa (1970). 5 µl of *E. coli* NM522 stored in glycerol at -20°C was used to inoculate 3 ml LB medium, which was incubated overnight at 37°C. 2 ml of the overnight culture was added to 200 ml LB liquid and shaken vigorously until the absorbance at 600 nm, A_{600} = 0.2-0.225 was reached. Cultures were kept on ice for 10 minutes and then centrifuged in a Heraeus-Christ centrifuge at 3000 rev min^{-1} for 10 minutes at 4°C. The cell pellet was resuspended in 80 ml ice-cold 0.1 M calcium chloride and incubated on ice for 30 minutes. The cells were harvested

again at 4°C and resuspended in 0.1 M calcium chloride to 1 ml final volume. These were then incubated for 1-24 hours at 0°C prior to the addition of 1/4 volume 0.1 M calcium chloride; 50% (v/v) glycerol and storage as 0.2 ml aliquots in 1.5 ml Nunc vials at -70°C.

In order to determine the transformation efficiency of a preparation of competent cells, 10 ng supercoiled pUC18 DNA (Yanish-Perron *et al.*, 1985) in a volume less than 30 µl was added to 0.2 ml competent cells and incubated at 0°C for 30 minutes. The transformation mixture was transferred to 42°C for 2 minutes, diluted with 1 ml LB and incubated at 37°C for 1 hour. A portion of the transformation mixture (0.2 ml) was spread over the surface of an LB agar plate supplemented with 50 µg ml^{-1} ampicillin, to select transformants. After 16 hours incubation at 37°C, the transformation efficiency was determined.

When DNA had been ligated into the α–peptide encoding DNA of pUC18, the ligation mix was used directly to transform competent cells as described above. However, the LB agar plates contained, in addition to 50 µg ml^{-1} ampicillin, 0.3 mM IPTG and 50 µg ml^{-1} X-gal.

After growth overnight at 37°C, white transformants were selected and spread onto LB agar plates containing 200 µg ml^{-1} ampicillin and incubated at 37°C. This secondary selection was required as the secreted ß-lactamase inactivated enough of the ampicillin in the transformation plates to allow the growth of plasmid-free cells.

11. Preparation of radiolabelled DNA fragments.

Radiolabelled probes for Southern analysis were prepared by a method using random hexanucleotide primers (Feinberg and Vogelstein, 1983).

The DNA fragment to be labelled was excised from a 1% (w/v) low melting point agarose gel after electrophoresis at 10 V cm^{-1} for 40 minutes and visualisation in ethidium bromide on a u.v. transilluminator. The weight of the gel slice was determined and 1.5 ml distilled water added per gramme of agarose. From this an approximate value for the DNA concentration was calculated. The gel slice was incubated at 37°C for 60 minutes. 25 ng DNA fragment was radiolabelled in a 25 µl reaction volume at room temperature for 16 hours. The reaction contained 50 mM Tris-HCl (pH 8.0 at 20°C), 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 20 µM each of dATP, TTP, dGTP, 200 mM HEPES-NaOH pH 6.6 at 20°C, 1.35 A_{260} units hexadeoxynucleotides, 0.4 mg ml^{-1} enzyme grade BSA, 25 µCi [α-^{32}P] dCTP (3000-4000 Ci $mmol^{-1}$) and 2 units large fragment of DNA polymerase I.

The reaction was stopped by addition of 0.1 ml buffer [20 mM Tris-HCl pH 7.5 at 20°C, 20 mM NaCl, 2 mM Na_2EDTA pH 8.0, 0.25% (w/v) SDS and 1 µM dCTP] and unincorporated nucleotides were removed by gel filtration through a Sephadex G50 column. The column was prepared in a siliconised 1 ml Gilson tip plugged with siliconised glass wool, with G50 that had been pre-incubated in 3 mM Tris-HCl (pH 7.0 at 20°C) plus 0.2 mM Na_2EDTA (pH 8.0 at 20°C). The column was

rinsed with this buffer and then allowed to drain. After the sample had been loaded and allowed to run into the column, more buffer was added. Ten 0.2 ml fractions were then collected and 1 µl of each was placed in a xylene-based scintillation fluid (Fisofluor Nº1 from Fisons) and the radioactivity estimated. Two peaks of radioactivity were obtained. Fractions from the first peak were pooled and denatured by incubation at 95°C for 10 minutes together with 1 mg salmon sperm DNA, immediately prior hybridisation.

12. Transcript mapping.

12.1. Preparation of total RNA.

S. lividans strains were grown in YEME medium supplemented with 34% sucrose and 0.5% glycine at 30°C for 40-48 hours. Mycelium was harvested by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature. Cells were washed twice by centrifugation through 10.3% sucrose, resuspended in 8 ml lysozyme solution [1 mg ml^{-1} lysozyme in P buffer] and incubated for 20 minutes at 30°C. The mycelium was then broken up by pipetting and incubated for a further 10 minutes at 30°C. The protoplasts that had formed were separated from mycelium by passage through a cotton wool plug. The protoplasts were collected by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature in an MSE bench top centrifuge, and resuspended in 8 ml lysis solution [25 mM Tris-HCl pH 7.4 at 20°C; 10 mM Na_2EDTA pH 8.0 at 20°C; 0.1 M NaCl; 1% (w/v) SDS] and agitated by vortex mixer for 2-3 minutes. The lysate was extracted with an equal volume of neutral phenol:chloroform and agitated on a vortex mixer for five minutes. The

phases were separated by centrifugation at 3000 rev min^{-1} for 10 minutes at room temperature. The aqueous phase was extracted a further twice with neutral phenol:chloroform and then twice with chloroform. The nucleic acid was precipitated using 300 mM ammonium acetate and 2 volumes of ethanol and incubation at -20°C for 3 hours. The precipitate was collected by centrifugation in 30 ml Corex tubes at 7000 rev min^{-1} for 20 minutes at 4°C in an HB4 rotor using a Sorvall RCB5 centrifuge. The pellet was dried *in vacuo* and resuspended in 2 ml sterile water. Specific precipitation of the RNA was achieved by the addition of 3 volumes of 4 M sodium acetate pH 6.0 at 20°C and incubation at 4°C for 16 hours. The precipitated nucleic acid was collected as above and the supernatant discarded. The pellet was left to dry and then resuspended in 1 ml sterile water.

The RNAs were stored as isopropanol precipitates in 350 µg aliquots at -20°C.

12.2. Labelling the 5' end of DNA.

The principle of this procedure is to label a DNA fragment at one specific site by replacement of the 5' terminal phosphate with ^{32}P. This was achieved by restriction of the DNA with a single enzyme and CIAP treatment. The DNA was then restricted with a second enzyme and the appropriate fragment was extracted from an agarose gel. In this way, DNA fragments with only one 5' terminal phosphate removed and therefore suitable for radiolabelling, were obtained.

The DNA was restricted as described earlier, extracted with

phenol:chloroform and precipitated with 300 mM sodium acetate, 3 volumes of ethanol, and incubation in an IMS/CO_2 bath for 10 minutes. The DNA was collected by centrifugation for 10 minutes at room temperature in an Burkard microcentrifuge, dried *in vacuo* and resuspended in 50 µl buffer containing 50 mM Tris-HCl (pH 9.0 at 20°C), 1 mM $MgCl_2$ and 0.1 mM $ZnCl_2$. The DNA was treated with CIAP by the addition of 1 unit of enzyme and incubation at 37°C for 1 hour. The alkaline phosphatase was inactivated using the method described earlier. The DNA was cleaved with a second restriction endonuclease and extracted, after electrophoresis, from a 1% low melting point agarose gel by the method described in section 8.

The DNA fragment was resuspended in a buffer containing 50 mM Tris-HCl (pH 7.6 at 20°C), 10 mM $MgCl_2$, 5 mM DTT and 0.1 mM Na_2 EDTA pH 8.0 at 20°C, with 70 µCi [γ-^{32}P] ATP (3000 Ci mmol^{-1}) and 10 units T4 polynucleotide kinase at 37°C for 1 hour in a reaction volume of 0.1 ml. The enzyme was inactivated by extraction with phenol:chloroform and the aqueous phase separated by centrifugation for 10 minutes at room temperature in a microcentrifuge. The unincorporated radiolabelled ATP was removed by passage through a Sephadex G50 column.

The column was prepared as described by Maniatis *et al.*, (1982). A suspension of Sephadex G50 in water was autoclaved at 15 psi for 10 minutes. This material was used to prepare a column in a 1 ml syringe plugged with siliconised glass wool. The column was packed

by centrifugation at 2000 rev min^{-1} for 1 minute at room temperature in a MSE bench top centrifuge. Finally 0.1 ml sterile water was passed through the column by centrifugation as described above to prepare the column for use.

The extracted aqueous phase was loaded on the column and the labelled nucleic acid collected by centrifugation at 2000 rev min^{-1} for 1 minute at room temperature. The unincorporated radiolabelled ATP, inorganic salts and any contaminating phenol:chloroform is retained on the column, while large molecules pass through. The DNA was precipitated by the addition of 300 mM sodium acetate and 3 volumes of ethanol and stored at -20°C.

12.3. Labelling the 3' end of DNA.

The DNA was digested with two enzymes. The site to be labelled required a 5' overhang while the other enzyme was chosen to have a 3' overhang to prevent the DNA being labelled at both ends.

The appropriate restriction fragment was extracted after electrophoresis from an agarose gel as described in section 8. The DNA was labelled in a 50 µl reaction volume containing approximately 1 µg DNA; 50 mM Tris-HCl (pH 7.2 at 20°C); 10 mM $MgSO_4$; 0.1 mM DTT; 0.1 mM dCTP, dGTP, TTP, 20 µCi [α-^{32}P] dATP (3000 Ci $mmol^{-1}$); 50 µg ml^{-1} BSA (restriction enzyme grade) and 3 units large fragment DNA polymerase I (Klenow). The reaction was incubated at room temperature for 1.5 hours before the addition of unlabelled dATP (0.1 mM final concentration) and incubation at room temperature was

continued for a further 45 minutes. The reaction was stopped by incubation at 70°C for 5 minutes. The volume was increased to 0.1 ml by the addition of sterile water and the unincorporated radiolabel was removed by passage through a Sephadex G50 column as described earlier. The radiolabelled DNA was stored in ethanol at -20°C.

12.4. Hybridisation.

The radiolabelled probe DNA and the appropriate RNA preparation were collected by centrifugation for 15 minutes at room temperature in a Burkard microcentrifuge. The pellets were dried *in vacuo* and resuspended in 10 µl and 20 µl of sterile water respectively. The nucleic acids were combined, 0.24 ml deionised formamide added, heated to 100°C for 3 minutes and the preparation was then placed immediately on ice. 30 µl of 10 x concentration hybridisation buffer [0.4 M NaCl; 20 mM PIPES-NaOH pH 6.4 at 20°C; 2 mM Na_2EDTA pH 8.0 at 20°C, final concentration] was added and incubation at 60°C was for 3 hours (Favarolo *et al.*, 1980). The hybridisation mix was placed on ice for 2-3 minutes before the nucleic acid was precipitated by the addition of 3 volumes of ethanol (precooled to -20°C) and incubation overnight at -20°C.

12.5. Digestion with nuclease S1.

The precipitate was collected by centrifugation for 15 minutes in a Burkard microcentrifuge. The pellet was washed with 70% ethanol (precooled to -20°C), dried *in vacuo* and resuspended in 87.4 µl water. S1 buffer components were added to give a final buffer concentration of 50 mM sodium acetate pH 4.6 at 20°C, 280 mM sodium chloride, 5

mM zinc chloride and 20 μg ml^{-1} denatured salmon sperm DNA. Nuclease S1 (400 units) were added and incubated at room temperature for 30 minutes. The reaction was stopped by extraction with an equal volume of phenol:chloroform. The phases were separated by centrifugation for 2 minutes in a Burkard microcentrifuge, and unwanted radiolabelled products and any contaminating phenol were removed by passage through a G50 column. The hybrids were precipitated by the addition of 300 mM sodium acetate, 3 volumes ethanol and 1 μg carrier DNA. This was incubated overnight at -70°C.

13. DNA sequencing.

This method is essentially as described in the Boehringer Mannheim manual (1986) and is based on the dideoxy chain termination method (Sanger *et al.*, 1977). The DNA (5-10 μg miniprep from a small scale preparation as described in Section 4.4; 2-3 μg DNA purified on a caesium gradient) was denatured by incubation with 0.2 M sodium hydroxide and 2 mM Na_2EDTA (pH 8.0 at 20°C) for 5 minutes at room temperature in a reaction volume of 40 μl. The DNA was then precipitated by the addition of 200 mM ammonium acetate (pH 4.6 at 20°C) and 90 μl ethanol and incubation for 15 minutes in an IMS/CO_2 bath. The precipitate was collected by centrifugation for 10 minutes at room temperature in a Burkard microcentrifuge, washed with 70% (v/v) ethanol in TE buffer and dried *in vacuo*. The pellet was resuspended in 15 μl solution containing 10 mM Tris-HCl pH 7.5 at 20°C, 6.6 mM $MgCl_2$, 60 mM NaCl, 5 ng oligonucleotide primer and [^{35}S]-dATP and incubated at 37°C for 15 minutes. The hybridisation

mixes were transferred to ice and 6 units large fragment DNA polymerase I (Klenow) added. 3 µl of this hybridisation mix was added to each of four tubes containing 2 µl of A mix, C mix, G mix or T mix (see Table 2.2) and incubated at 50°C for 20 minutes. 2 µl 1.25 mM dNTP solution was then added to each tube and incubation at 50°C was continued for a further 20 minutes. The reaction was terminated by the addition of 4 µl of a buffer containing 90% deionised formamide, 10 mM Na_2EDTA pH 8.3, 0.3% (w/v) BPB, 0.3% (w/v) xylene cyanol.

MATERIALS

1. Enzymes.

The following enzymes were obtained from the Sigma Chemical Company: lysozyme (from chicken egg white), ribonuclease A (from bovine pancreas) and pyruvate kinase (type III from rabbit muscle). Calf intestinal alkaline phosphatase (molecular biology grade), deoxyribonuclease I (from bovine pancreas) and micrococcal nuclease (from *Staphylococcus aureus*) were purchased for Boehringer Mannheim. Restriction endonucleases, T4 DNA ligase, large fragment of DNA polymerase I and nuclease S1 from Bethesda Research Laboratories.

2. Biochemicals.

The following biochemicals were purchased from Sigma Chemical Company: ATP, CTP, GTP, UTP, dATP, dCTP, dGTP, TTP, ddATP, ddCTP, ddGTP, ddTTP, PEP (all sodium salts), DTT, L-aminoacids,

Table 2.2. Composition of sequencing mixes.

Nucleotide	A mix		C mix		G mix		T mix	
	µl	mM	µl	mM	µl	mM	µl	mM
dCTP	25	0.5	6.4	0.5	64	0.5	35	0.5
dGTP	25	0.5	64	0.5	6.4	0.5	35	0.5
TTP	25	0.5	64	0.5	64	0.5	3.5	0.5
ddATP	16	0.0625	-	-	-	-	-	-
ddCTP	-	-	18	1	-	-	-	-
ddGTP	-	-	-	-	18	1	-	-
ddTTP	-	-	-	-	-	-	100	2
10 x Annealing buffer	20		20		20		20	
H_2O	89		27.6		27.6		6.5	

calcium folinate spermidine trihydrochloride, IPTG, HEPES, TES, Tris, 2-mercaptoethanol, polyuridylate (potassium salt), *E. coli* unfractionated tRNA and *E. coli* tRNA (phenylalanine specific). Urea, low melting point agarose and BSA were obtained from Bethesda Research Laboratories. SDS and acrylamide were from Serva Ltd., while Sephadex G50 and hexanucleotides (cat nº 2166) were from Pharmacia fine chemicals. PEG 1000 (Koch-Light), sodium triisopropylnaphthalene sulphonate (Kodak), sodium 4-aminosalicylate (Aldrich), X-gal (Anglia Biotechnology) and Triton X-100 (Imperial Chemical Industries Ltd.) were obtained from the sources indicated.

2.1. Preparation of 3 M potassium acetate, pH 4.8.

This was prepared by mixing the following solutions: 60 ml 5 M potassium acetate, 11.5 ml glacial acetic acid, and 28.5 ml water. The resulting solution is 3 M with respect to potassium and 5 M with respect to acetate.

2.2. Preparation of butanol/CTAB and water/CTAB.

Butanol (125 ml) and water (125 ml) were mixed and the phases allowed to separate. 1 g of CTAB was then dissolved in 100 ml of the water-saturated butanol. 100 ml of the butanol-saturated water was added, mixed thoroughly and allowed to separate at room temperature (approximately 16-20 hours). The upper, butanolic phase was removed and both solutions kept in tightly capped bottles at 37°C.

3. Reagents for radiochemical analysis.

L-[U-^{14}C] phenylalanine, [^{35}S] methionine, [α-^{32}P] dCTP, [α-^{32}P]

dATP, [γ-^{32}P] ATP, [α-^{35}S] dATP, [^{14}C] methylated protein mixture and AMPLIFY were obtained from the Radiochemical Centre, Amersham. Polaroid Type 57 4 x 5" land film was used to photograph agarose gels and Fuji RX film was used for autoradiography and fluorography. Fisofluor Nº3 (Fisons) was used for liquid scintillation spectrometry.

4. Antibiotics.

The antibiotics used in this work were obtained from the sources indicated: ampicillin, kanamycin and gentamicin sulphate (Sigma Chemical Co.); thiostrepton (Squibb Institute, Princeton, N.J.,USA); sisomicin sulphate (J. Davies, Pasteur Institute, Paris, France); neamine and neomycin sulphate (Upjohn Co., Kalamazoo, Mich., USA); butirosin sulphate (Warner-Lambert Co., Ann Arbor, Mich., USA); ribostamycin (Meiji Seika Kaisha, Ltd., Yokohama, Japan); tobramycin and apramycin (Eli Lilly Cº., Indianapolis).

CHAPTER 3

RIBOSOMAL RESISTANCE TO AMINOGLYCOSIDES IN ACTINOMYCETES.

Introduction.

The aminoglycoside antibiotics are a group of compounds that inhibit protein synthesis and cause mistranslation of mRNA in prokaryotes. Resistance to these antimicrobial agents in clinical isolates and some producers has been ascribed to inactivating enzymes which modify specific sites of the molecules by phosphorylation, acetylation or adenylylation, rendering the compounds impotent as antibiotics (for review see Benveniste and Davies, 1973). Indeed, the enzymes present in clinical isolates are thought, by some, to have originated in the antibiotic producers (Walker and Walker, 1970; Benveniste and Davies, 1973). Moreover, unequivocal evidence that some of these enzymes can confer resistance to aminoglycosides (at varying levels) has been obtained by cloning the relevant genes in a sensitive host (Thompson *et al.*, 1982; Crameri and Davies, 1986). However, examination of some aminoglycoside producers failed to identify any antibiotic modifying enzymes (which of course does not refute the possibility that they actually exist) but revealed that the ribosomes of various bacteria were refractory to their toxic product (Piendl and Böck, 1982; Yamamoto *et al.*, 1982; Matkovic *et al.*, 1984).

Specifically, resistance to gentamicin in the producer *M. purpurea* was manifested as a property of the 30S ribosomal subunit (Piendl and Böck, 1982). Subsequently, resistance determinants from *M. purpurea* and *S. tenjimariensis* were cloned in *S. lividans*, in which, similarly, resistance was mediated as a function of the small ribosomal subunit (Skeggs *et al.*, 1985; Thompson *et al.*, 1985). The genes from *M. purpurea* and *S. tenjimariensis* are now designated *kgmA* (kanamycin

gentamicin methylation) and *kamA* (kanamycin-apramycin methylation) respectively. Furthermore, it was shown that resistance in each case was due to modification of the 16S rRNA by S-adenosylmethionine dependent methylases, each of which, acted at a specific but different site. The precise locations of these sites of modification were identified in an elegant set of experiments, ultimately using the ability of modified bases to pause elongation by reverse transcriptase (Beauclerk and Cundliffe, 1987). In this way the altered residues were identified as G-1405 (in the case of the resistance mechanism from *M. purpurea*) and A-1408 (in the case of that from *S. tenjimariensis*) using the numbering system based on *E. coli* rRNA. Interestingly, the region of 16S rRNA modified by these enzymes has a conserved primary sequence which is not sequestered in secondary structure according the models currently proposed (Moazed *et al.*, 1986). In fact, three such conserved single stranded sequences exist within 16S rRNA and it would not be unreasonable to suggest that they might have functional roles. Evidence already exists to suggest that the sequence of 16S rRNA around C-1400 might be involved in codon-anticodon interaction (Prince *et al.*, 1982).

Streptomyces tenebrarius is another aminoglycoside producer which, in addition to any other resistance mechanism, has ribosomes that are insensitive to a range of antibiotics (Yamamoto *et al.*, 1982; Piendl *et al.*, 1984). This streptomycete produces the nebramycin complex which includes tobramycin and other derivatives of kanamycin B plus the novel aminoglycoside, apramycin. Ribosomes from this organism were found to be highly resistant *in vitro* to

kanamycins A, B and C, to the gentamicin C complex and also to neamine and ribostamycin (which comprise, respectively, two or three rings of the four ring neomycin molecule), although they were somewhat more sensitive to neomycin (Yamamoto *et al.*, 1982). Subsequently, these findings were confirmed and extended (Piendl *et al.*, 1984) when ribosomes from *S. tenebrarius* were shown to be resistant to apramycin. The 30S ribosomal subunit was implicated in the expression of these effects and, furthermore, when hybrid 30S particles (containing combinations of rRNA and r-proteins from *S. tenebrarius* and *S. lividans*) were reconstituted *in vitro* and assayed in the presence of gentamicin, it was shown that some property of the 16S rRNA of *S. tenebrarius* was responsible for ribosomal resistance to that drug (Piendl *et al.*, 1984).

Resistance determinants were isolated from *S. tenebrarius* (by P.A. Skeggs) and analysed at a number of levels and the results in this chapter have, in part, been published (Skeggs *et al.*, 1987). Those experiments performed by Dr. P.A. Skeggs are clearly indicated.

Results.

When *S. tenebrarius* DNA was cloned in *S. lividans* (by P.A. Skeggs), selection with various aminoglycoside antibiotics identified two quite distinct phenotypes. On the one hand, the first group resembled the *S. lividans* clones containing *M. purpurea* DNA (*S. lividans* JR14) (Thompson *et al.*, 1985) while on the other, some clones of *S. lividans* harbouring *S. tenebrarius* DNA resembled *S. lividans* strain TSK41 that contain DNA derived from *S. tenjimariensis* (Skeggs

et al., 1985). Accordingly, the minimum inhibitory concentrations of a number of aminoglycoside antibiotics for the various strains of *S. lividans* were determined and those patterns of collateral resistance together with comparable data relating to *S. tenebrarius, S. tenjimariensis, M. purpurea,* and *S. lividans* containing pIJ702 (TSK1) are given in Table 3.1. Evidently, strains TSK31 and TSK51 were both highly resistant to kanamycin, whereas gentamicin and apramycin discriminated between them. Given that *S. tenebrarius* was highly resistant to each of the drugs listed in Table 3.1, it was interesting that fairly high levels of resistance to most of those compounds were also displayed by *S. lividans* TSK31 or by strain TSK51 or by both. On the other hand, strain TSK31 was only marginally resistant to neomycin and, although the resistance level of strain TSK51 was significantly higher, it did not approach that seen with *S. tenebrarius.* Accordingly, repeat attempts were made (by P. Skeggs) to generate clones, using *S. tenebrarius* DNA, that would grow on higher concentrations of neomycin (e.g. 50 $\mu g\ ml^{-1}$), but without success. Certainly, when the DNA cloned in TSK31 and TSK51 were combined on a single plasmid, it failed to transform *S. lividans* to neomycin resistance.

When the DNA fragments from *S. tenebrarius* cloned in pLST31 and pLST51 had been subjected to restriction analysis, they were compared with maps of the *M. purpurea* DNA fragment present in pLST14 (isolated from *S. lividans* JR14) and *S. tenjimariensis* DNA contained in pLST41 (Fig. 3.1). The latter DNA fragment (about 3.7 kb) contained the kanamycin-apramycin resistance gene, which was subsequently subcloned to a smaller insert size in pLST412 (see

Table 3.1. Minimal inhibitory concentrations (μg ml^{-1}) of aminoglycoside antibiotics for *S. lividans* strains, *M. purpurea*, *S. tenjimariensis*, and *S. tenebrarius*.*

				S. lividans				
Antibiotic	*M. purpurea*	*S. tenjimariensis*	*S. tenebrarius*	TSK1	JR14	TSK31	TSK41	TSK51
Kanamycin	> 1000	> 1000	> 1000	10-30	> 1000	> 1000	> 1000	> 1000
Tobramycin	> 500	100-300	> 1000	3-10	300-1000	300-1000	100-300	100-300
Apramycin	< 1	300-1000	> 1000	< 1	< 1	1-3	300-1000	300-1000
Gentamicin	> 1000	5-10	> 1000	3-10	> 1000	> 1000	10-30	10-30
Sisomicin	> 1000	100-300	> 1000	3-10	> 1000	> 1000	100-300	100-300
Neamine	N.T.	N.T.	> 1000	10-30	10-30	10-30	> 1000	> 1000
Ribostamycin	N.T.	N.T.	> 1000	3-10	3-10	3-10	> 1000	> 1000
Butirosin	N.T.	N.T.	> 1000	< 1	< 1	< 1	> 1000	> 1000
Neomycin	< 1	10-30	> 1000	3-10	3-10	3-10	30-100	30-100

* Spores were streaked on NE agar plates (see methods) containing antibiotics and growth was monitored after 3-5 days incubation at 30°C.

N.T. Not tested.

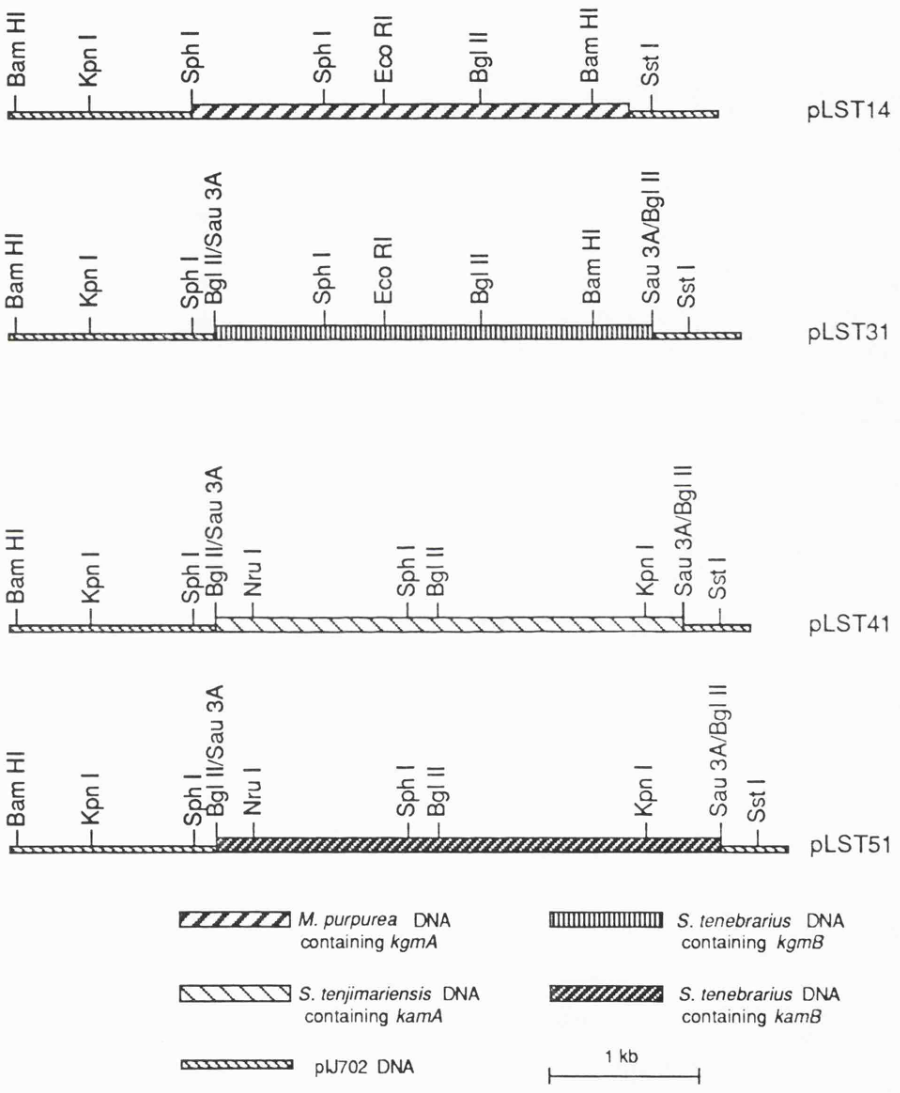


Fig. 3.1 Restriction maps of DNA fragments carrying aminoglycoside resistance determinants.

Chapter 5). The similarities between the inserts in pLST31 compared with pLST14 and pLST51 compared with pLST41 were obvious, suggesting that the similar patterns of aminoglycoside resistance linking *S. lividans* strain JR14 with TSK31 and strain TSK41 with TSK51 resulted from their possession of related (and possibly homologous) resistance genes. Further evidence in support of this idea was obtained when a fragment of *M. purpurea* DNA, isolated from pLST14 and known to contain most of the kanamycin-gentamicin resistance determinant (see Fig. 4.5), was used to probe Southern blots (using the GeneScreen method; see Materials and Methods) containing restriction fragments from pIJ702, pLST41, pLST31, pLST51, and pLST14. Hybridisation at high stringency was observed with the piece of *S. tenebrarius* DNA present in pLST31 but not with any part of pLST41, pLST51, or pIJ702 (Fig. 3.2). Conversely, when a similar blot was probed with a radiolabelled restriction fragment derived from pLST41 and containing most of the kanamycin-apramycin resistance determinant (see Fig. 5.5), hybridisation at high stringency was observed with the *S. tenebrarius* DNA insert in pLST51 but not with any part of pLST31. Accordingly, the aminoglycoside resistance gene in pLST31 was designated *kgmB* and that from pLST51, *kamB*. Another aspect of the homologies evident in Fig. 3.1 that deserves further comment is the observation that the pieces of DNA being compared, between pLST14 and pLST31 on the one hand and between pLST41 and pLST51 on the other, appeared similar over greater distances than those normally expected to contain a single gene at least at this crude level analysis. In the absence of further data, the significance of these additional homologies cannot be assessed and they may well be

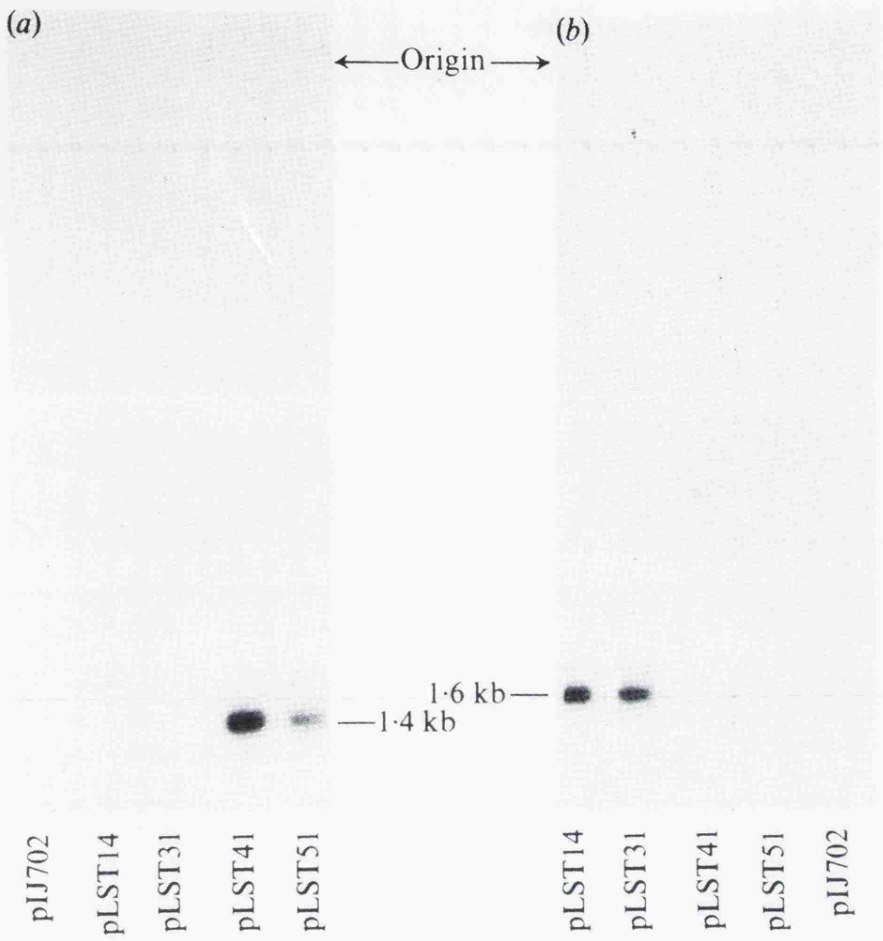


Fig. 3.2 Southern analysis of aminoglycoside resistance determinants from various actinomycetes.

Radiolabelled probes were prepared using (a) the *Bgl*II-*Kpn*I fragment from the insert in pLST41 and (b) the *Eco*RI-*Bam*HI fragment from the insert in pLST14 (see Fig. 3.1). Hybridisation was carried out with filters containing linearised pIJ702 plus the recombinant plasmids pLST14, and pLST31 restricted with *Eco*RI and *Bam*HI as well as pLST41 and pLST51 cleaved with *Bgl*II and *Kpn*I.

trivial. Nevertheless, it might be instructive to know whether resistance determinants in these various aminoglycoside-producing organisms are physically adjacent to other common genes involved in aminoglycoside metabolism.

Subsequent *in vitro* analysis (carried out by P. A. Skeggs), confirmed that *kgmB* and *kamB* encoded S-adenosylmethionine dependent methylases which modified 16S rRNA. Indeed the products of these genes shown to modify residues G-1405 and A-1408 in 16S rRNA in an identical manner to the enzymes encoded by *kgmA* and *kamA* (A.A.D. Beauclerk, personal communication).

Examination of the data concerning *S. tenebrarius* raises the question of why this organism carries an additional methylase that appears not to be required. Resistance to all of the antibiotics thus far attributed to *S. tenebrarius* can be accounted for by the expression of the gene cloned in *S. lividans* TSK51, making that cloned in *S. lividans* TSK31 seemingly redundant, assuming that the effect of methylation apart from preventing drug binding, is neutral. If in the case of the gene cloned in strain TSK31, this was not true, then this could explain its apparently anomalous presence in *S. tenebrarius*. Alternatively, it could be that the presence of multiple resistance determinants reflects a pattern, observed elsewhere, concerning organisation of genes involved in antibiotic production and resistance which are physically linked. Thus in *S. tenebrarius* genes involved in the production of apramycin and tobramycin might form separate clusters, each associated with a methylase gene. Whether such genetic organisation

exists in the chromosome or on a plasmid (or both) in *S. tenebrarius* remains to be established.

While it has been shown that homologous proteins are shared by at least three actinomycetes the possibility remained that rRNA methylation, as a means of immunity to the gentamicin class of antibiotics in those organisms which produce aminoglycosides, might be a more widespread phenomenon than hitherto suspected. Indeed, ribosomal resistance was demonstrated in various *Micromonospora* species which produce 4,6-disubstituted deoxystreptamine (i.e. gentamicin-like) compounds (Matkovic *et al.*, 1984). Accordingly, chromosomal DNAs were isolated from a number of aminoglycoside producers and after digestion with restriction endonucleases and electrophoresis in agarose gels, were blotted onto nitrocellulose membranes (Hybond-N method; see Materials and Methods). The blot was probed with a radiolabelled restriction fragment from pLST14. The results showed that, apart from *M. purpurea* and *S. tenebrarius*, *Micromonospora inyoensis* and *Micromonospora rhodorangea* possess sequences which will hybridise at high stringency with the DNA from pLST14 containing the methylase gene (Fig. 3.3). *M. inyoensis* and *M. rhodorangea* produce gentamicin-like molecules, namely sisomicin and gentamicin G418 respectively. The hybridising band in the case of *M. rhodorangea* was identical in size to that from *M. purpurea* and *S. tenebrarius*, an observation consistent with the possibility that a very similar DNA fragment might exist in that organism. On the other hand, the hybridisation pattern shown by *M. inyoensis* gave two bands implying that at least, a different restriction pattern surrounds the putative methylase gene in this organism. While these actinomycetes

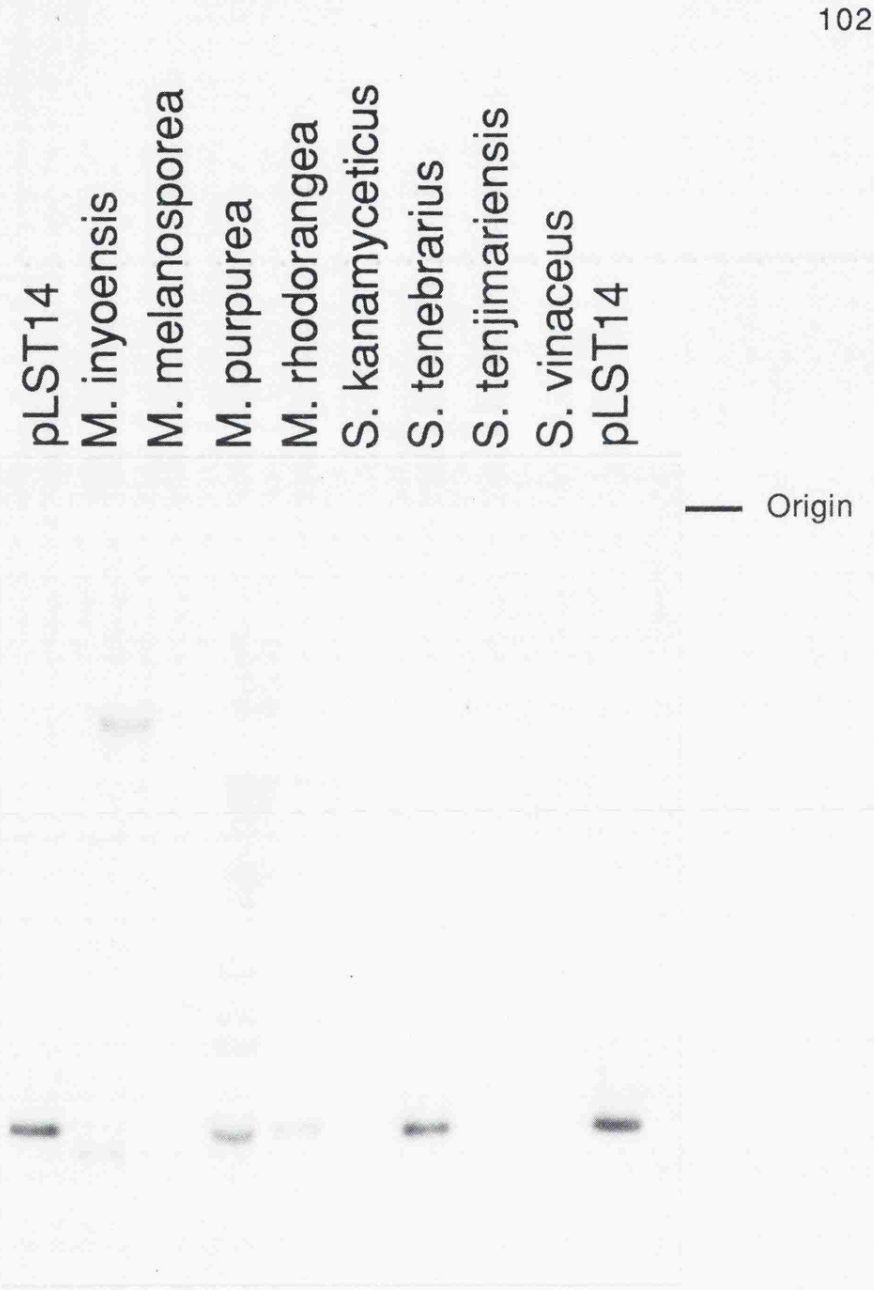


Fig. 3.3 Southern Analysis of Chromosomal DNAs from Aminoglycoside Producers Using *kgmA* as Probe.

Various chromosomal DNAs were digested with *Bam* HI and *Eco* RI and probed with the *Bgl* II -*Eco* RI fragment of pLST14. In each case 5 µg of DNA was loaded except that 25 µg of *M. purpurea* DNA was necessary to get an equivalent signal.

possess DNA which hybridises at high stringency with the probe, chromosomal DNA from some other aminoglycoside producers failed to do so. These included DNA from *S. tenjimariensis* which produces istamycin, *S. kanamyceticus* which makes kanamycin and the viomycin producer, *S. vinaceus*. Of course, this does not mean that these *Streptomyces* lack of enzymes similar to those identified in clones *S. lividans* JR14 and *S. lividans* TSK31. In fact, a kanamycin resistance determinant isolated from *S. kanamyceticus* (Nakano *et al.*, 1984) failed to hybridise with *kgmA* (data not shown) despite the fact that both encode methylase enzymes which modify 16S rRNA in the same way (A.A.D. Beauclerk, personal communication). However, the fact that positive results were shown for a number of gentamicin producers illustrates that this may be a common method of resistance to aminoglycoside perhaps acquired along with genes for biosynthesis.

The horizontal transfer of genetic information, not only between actinomycetes, is a topic of intensive study. Gene cloning experiments revealed that aminoglycoside modifying enzymes of antibiotic producing bacteria could behave as resistance determinants in susceptible heterologous hosts (Courvalin, *et al.*, 1977; Thompson and Gray, 1983).

The observation that some *Streptomyces* possess antibiotic inactivating enzymes similar to those from clinical isolates (Walker and Walker, 1970; Miller and Walker, 1969; Nimi *et al.*, 1971), has led to the suggestion that this might be the origin of the genetic pool of R-determinants (Watanabe, 1971). Circumstantial evidence in support

of this hypothesis includes the discovery that some antibiotic resistance/biosynthetic genes are located on plasmids as well as the notion that transposable elements exist in *Streptomyces*. Plasmids might be transferred by conjugation, perhaps across generic boundaries, while many of the clinical resistance determinants are carried on transposons which could conceivably have originated in actinomycetes.

The question remains then, why rRNA methylase mediated resistance to aminoglycosides has not been discovered in clinical isolates. It could be a result of the screening methods employed to identify various determinants since antibiotic resistant isolates are often simply assayed for growth in the presence of a number of aminoglycoside antibiotics in the hope of revealing a resistance profile characteristic of a previously identified modifying enzyme. The possibility remains that rRNA methylation might result in a similar or identical resistance pattern to one already assigned to a particular inactivating enzyme and perhaps these rRNA methylase genes are already present but yet to be discovered in some clinical isolates.

CHAPTER 4

NUCLEOTIDE ANALYSIS AND EXPRESSION IN *E. coli* OF A GENTAMICIN RESISTANCE DETERMINANT FROM *Micromonospora purpurea*

Introduction.

Of the order *Actinomycetales*, the genus *Streptomyces* is by far the most intensively studied. Research on other members of the order includes the isolation, characterization and DNA sequencing of a gene, *ermE,* involved in erythromycin resistance in *Saccharopolyspora erythraea* (though formerly *Streptomyces erythraeus*) (Skinner *et al.*, 1983; Bibb *et al.*, 1985); the identification of rifampicin-resistant RNA polymerase in *Amylocolatopsis mediterranei* (formerly *Nocardia*, latterly *Streptomyces mediterranei*) (Blanco *et al.*, 1984) and ribosomal resistance to aminoglycosides was illustrated in several *Micromonospora* species (Matkovic *et al.*, 1984) although the vast majority of this work has been done in *Streptomyces.*

This chapter focuses on *Micromonospora purpurea*, the producer of gentamicin, an antibiotic which normally inhibits protein synthesis and causes mistranslation in prokaryotes. Like several of the other *Micromonospora* species, the ribosomes from this organism were shown to be resistant to kanamycin and gentamicin in cell-free protein synthesis (Piendl *et al.*, 1985). Subsequently, a resistance determinant from *M. purpurea* was cloned in *S. lividans* (Thompson *et al.*, 1985) and ribosomes from the clone were shown to resemble those from *M. purpurea* in their aminoglycoside resistance characteristics. The cloned gene, probably the first to be isolated from this genus encodes a methylase which, as discussed previously, modifies 16S rRNA at position G-1405 (Beauclerk and Cundliffe, 1987). Accordingly, the designation *kgmA* (kanamycin gentamicin methylation) has been proposed for that gene (see Chapter 3).

Results.

4.1. Expression of the resistance gene (*kgmA*) in *E. coli*.

In an attempt to express the *kgmA* gene in *E. coli* NM522, the 3.6 kb *Sst*I-*Pst*I fragment of pLST14 (Fig. 4.1) was subcloned in pUC18 and pUC19. Transformants were selected on 50 μg ml^{-1} ampicillin, and recombinants identified by insertional inactivation of the *lacZ'* gene of pUC18/19. Colonies which were white in the presence of X-gal and IPTG were grown on LB agar plates supplemented with either IPTG and 200 μg ml^{-1} ampicillin or IPTG plus 5 μg ml^{-1} gentamicin (the minimal inhibitory concentration of gentamicin for *E. coli* NM522 had previously been determined as 1-2 μg ml^{-1}). Of the 32 white primary transformants selected, 31 grew on 200 μg ml^{-1} ampicillin. However, only 5 grew (and then poorly) on 5 μg ml^{-1} gentamicin and these all contained pUC18 derived plasmids. Mutants of *E. coli* resistant to aminoglycosides have been isolated (Gorini, 1974; Benveniste and Davies, 1973; Wittmann and Wittmann-Liebold, 1974; Pestka, 1977) and, more specifically, some spontaneously arising gentamicin-resistant mutants were shown to possess ribosomes with an altered ribosomal protein L6 (Buckel *et al.*, 1977). Resistance in these strains was as high as 20 μg ml^{-1} gentamicin and the frequency of appearance of resistant colonies was about 1 per 10^8 cells plated.

Given that *kgmA* confers a high level (>1 mg ml^{-1}) of resistance on *Streptomyces*, *E. coli* strains expressing that gene might be expected to grow on medium containing concentrations of antibiotic which would inhibit any spontaneously arising mutants. Accordingly, the colonies

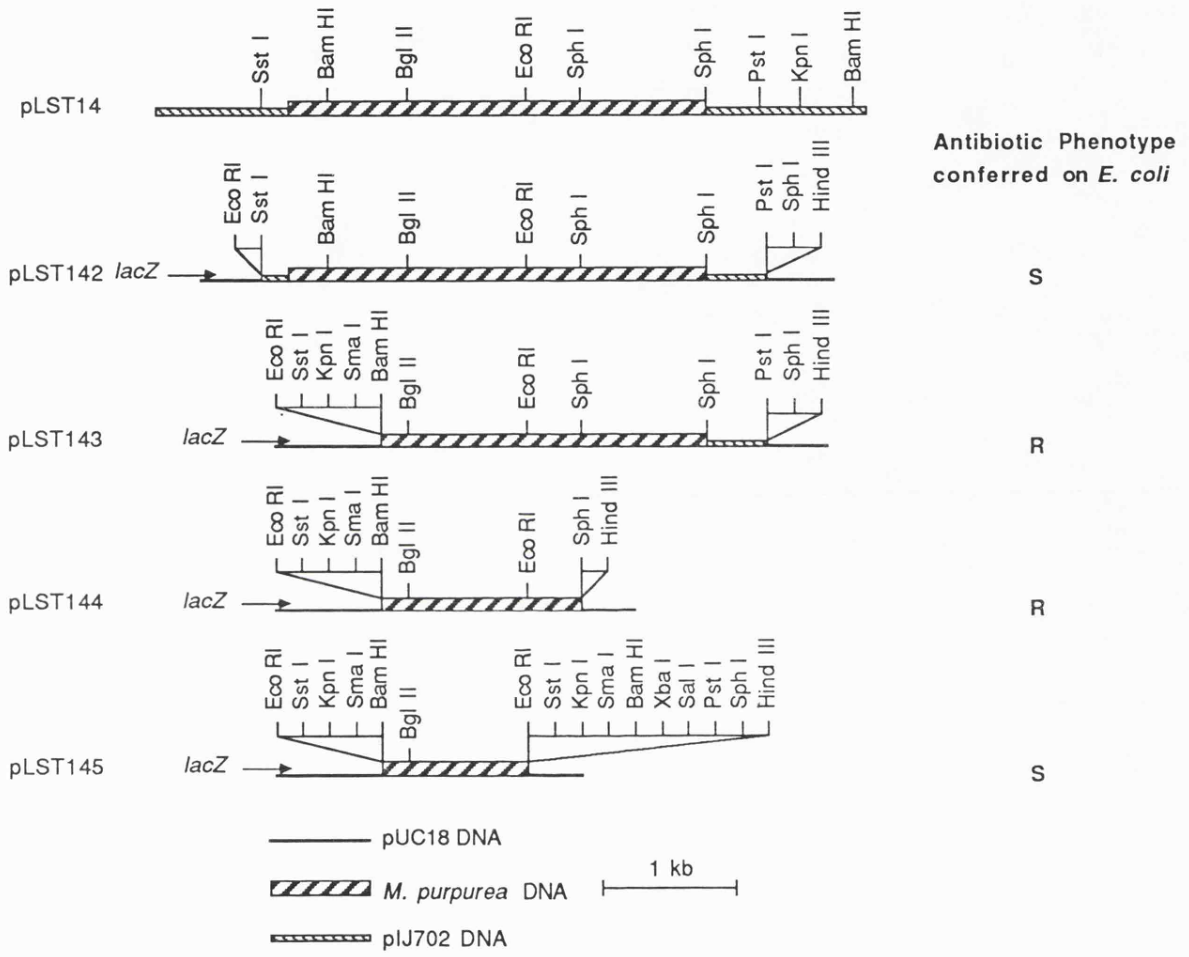


Fig. 4.1 Restriction maps of pLST14 and *E. coli* plasmids containing *M. purpurea* DNA.

The phenotype conferred by the *E. coli* plasmids is indicated (plasmids which confer aminoglycoside resistance are indicated by R, and those that do not by S) and since *kgmA* is expressed from the *lacZ'* promoter in pLST144 but not in pLST145, the carboxy-terminus of the gene must lie between the *Eco* RI and *Sph* I sites of pLST144.

growing at 5 μg ml^{-1}, were tested against a variety of aminoglycoside antibiotics (Table 4.1) which revealed that one of the isolates grew at high concentrations of gentamicin and kanamycin but was sensitive to apramycin, a resistance pattern reminiscent of *S. lividans* JR14, the primary clone (see Chapter 3).

Analysis of the recombinant plasmid in this strain of *E. coli* indicated that it had undergone some sort of rearrangement. Subsequent restriction mapping showed that approximately 1.4 kb of the inserted DNA had been deleted, removing the *Bam*HI site from the *M. purpurea* DNA insert and bringing the *Bgl*II site into closer proximity to the promoter of the *lacZ* α-fragment (Fig. 4.1). In addition, some "extra" restriction sites were detected. According to the cloning strategy, all of the restriction sites in the "polylinker" between *Sst*I and *Pst*I ought to have been removed, but the mapping suggested that the *Kpn*I, *Sma*I and *Bam*HI sites from the polylinker were still present (Fig. 4.1). Although there is no unequivocal explanation for the mechanism of this deletion, the presence of these anomalous restriction sites and the nature of the sequence at the point of the deletion (data presented below) would at least be consistent (on purely *ad hoc* basis) with a model first proposed to account for homologous recombination following transformation of *E. coli* with linear molecules (Conley *et al.*, 1986 a, b; see discussion Chapter 6). Be that as it may, the plasmid conferring resistance on *E. coli* was designated pLST143, and the strain harbouring it DH143. Subsequently, when the complete *Sst*I-*Pst*I fragment from pLST14 was successfully subcloned in pUC18, to give pLST142 (Fig. 4.1) it failed to give gentamicin resistance in *E. coli*. The

Table 4.1. Minimal inhibitory concentrations (μg ml^{-1}) of aminoglycoside antibiotics for *S. lividans* JR14, *E. coli* DH143 and *E. coli* NM522.

Antibiotic	*S. lividans* JR14	*E. coli* DH143	*E. coli* NM522
Gentamicin	› 1000	› 1000	3
Kanamycin	› 1000	› 1000	3
Tobramycin	› 1000	› 1000	3
Apramycin	1	1	1

reason for its failure to do so remains unresolved, and although it is obvious that the *kgmA* promoter was not utilised in *E. coli*, there appears to be an additional factor operating which does not allow RNA polymerase, initiating at the *lacZ* promoter, to read through the upstream sequence and transcribe *kgmA*. More of this later.

In order to confirm that the resistance phenotype of *E. coli* DH143 was indeed due to the expression of the *M. purpurea* gene (*kgmA*), it was necessary to show that the ribosomes in that strain were resistant to specific aminoglycosides, such as kanamycin and gentamicin. This was achieved using an *in vitro* protein synthesising system directed by polyuridylic acid. The mode of action of aminoglycosides causes apparently anomalous effects on these *in vitro* systems which have plagued researchers working with this group of antibiotics over the years. On the one hand, kanamycin and gentamicin inhibit ribosomal functions resulting in a reduction in protein synthesis; on the other, they cause misreading of mRNA codons. The latter phenomenon is normally affected by kanamycin in two distinct ways; the drug increases the error frequency in the initial selection of aminoacyl-tRNA by the programmed ribosome and also decreases the efficiency with which ribosomes proof-read their choice of aminoacyl-tRNA prior to peptide bond formation (Jelenc and Kurland, 1979), which can, under certain circumstances, result in an increase in protein synthesis in these assays. Despite this well documented multiphasic response (Tai and Davis, 1979; Piendl *et al.*, 1984; Skeggs *et al.*, 1985), the ribosomes from strain DH143 were highly resistant to kanamycin compared with those of the control (Fig. 4.2). In addition,

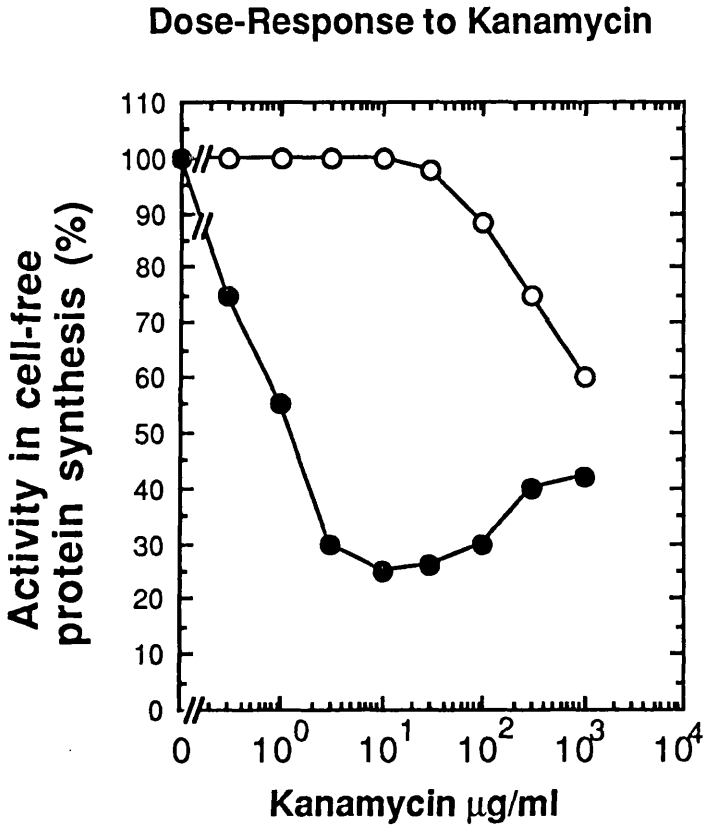


Fig. 4.2 Effects of kanamycin on cell-free protein synthesis

Assays contained S100* from *E. coli* A19 (RNaseI$^-$; kanamycin sensitive) together with 5 pmol 70S ribosomes from *E. coli* NM522 (containing pUC18) (—●—) [100% activity = 64 phe per ribosome h^{-1}] or from *E. coli* DH143 (—○—) [100% activity = 70 phe per ribosome h^{-1}].

extracts of strain DH143 were shown to possess methylase activity (identical to that of *S. lividans* JR14) for which 30S ribosomal subunits from *E. coli* NM522, but not those from DH143 were substrates (Fig. 4.3).

4.2. Subcloning the resistance gene.

The *kgmA* gene was localised within 1.1 kb of DNA by deleting a *Sph*I fragment of pLST143, generating a smaller plasmid, pLST144 (Fig. 4.1), which conferred the same properties on *E. coli* NM522 as did its parent plasmid. However, when the *Eco*RI fragment was subcloned in pUC18 to give pLST145 (Fig. 4.1), this construct failed to give drug resistance, thereby locating the 3' end of the structural gene between the *Eco*RI and the *Sph*I sites.

4.3. Nucleotide sequence of *kgmA*.

The sequence of a 1.5 kb fragment of *M. purpurea* DNA embracing *kgmA* and isolated from pLST142 (in which neither deletion nor rearrangement had occurred), was determined using the dideoxy method. Double stranded DNA was used to minimise the chance of sequencing aberrant subclones. Also, by employing this method fewer premature terminations occurred during the sequencing reactions, than when single stranded DNA was used, thereby limiting gel artifacts. The sequencing strategy (Fig. 4.4) is overlapping and sequences through all restriction sites used to generate subclones. An open reading frame was observed from nucleotide 259 to 1098 (Fig. 4.5) consisting of 280 codons that could code for a predicted protein of molecular weight 30.9 kDa. This open reading frame showed the biased codon usage

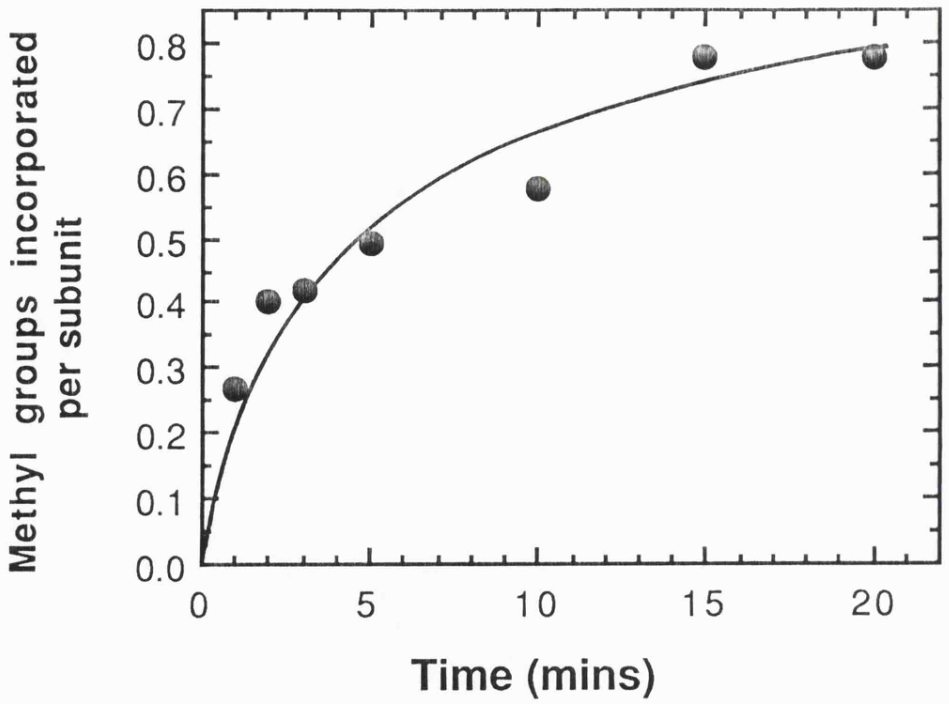


Fig. 4.3 Assay of methylase activity

Methylation of 30S ribosomal subunits from *E. coli* NM522 containing pUC18 by S100* from *E. coli* DH143 using radiolabelled S-adenosylmethionine as cofactor. Time independent background "incorporation" into 30S particles from strain DH143 has been subtracted

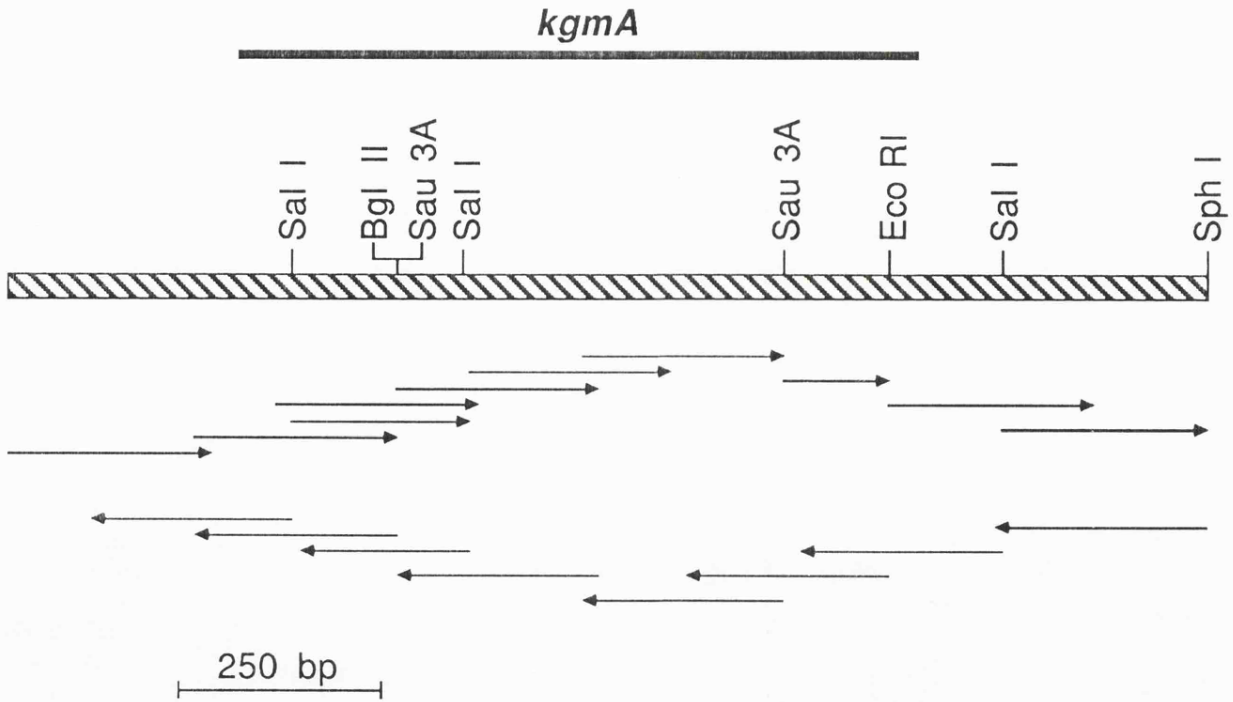
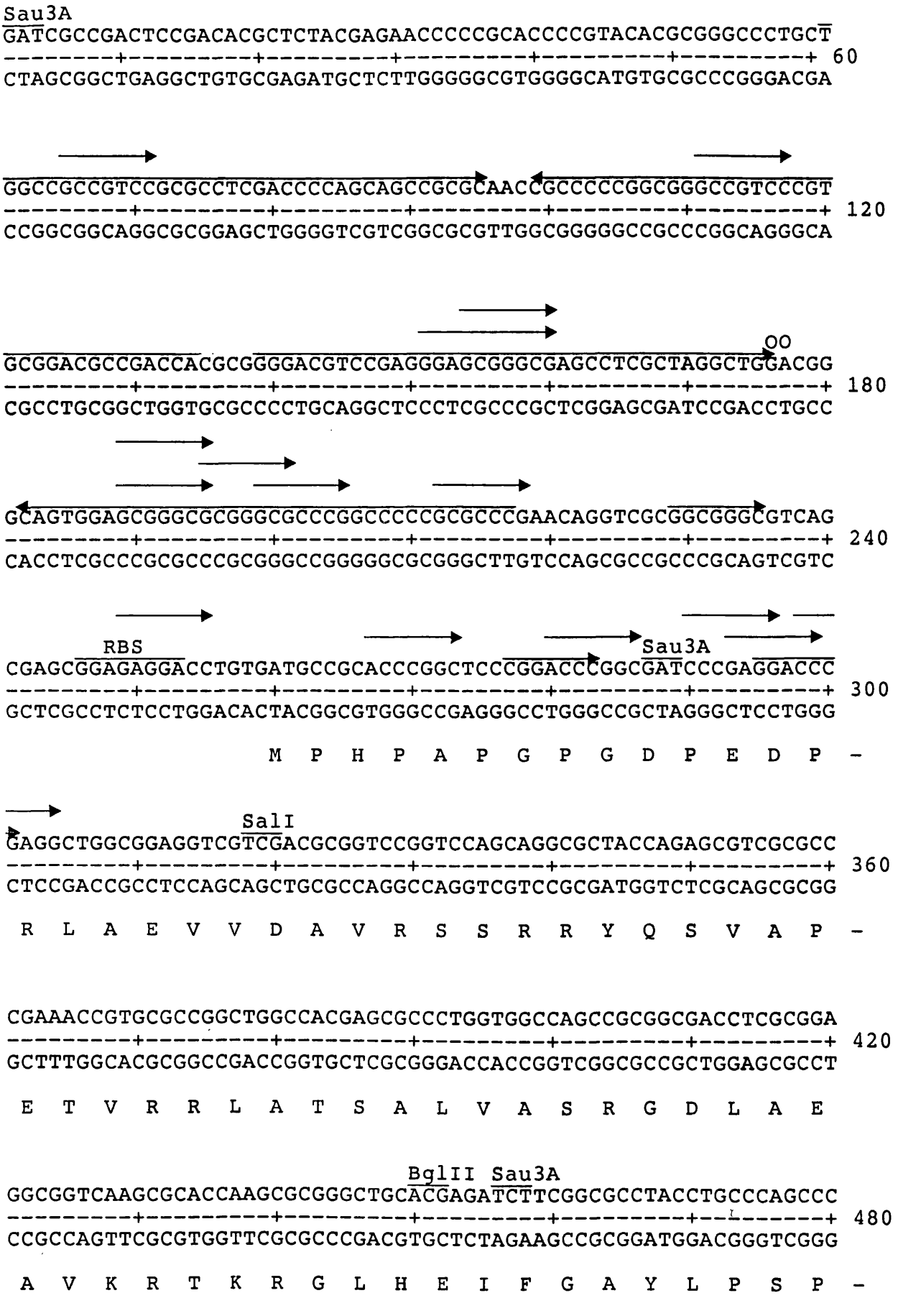


Fig. 4.4 Strategy employed to determine the nucleotide seq uence of*kgmA*

Sequence was determined through all restriction sites used to generate subclones. The sequence revealed the presence of a cluster of codons ending in A or U close to the *Bgl* II site (see also Fig. 4.5). This region of DNA was sequenced with particular care.


Sau3A
GATCGCCGACTCCGACACGCTCTACGAGAACCCCCCGCACCCCGTACACGCGGGCCCTGCT
CTAGCGGCTGAGGCTGTGCGAGATGCTCTTGGGGGCGTGGGGCATGTGCGCCCGGGACGA
60
GGCCGCCGTCCGCGCCTCGACCCCAGCAGCCGCGCAACCGCCCCCGGCGGGCCGTCCCGT
CCGGCGGCAGGCGCGGAGCTGGGGTCGTCGGCGCGTTGGCGGGGGCCGCCCGGCAGGGCA
120
OO
GCGGACGCCGACCACGCGGGGACGTCCGAGGGAGCGGGCGAGCCTCGCTAGGCTGGACGG
CGCCTGCGGCTGGTGCGCCCCTGCAGGCTCCCTCGCCCGCTCGGAGCGATCCGACCTGCC
180
GCAGTGGAGCGGGCGCGGGCGCCCGGCCCCCGCGCCCGAACAGGTCGCGGCGGGCGTCAG
CACCTCGCCCGCGCCCGCGGGCCGGGGGCGCGGGCTTGTCCAGCGCCGCCCGCAGTCGTC
240
RBS
Sau3A
CGAGCGGAGAGGACCTGTGATGCCGCACCCGGCTCCCGGACCCGGCGATCCCGAGGACCC
GCTCGCCTCTCCTGGACACTACGGCGTGGGCCGAGGGCCTGGGCCGCTAGGGCTCCTGGG
300
M P H P A P G P G D P E D P -
SalI
GAGGCTGGCGGAGGTCGTCGACGCGGTCCGGTCCAGCAGGCGCTACCAGAGCGTCGCGCC
CTCCGACCGCCTCCAGCAGCTGCGCCAGGCCAGGTCGTCCGCGATGGTCTCGCAGCGCGG
360
R L A E V V D A V R S S R R Y Q S V A P -
CGAAACCGTGCGCCGGCTGGCCACGAGCGCCCTGGTGGCCAGCCGCGGCGACCTCGCGGA
GCTTTGGCACGCGGCCGACCGGTGCTCGCGGGACCACCGGTCGGCGCCGCTGGAGCGCCT
420
E T V R R L A T S A L V A S R G D L A E
BglII Sau3A
GGCGGTCAAGCGCACCAAGCGCGGGCTGCACGAGATCTTCGGCGCCTACCTGCCCAGCCC
CCGCCAGTTCGCGTGGTTCGCGCCCGACGTGCTCTAGAAGCCGCGGATGGACGGGTCGGG
480
A V K R T K R G L H E I F G A Y L P S P -

SalI
GCCCAAGTACGACGCCCTCCTCCGCCAGCTCAGGGGGGCGGTCGACGCGGCGACGACGAG
---------+---------+---------+---------+---------+---------+ 540
CGGGTTCATGCTGCGGGAGGAGGCGGTCGAGTCCCCCCGCCAGCTGCGCCGCTGCTGCTC

P K Y D A L L R Q L R G A V D A A T T R -

GCCGTGCGGGCACCCTGCACCGCGCCATGTCCACGCACGCCTCCACCCGCGAGCGCTGCC
---------+---------+---------+---------+---------+---------+ 600
CGGCACGCCCGTGGGACGTGGCGCGGTACAGGTGCGTGCGGAGGTGGGCGCTCGCGACGG

P C G H P A P R H V H A R L H P R A L P -

CATCCTCGACGAGTTCTACCGCGAGGTCTTCGCCCGGTGCGCCGACCCGGCCAGCGTGCG
---------+---------+---------+---------+---------+---------+ 660
GTAGGAGCTGCTCAAGATGGCGCTCCAGAAGCGGGCCACGCGGCTGGGCCGGTCGCACGC

I L D E F Y R E V F A R C A D P A S V R -

TGACCTGGCCTGCGGGATGAACCCGCTCGCCGCGCCGTGGATGCCCGGCTCGGACGCGTT
---------+---------+---------+---------+---------+---------+ 720
ACTGGACCGGACGCCCTACTTGGGCGAGCGGCGCGGCACCTACGGGCCGAGCCTGCGCAA

D L A C G M N P L A A P W M P G S D A F -

CACCTACCACGCGTCCGACATCGACACCCGGCTCATGGAGTTCCTCGACGCCGCCCTGGA
---------+---------+---------+---------+---------+---------+ 780
GTGGATGGTGCGCAGGCTGTAGCTGTGGGCCGAGTACCTCAAGGAGCTGCGGCGGGACCT

T Y H A S D I D T R L M E F L D A A L E -

GACGCTCGGGGTCGCGCACGACGTCCGGGTGCGCGACCTGATGACCGGGGTCGGCGAGGT
---------+---------+---------+---------+---------+---------+ 860
CTGCGAGCCCCAGCGCGTGCTGCAGGCCCACGCGCTGGACTACTGGCCCCAGCCGCTCCA

T L G V A H D V R V R D L M T G V G E V -

CGAGACCGACGTGACGCTGCTGCTCAAGACCGTGCCCTGCATCGAGGCGCAGGGGAGGGG
---------+---------+---------+---------+---------+---------+ 900
GCTCTGGCTGCACTGCGACGACGAGTTCTGGCACGGGACGTAGCTCCGCGTCCCCTCCCC

E T D V T L L L K T V P C I E A Q G R G -

Sau3A
GCAGGGCTGGGACCTCATCGACGCGATCCGCTCGCCGCTGGTCGTGGTGAGTTTCCCGAC
---------+---------+---------+---------+---------+---------+ 960
CGTCCCGACCCTGGAGTAGCTGCGCTAGGCGAGCGGCGACCAGCACCACTCAAAGGGCTG

Q G W D L I D A I R S P L V V V S F P T -

```
GAAGTCCCTCGGCCAGCGTTCCAAGGGGATGTTCAACACCTACTCGGCGAATTTCGACGC
---------+---------+---------+---------+---------+---------+ 1020
CTTCAGGGAGCCGGTCGCAAGGTTCCCCTACAAGTTGTGGATGAGCCGCTTAAAGCTGCG

 K  S  L  G  Q  R  S  K  G  M  F  N  T  Y  S  A  N  F  D  A  -


                                          EcoRI
CTGGCTGGAGAACCGGCCGCACGACGTCGAGCAGCTCGAATTCAGGAACGAACTGGTCTA
---------+---------+---------+---------+---------+---------+ 1080
GACCGACCTCTTGGCCGGCGTGCTGCAGCTCGTCGAGCTTAAGTCCTTGCTTGACCAGAT

 W  L  E  N  R  P  H  D  V  E  Q  L  E  F  R  N  E  L  V  Y  -


TTTCGTGCGGAAGAACGCGTGAGGAAACGCCGCGCTGGATGCGTGGACCCCGTCGTCGCG
---------+---------+---------+---------+---------+---------+ 1140
AAAGCACGCCTTCTTGCGCACTCCTTTGCGGCGCGACCTACGCACCTGGGGCAGCAGCGC

 F  V  R  K  N  A  *


                                                    SalI
TCGGTGTCCCCTCGCCGCCGCACTAGACTCCTGGCGTGGGTTCCCCCGTCGACTCCGGCC
---------+---------+---------+---------+---------+---------+ 1200
AGCCACAGGGGAGCGGCGGCGTGATCTGAGGACCGCACCCAAGGGGGCAGCTGAGGCCGG


GGTTGGTCCTGGCCGCCACGCCCCTCGGTGACTTCCGGGACGCCTCGCCCCGGCTGGTCG
---------+---------+---------+---------+---------+---------+ 1260
CCAACCAGGACCGGCGGTGCGGGGAGCCACTGAAGGCCCTGCGGAGCGGGGCCGACCAGC


AGATGCTGGCCACCGCCGACGTCGTCGCGGCGGAGGACACCCGGCACCTGCGCACCCTGA
---------+---------+---------+---------+---------+---------+ 1320
TCTACGACCGGTGGCGGCTGCAGCAGCGCCGCCTCCTGTGGGCCGTGGACGCGTGGGACT


CCGGCGTCCTCGGCGTGACGCCACCGGGCGCGTGGTCAGCTTCTACGACGCCGTCGAGGC
---------+---------+---------+---------+---------+---------+ 1380
GGCCGCAGGAGCCGCACTGCGGTGGCCCGCGCACCAGTCGAAGATGCTGCGGCAGCTCCG


GGCGCGGGTGCCCGGCCTGGTCGCGCCATGCGGTCGGGCGCACCGTGTTGCTGGTGACCG
---------+---------+---------+---------+---------+---------+ 1440
CCGCGCCCACGGGCCGGACCAGCGCGGTACGCCAGCCCGCGTGGCACAACGACCACTGGC
```

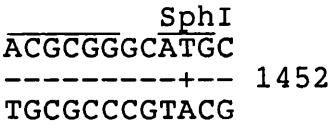


Fig. 4.5 Nucleotide sequence of DNA embracing kgmA together with the predicted amino acid sequence (below) and the positions of key restriction endonuclease cleavage sites (above). The putative ribosome binding site (RBS) is overlined and the transcription initiation site is indicated (O). The positions of some direct and inverted repeats are shown using arrows.

normally associated with DNA of high G+C content (Table 4.2) as illustrated using the UWGCG programme CODONPREFERENCE (Gribskov *et al.*, 1984) (Fig. 4.6). Curiously, this analysis revealed the presence within that ORF of a small cluster of codons ending with A or U from position 556 to 591 (Fig. 4.5 and 4.6), and consequently this part of the sequence was examined with particular care (see strategy in Fig. 4.4). Despite this anomalous region, the overall G+C content in each of the codon positions conforms with those from other actinomycete genes, in that the total G+C content of the coding sequence is 71.1%; for codon positions 1,2 and 3 the G+C contents were 70.7%, 48.2% and 94.3% respectively.

4.4. Nucleotide sequence at the site of deletion.

The sequence of the insert in pLST144 confirmed the previous observations concerning restriction endonuclease sites which should not have been present (i.e. they were not present in pLST142). Examination of the DNA sequence at the N-terminal end of the methylase gene in pLST144 strongly suggested that expression was due to a translational fusion with the N-terminus of the *lacZ* α-fragment. Accordingly, a frame shift was introduced at the beginning of the gene, in order to interrupt the sense in this putative fusion by restriction at the *Sst*I site, removal of the single stranded DNA using the 3'→ 5' exonuclease activity of the large fragment of DNA polymerase I (Klenow), ligation of the blunt ends and subsequent transformation of *E. coli* (Fig. 4.7). The frame shifted construct (designated pLST146) was verified by DNA sequencing and failed to transform *E. coli* NM522 to gentamicin resistance. The conclusion was

Table 4.2. Codon usage of *kgmA* .

aa	codon		aa	codon		aa	codon		aa	codon	
Phe	TTT	0	Ser	TCT	0	Tyr	TAT	1	Cys	TGT	0
	TTC	10		TCC	4		TAC	6		TGC	4
	TTA	0		TCA	0	Term	TAA	0	Term	TGA	1
Leu	TTG	0		TCG	3		TAG	0	Trp	TGG	3
Leu	CTT	0	Pro	CCT	1	His	CAT	1	Arg	CGT	2
	CTC	14		CCC	9		CAC	8		CGC	11
	CTA	0		CCA	0	Gln	CAA	0		CGA	1
	CTG	14		CCG	13		CAG	6		CGG	7
Ile	ATT	0	Thr	ACT	0	Asn	AAT	1	Ser	AGT	1
	ATC	6		ACC	8		AAC	5		AGC	6
	ATA	0		ACA	0	Lys	AAA	0	Arg	AGA	0
Met	ATG	6		ACG	6		AAG	7		AGG	6
Val	GTT	0	Ala	GCT	1	Asp	GAT	1	Gly	GGT	0
	GTC	15		GCC	13		GAC	19		GGC	7
	GTA	0		GCA	2	Glu	GAA	3		GGA	1
	GTG	9		GCG	17		GAG	13		GGG	9

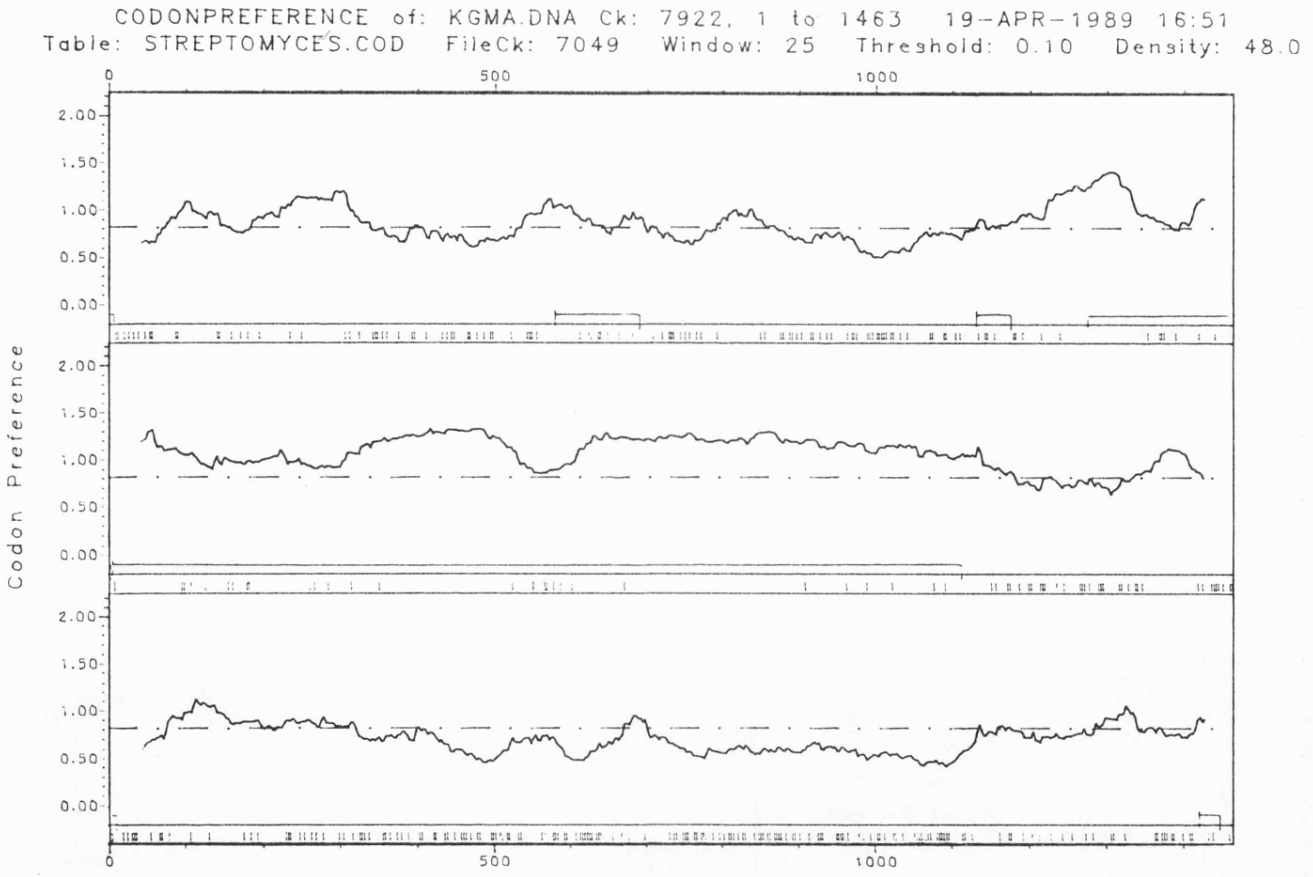


Fig. 4.6 Analysis of the sequence of *kgmA* using the UWGCG programme CODONPREFERENCE

The analysis used the default settings (i.e. a moving window of 25 codons) and a codon usage table based on a number of *Streptomyces* gene sequences. Open reading frames are shown as boxes below their respective reading frames. "Rare" codons are marked below the reading frame plot. Open reading frames showing the biased codon usage typical of *Streptomyces* genes is revealed by a codonpreference index of greater than 0.8. This analysis revealed the occurrence in *kgmA* of a cluster of codons ending in A or U.

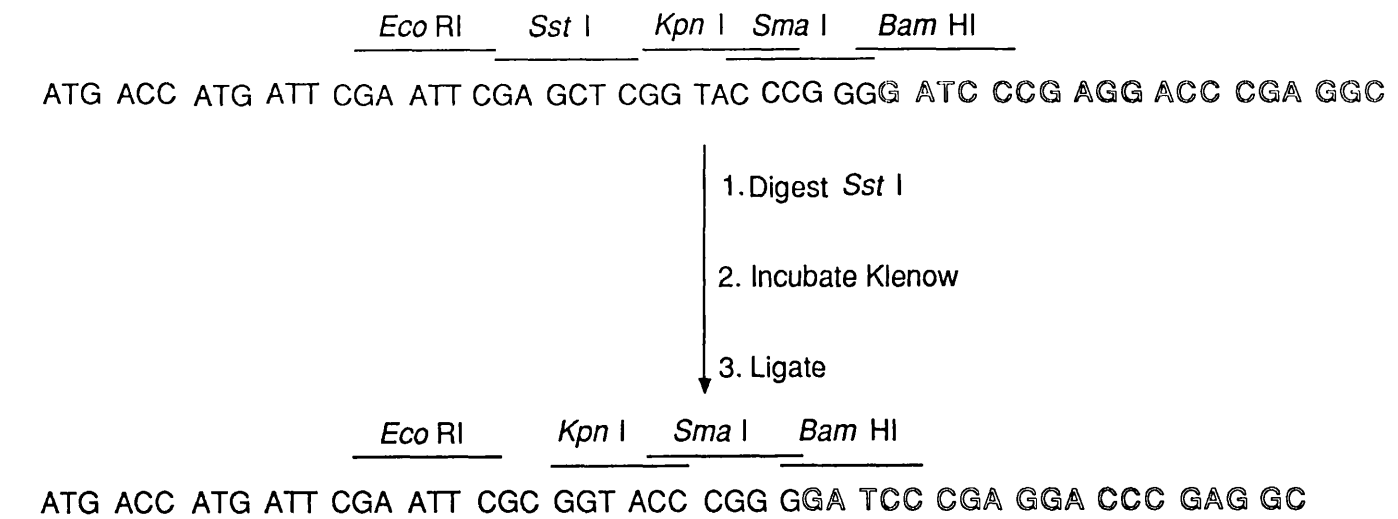


Fig. 4.7 *kgmA* is expressed as a translational fusion in *E. coli*

By changing the reading frame of the *M. purpurea* DNA in pLST144 with respect to the α -peptide translational start and illustrating the subsequent loss of the resistance phenotype, it was possible to show that the methylase gene was being expressed as a translational fusion.

GATC pUC18 sequence

GATC *kgmA* sequence

therefore, that gentamicin resistance methylase of *E. coli* DH143 had been expressed as a result of a translational fusion in which the first nine amino acids of *kgmA* had been replaced by the first 12 amino acids of the *lacZ* α-fragment of pUC18.

4.5. *In vitro* coupled transcription-translation.

Plasmids pLST142, pLST143, pLST144 and pLST146 were used to direct coupled transcription-translation in cell-free extracts of *Streptomyces*, using [^{35}S]-methionine to label the protein products. The latter were separated using SDS-polyacrylamide gel electrophoresis and visualised using fluorography (Fig. 4.8). Only plasmids pLST143 and pLST144 gave rise to an extra band when compared with pUC18 control. The gene product ran slightly faster than the 30 kDa protein marker in contrast to the predicted molecular weight 30.9 kDa. This band was confirmed as representing the fusion product in similar analysis using linearised plasmids. For example, when pLST144 was linearised with *Sca*I the band associated with the β-lactamase preprotein was absent (Fig. 4.8 track g); likewise, the 29.9 kDa protein (the putative *kgmA* product) was not expressed from pLST144 linearised with *Bgl*II (Fig. 4.8 track h). Inexplicably, however, the *kgmA* gene was not obviously expressed from pLST142 in such *Streptomyces* extracts (Fig. 4.8 track a) despite being expressed, apparently constitutively, from its own promoter *in vivo*. Presumably a factor required for expression of the gene was not present at the stage of growth at which active extracts can be made, or, alternatively, such a factor might have been inactivated during preparation.

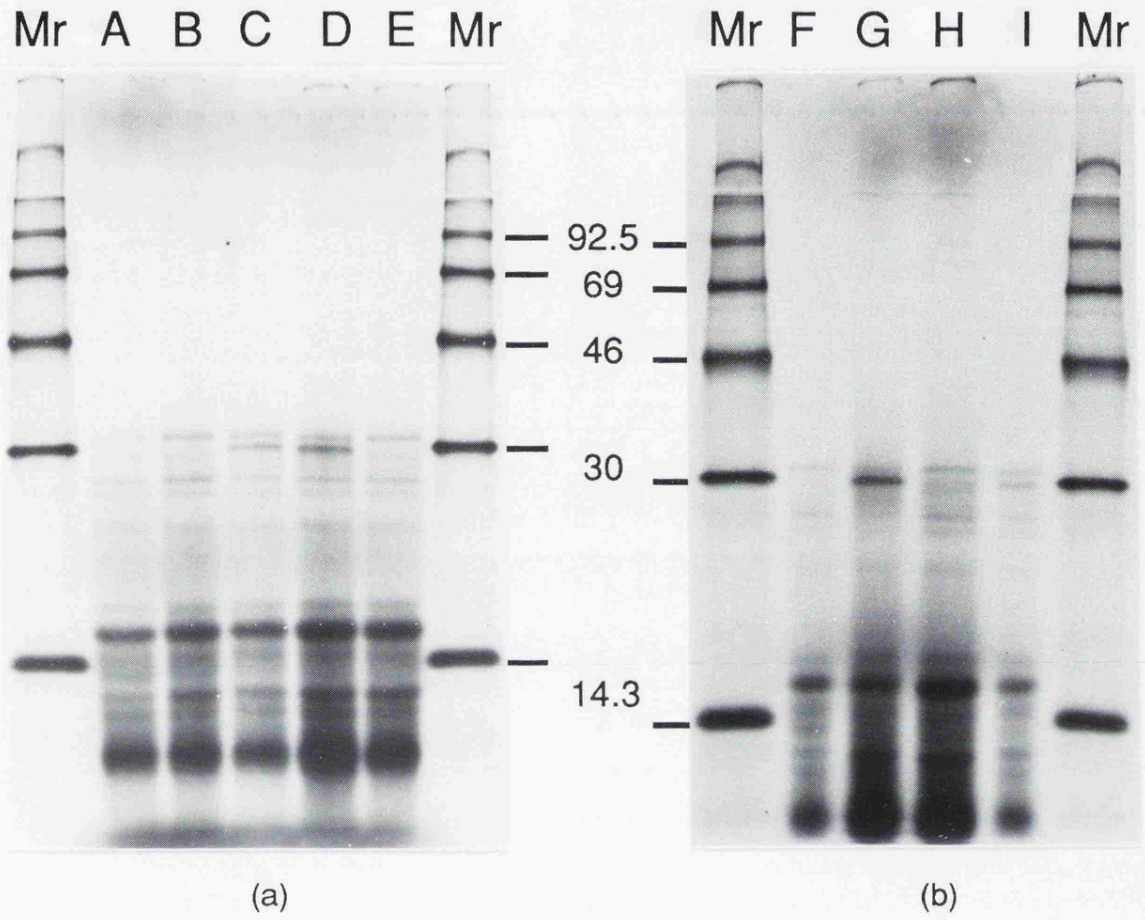


Fig. 4.8 Coupled Transcription-Translation *in vitro*.

(a) Electrophoretic analysis of the products of coupled transcription-translation directed by: (A) pUC18; (B) pLST142; (C) pLST143; (D) pLST144; and (E) pLST146. (b) Similar analysis using; (F) pUC18; (G) pLST144 linearised with *Sca*I; (H) pLST144 linearised with *Bgl*II; and (I) pLST144.
Products were subjected to electrophoresis on SDS-polyacrylamide gels together with [^{14}C] protein markers whose molecular weights (in kDa) are indicated.

Streptomyces genes that fail to programme coupled transcription-translation in these extracts have been observed before. A determinant from the bialaphos pathway was isolated from *S. hygroscopicus* and used to direct coupled transcription-translation in extracts of *S. lividans*, prepared after various growth periods. By analysis using polyacrylamide gels, the gene was apparently expressed only in late phase cultures at which time the activity of extracts is very low (C.J. Thompson, personal communication). Of course, for the most part, genes introduced into the coupled transcription-translation system are expressed, but it would not be surprising if some of the minor sigma factor(s), essential for the transcription of various determinants, were absent from extracts derived from relatively early phase cultures.

4.6. Transcript analysis of *kgmA*.

The transcriptional start site of *kgmA* was determined by high resolution S1 nuclease mapping, using as probe the *Bam*HI-*Bgl*II fragment of pLST142 (Fig. 4.1) radiolabelled with [γ-^{32}P] ATP at the *Bgl*II site, together with RNA from *S. lividans* JR14. The product of digestion was subjected to electrophoresis on a 7 M urea, 6% polyacrylamide gel together with a sequencing ladder (pUC18) as a size marker, and the suitable controls. The results show a band corresponding to 282 nucleotides in length in addition to a band equal in size to the intact probe (Fig. 4.9). The appearance of the latter was presumably due either to the presence of an additional promoter upstream of the *Bam*HI site or to re-annealing of the probe under the hybridisation conditions. The *kgmA* promoter (Fig. 4.10) has a

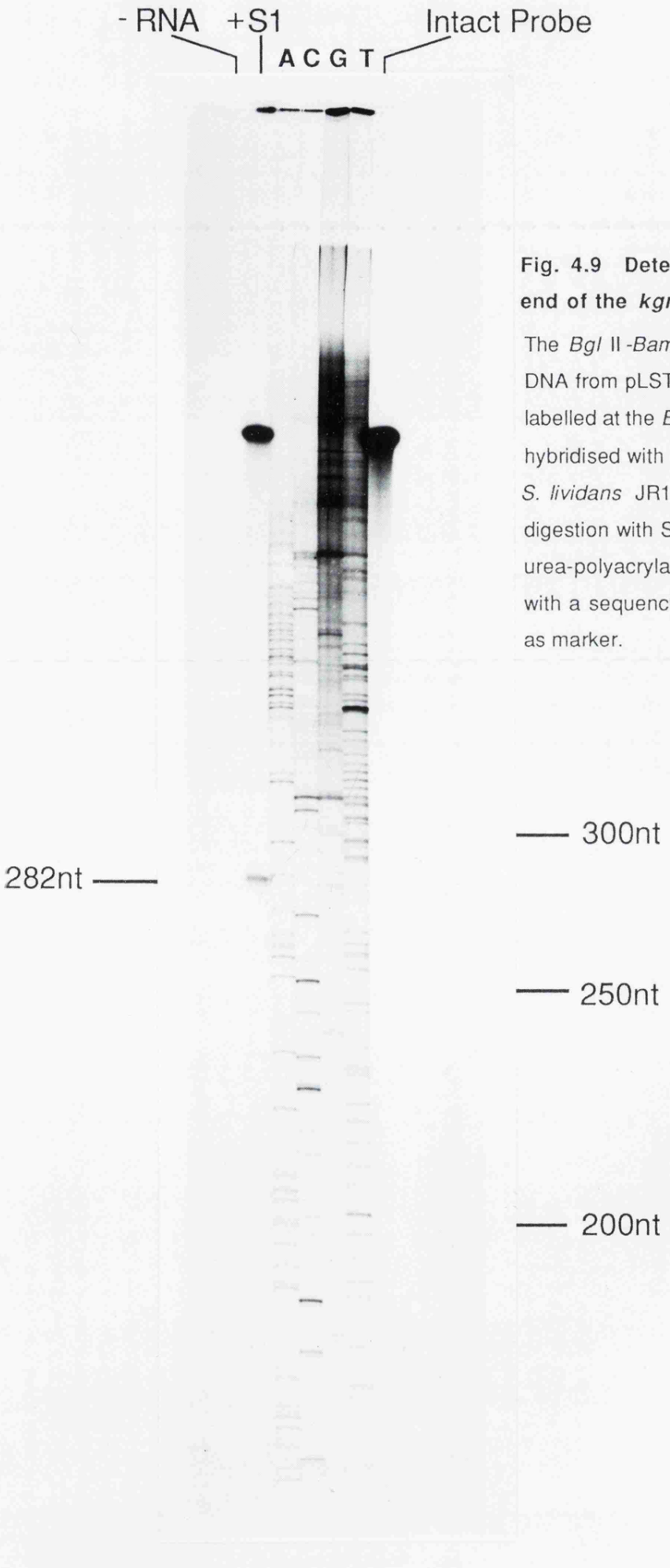


Fig. 4.9 Determination of the 5' end of the *kgmA* transcript.

The *Bgl* II-*Bam* HI fragment of DNA from pLST142 (see Fig.4.1) labelled at the *Bgl* II site was hybridised with RNA from *S. lividans* JR14. The product of digestion with S1 was run on a urea-polyacrylamide gel together with a sequencing ladder (pUC18) as marker.

recognisable -10 sequence but, like many *Streptomyces* promoters, it does not contain a sequence which is obviously similar to the prokaryotic consensus at -35. Some other promoters are shown alongside that for *kgmA*, however, comparisons of this nature are limited in value and a more logical system might be to classify promoters according to the sigma factor(s) that recognises them. This is now possible as several different forms of *Streptomyces* RNA polymerase holoenzyme have been identified (Westpheling *et al.*, 1985; Buttner *et al.*, 1988).

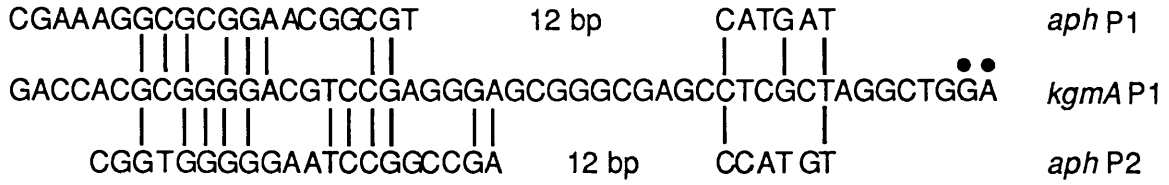


Fig. 4.10 Nucleotide sequence of the promoter directing transcription of *kgmA*.

The nucleotide sequence immediately upstream of the transcriptional start site (indicated by o) is shown and compared with those for the two promoters from the aminoglycoside phosphotransferase gene of *S. fradiae*.

The location of the transcript initiation point means that the *kgmA* gene has a leader sequence of 85 nucleotides containing a stable presumptive ribosome binding site at positions 245 to 252 in Fig. 4.5. Although, like those of *E. coli*, *Streptomyces* ribosomes do not require a particularly stable Shine-Dalgarno interaction, the putative ribosome binding site on the *kgmA* mRNA has a free energy of interaction with the 3' end of 16S rRNA of ΔG=-11.6 kcal mol^{-1} (Tinoco *et al.*, 1973). In *E. coli* the corresponding free energies span from ΔG=-4 to -22 kcal mol^{-1} (Gold *et al.*, 1981) whereas a more stringent interaction has been

observed in *Bacillus* with the free energies ranging between $\Delta G = -11.6$ and -21.0 kcal mol^{-1} (McLaughlin *et al.*, 1981).

An attempt was made to determine the 3' end of the transcript using, as probe, the 395 bp *Eco*RI-*Sph*I fragment of pLST144 (Fig. 4.1), radiolabelled at the *Eco* RI site using [α-^{32}P] dATP. However, when this was hybridised with RNA from *S. lividans* JR14, the only product of nuclease S1 digestion was equal in size to the intact probe (data not shown). The conclusion, therefore was that *kgmA* is transcribed beyond the proximal *Sph*I site in pLST14 (Fig. 4.1). An alternative explanation might have been that the probe failed to hybridise to the RNA. However, given that this method has been successfully employed in the context of another gene (see Chapter 5) and that the conditions for hybridisation and S1 nuclease treatment were identical to those used to determine the 5' end of the transcript, the former explanation is favoured.

Accordingly, the transcript apparently begins 85 nucleotides upstream of the translational start site and continues for more than 350 bases after the translational stop codon, past a large inverted repeat which fails to act as a terminator in *Streptomyces*.

Some features of note, at a purely superficial level, can be seen in the nucleotide sequence which may explain some of the phenomena associated with this DNA. The *kgmA* gene from pLST142 is not expressed in *E. coli*, either from its own promoter or from the *lacZ* promoter located further upstream. The possible reasons for the failure

of the gene to be expressed from the latter have already been discussed with the most likely candidate being termination of transcription prior to the completion of the *kgmA* mRNA. Factor independent termination of transcription requires the formation of stable secondary structure in the mRNA, often followed by a T-rich sequence (at least in *E. coli*). Within the DNA sequenced, prior to the *kgmA* translational start, there exists two stable inverted repeats (positions 60 to 218 in Fig. 4.5) capable of forming hairpin loops (as determined using the UWGCG programme FOLD (Zuker and Stiegler, 1981) (Fig. 4.11) which could conceivably act as transcriptional terminators, although additional sequences not determined, upstream of these features, could present a block to transcription. Alternatively, these inverted repeats might be recognised by RNA polymerase for the initiation of transcription in *Micromonospora* as could the numerous direct repeats which exist from position 62 to 304 in Fig. 4.5. Direct repeats in DNA sequences have been associated with a number of observations including protein-nucleic acid interaction in *E. coli* as well as being "hot spots" for DNA rearrangements (Streisinger *et al.*, 1986; Ripley *et al.*, 1985). The latter is particularly relevant since the deletion that leads to pLST143 results in the removal of all but two of these direct repeats.

Due to the G+C rich nature of *Streptomyces* DNA, the occurrence of inverted repeats, capable of forming stable hairpin loops is fairly frequent. Often, these are referred to as significant features both prior to and following open reading frames and while those preceding the coding sequence are "of unknown function" those after the structural

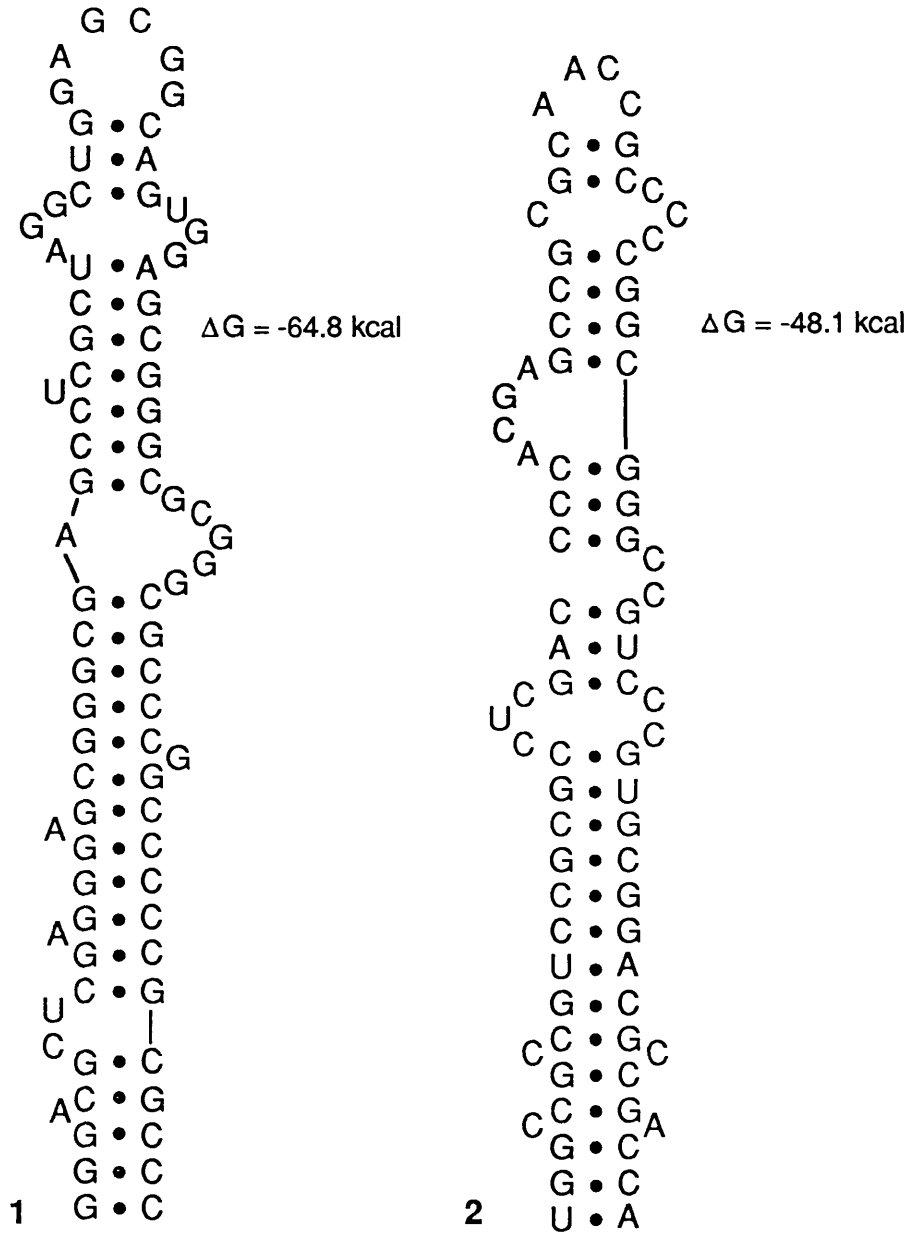


Fig. 4.11 **Inverted repeat structures upstream of *kgmA***

While inverted repeats are commonly observed upstream of coding sequences in *Streptomyces* they could present a barrier to transcriptional readthrough in *E. coli.* The structures shown might be formed if RNA polymerase initiating at the *lacZ* promoter was able to transcribe the DNA in pLST142. Perhaps they act as transcriptional terminators in that context.

genes are commonly referred to as "putative transcriptional terminator sequences" (Neal and Chater, 1987; Bibb *et al.*, 1985; Epp *et al.*, 1987; Tohyama *et al.*, 1987; Dehottay *et al.*, 1987; Long *et al.*, 1987).

4.7. An attempt to express *kgmA* in *Bacillus subtilis*.

Plasmids pC194 and pLST142 were linearised at their unique *Hin*dIII sites, ligated and used to transform *E. coli* NM522. Some ampicillin transformants were analysed and found to contain plasmids pC142a or pC142b (Fig. 4.12). These were used to transform *Bacillus subtilis* BD170 protoplasts, and chloramphenicol resistant colonies were analysed. These were confirmed as containing one of the above plasmids but were unable to grow on kanamycin or gentamicin. Since the ribosome binding site of *kgmA* should not preclude expression of this gene, and, if it were synthesised, the methylase ought to be able to modify *Bacillus* ribosomes, the simplest explanation is that this *Micromonospora* promoter is not recognised by any of the RNA polymerase holoenzymes in *B. subtilis* BD170. However, in order to establish any firm conclusions of this nature Northern hybridisation experiments would need to be performed using RNA from the *Bacillus* transformants to determine whether any *kgmA* transcript was detectable. However, since this experiment was designed to establish whether the *kgmA* promoter could act as such in *Bacillus*, it could perhaps more simply be cloned in an appropriate promoter probe plasmid.

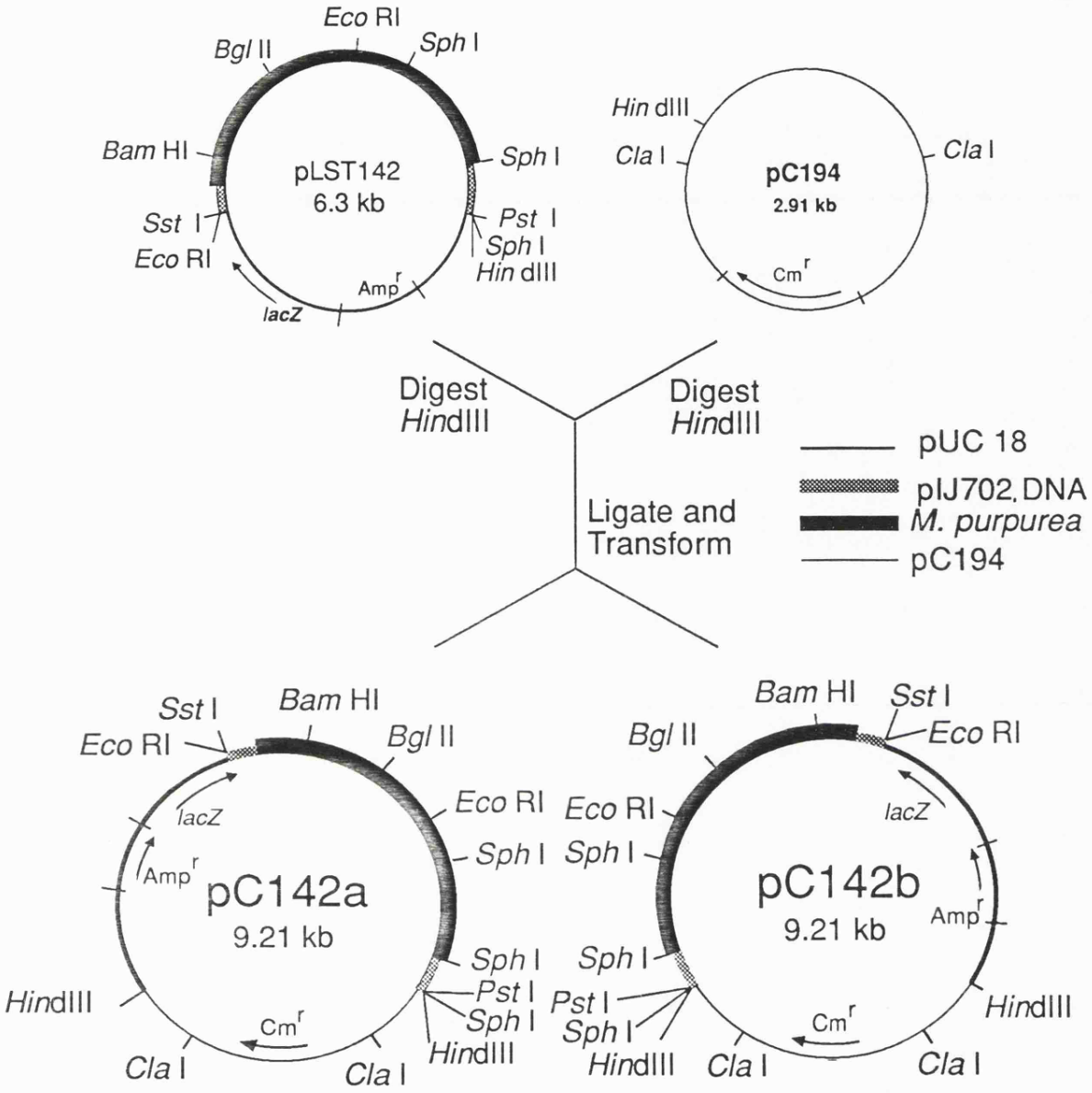


Fig. 4.12 An attempt to express *kgmA* in *B. subtilis*

Plasmid pLST142 was linearised with *Hin*dIII and ligated with the staphylococcal plasmid pC194 similarly restricted. The ligation was used to transform *E. coli* and the two plasmids pC142a and pC142b were identified. These were then used to transform *B. subtilis* BD170 and chloramphenicol resistant colonies were tested for *kgmA* expression. While the two plasmids appeared to be stable in *Bacillus*, they failed to confer kanamycin resistance.

The nucleotide sequence of *kgmA* represents the first of a gene isolated from a *Micromonospora* species. While the gene appears to be transcribed from a single promoter, it might well be part of a polycistronic message since the DNA is transcribed at least 350 bp beyond the translational stop codon.

The gene is expressed as a translational fusion in *E. coli* and, while conferring kanamycin resistance, has no other obvious effect. Since it has been shown that ribosomes, modified by this methylase, are refractory to the inhibitory effect of kanamycin on protein synthesis, it would be interesting perhaps to examine the ability of certain aminoglycosides to cause such ribosomes to mistranslate mRNA.

CHAPTER 5

NUCLEOTIDE ANALYSIS OF AN AMINOGLYCOSIDE RESISTANCE DETERMINANT FROM *S. tenjimariensis*

Introduction.

At approximately the same time that *kgmA* was isolated from *M. purpurea* another aminoglycoside resistance determinant (*kamA*), which conferred resistance to apramycin as well as kanamycin, was isolated from *S. tenjimariensis* (the istamycin producer) (Skeggs *et al.*, 1985). The phenotype of *S. lividans* TSK41, the primary clone harbouring *kamA*, has been discussed in detail previously (Chapters 3), and this introduction will focus on attempts (by P.A. Skeggs) to subclone and the subsequent localisation of the resistance gene in pLST41.

The original subcloning strategy involved partial digestion of pLST41 with *Sau*3A and ligation of the fragments with pIJ702 linearised with *Bgl*II. A smaller plasmid, pLST411, which conferred identical aminoglycoside resistance characteristics as did its parent plasmid was mapped with restriction enzymes (Fig. 5.1). An *Sph*I fragment was deleted from pLST411 to reduce the inserted DNA to less than 1 kb without interfering with the resistance profile, the smaller plasmid being designated pLST412 (Fig. 5.1). However, the subsequent mapping of pLST41 and comparison with maps already generated for pLST411 and pLST412 revealed the presence of a *Pst*I site in the subcloned plasmids not present in pLST41. There is only one such site in pLST41 and this is situated in the vector, not in the inserted DNA. The hypothesis then must be that pLST411 is the result of the ligation of at least 3 DNA fragments, one of which is the *Sau*3A piece containing the unique *Pst*I site of pIJ702. This would explain why, when pLST412 was digested with *Bcl*I and a Southern blot probed with the insert DNA

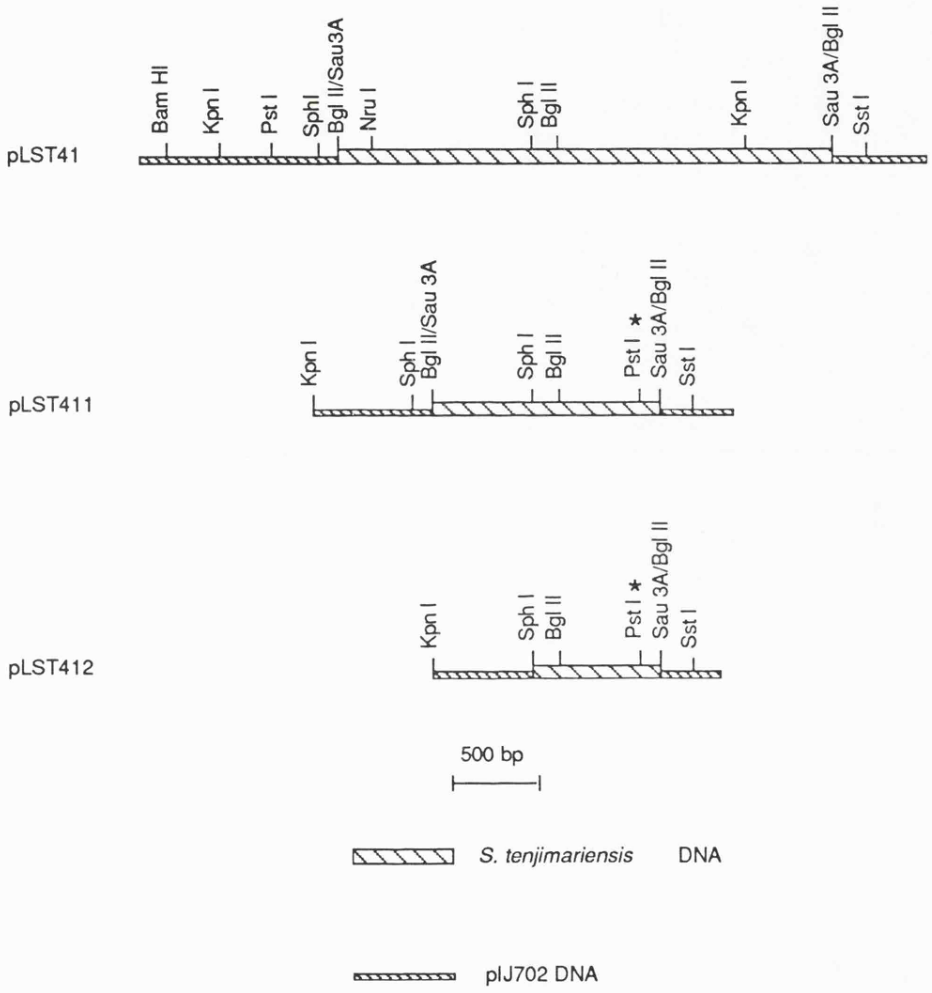


Fig. 5.1 Restriction maps of plasmids pLST41, pLST411, and pLST412. Plasmid pLST411 was generated by subcloning the DNA fragments generated by a partial *Sau* 3A digest of pLST41 in the *Bgl* II site of pIJ702. As can be seen, this resulted in the introduction of an extra *Pst* I site (marked ★) presumably derived from the vector.

of the plasmid, more than one band was observed after autoradiography (Skeggs, 1986). The quite reasonable explanation suggested at that time was that the probe might be contaminated with vector DNA, but we can now say that it is almost certainly because the "insert" itself contains vector sequences.

Nevertheless, pLST412 helped to localise *kamA* within pLST41, but, for the purpose of sequencing, a decision was made to isolate a fragment of the latter which presumably contains sequences contiguous in the *S. tenjimariensis* chromosome.

Results.

5.1. Subcloning and nucleotide sequence of *kamA*.

Since the plasmid pLST412 had pinpointed the site of *kamA* in pLST41, the *Sph*I-*Kpn*I fragment was subcloned in pUC18 and pUC19 (Fig. 5.2). Neither of the pUC derived constructs conferred kanamycin resistance on *E. coli* NM522, so further subclones were generated which contained the *Bgl*II-*Kpn*I fragment in pUC18 and pUC19 (Fig. 5.2) but again *kamA* was not obviously expressed. In order to confirm that *kamA* lay within the subcloned DNA, the *Sph*I-*Kpn*I fragment was introduced into pIJ702 and used to transform *S. lividans* TK21. Although kanamycin resistant transformants were obtained, the recombinant plasmids isolated from those examined had all undergone apparently identical but unresolved rearrangements. An alternative strategy was employed using the *Sph*I and *Sst*I sites to excise the insert from pLST413 (Fig. 5.2) which was similarly combined with

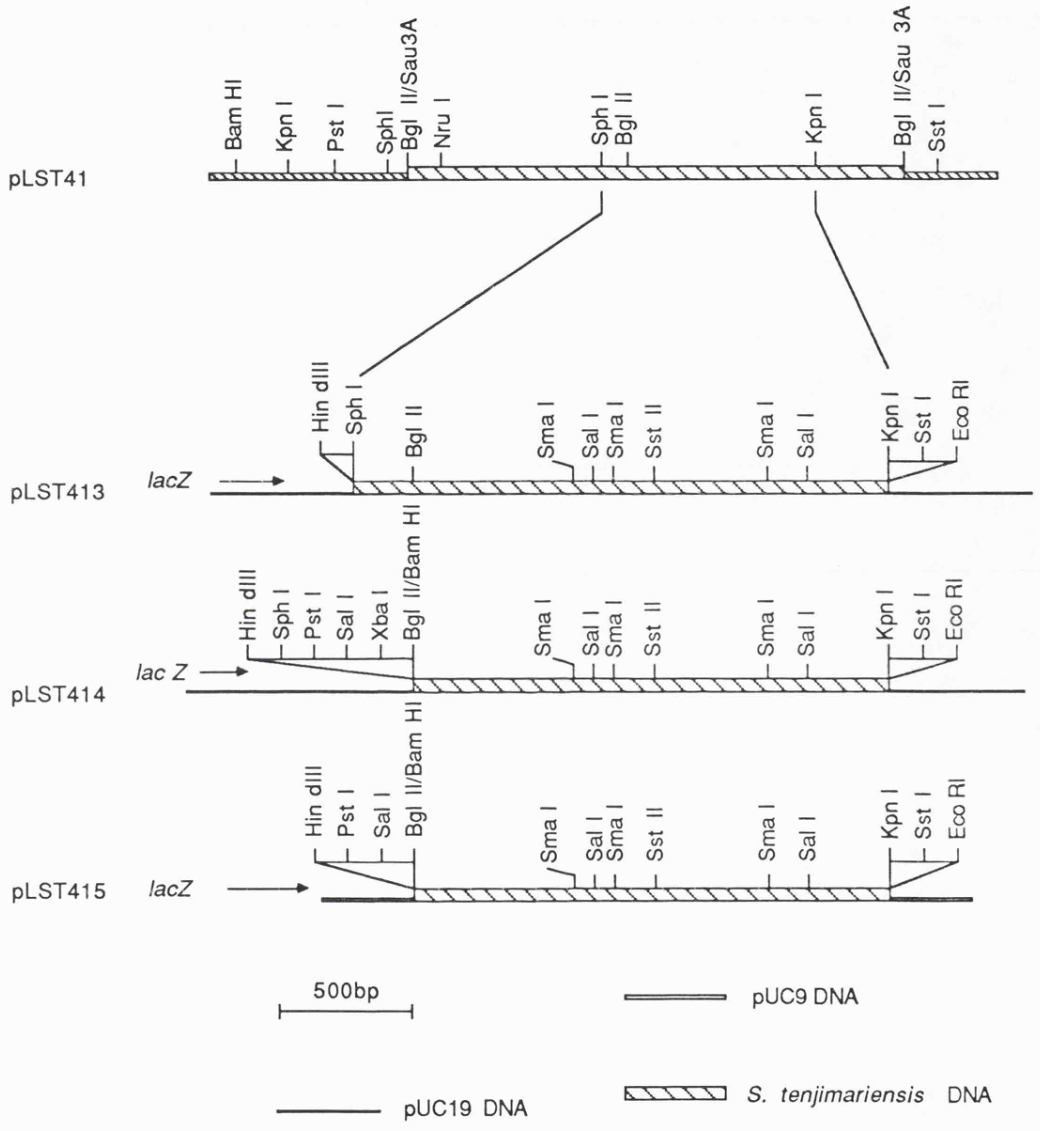


Fig. 5.2 Subclones generated from pLST41

Due to the uncertain nature of plasmid pLST412, other subclones were generated which contained contiguous sequences from the *S. tenjimariensis* genome. The insert in pLST413 was shown to embrace *kamA* (see Fig. 5.3).

pIJ702 and kanamycin resistant *Streptomyces* transformants were obtained. On this occasion, the anticipated recombinant plasmid was identified (pLST416), thereby confirming that *kamA* was contained within the subcloned DNA (Fig. 5.3).

The 1.6 kb *Sph*I-*Kpn*I insert in pLST413 was then sequenced according to the strategy shown in Fig 5.4 using the same dideoxy method and criteria described in the previous chapter. The sequence revealed an open reading frame from nucleotide 183 to 641 (Fig. 5.5) consisting of 155 codons that would code for a predicted protein of molecular weight 19.5 kDa. This open reading frame showed the biased codon usage normally associated with high G+C DNA (Table 5.1) as illustrated using the UWGCG programme CODONPREFERENCE (Gribovsky *et al.*, 1984) (Fig. 5.6). The overall G+C content of the coding sequence is 74.8% and those for codon positions 1,2, and 3 are 71.6%, 55.4% and 97.4% respectively.

In order to confirm that this, relatively small, sequence encoded the *kamA* methylase, the DNA was further manipulated. Plasmid pLST413 was linearised with *Bgl*II (a unique site at position 190 in Fig. 5.5) and the 5' protruding ends were filled in using DNA polymerase I (Klenow) and religated, introducing a frame shift in the coding sequence. This construct was transferred as an *Sph*I-*Sst*I fragment into pIJ702 and, as anticipated, it failed to confer kanamycin resistance on *S. lividans*.

It can be seen from the sequence that the *Sal*I sites at positions

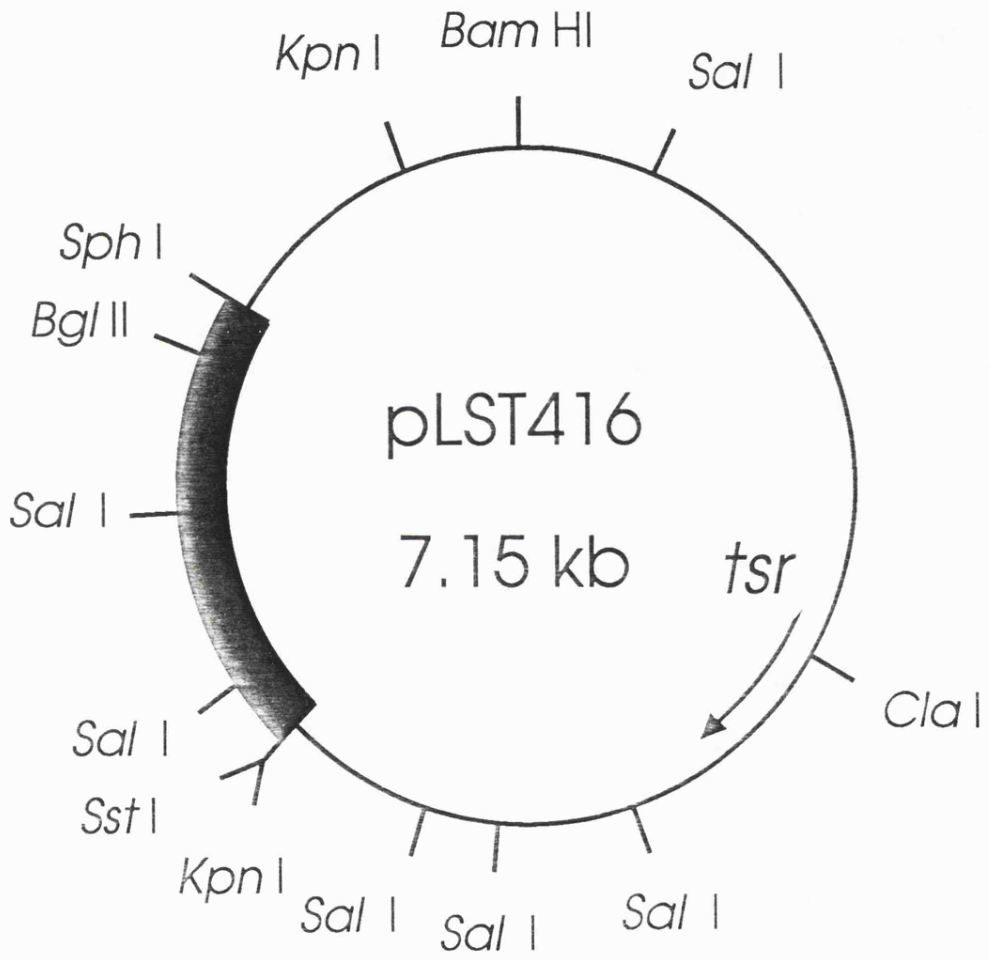


Fig. 5.3 Restriction endonuclease cleavage map of pLST416.

The insert in pLST413 (Fig. 5.2) was excised as an *Sph*I - *Sst*I fragment and inserted into pIJ702. This plasmid evidently contains *kamA* since it confers the same resistance profile on *S. lividans* as does pLST41.

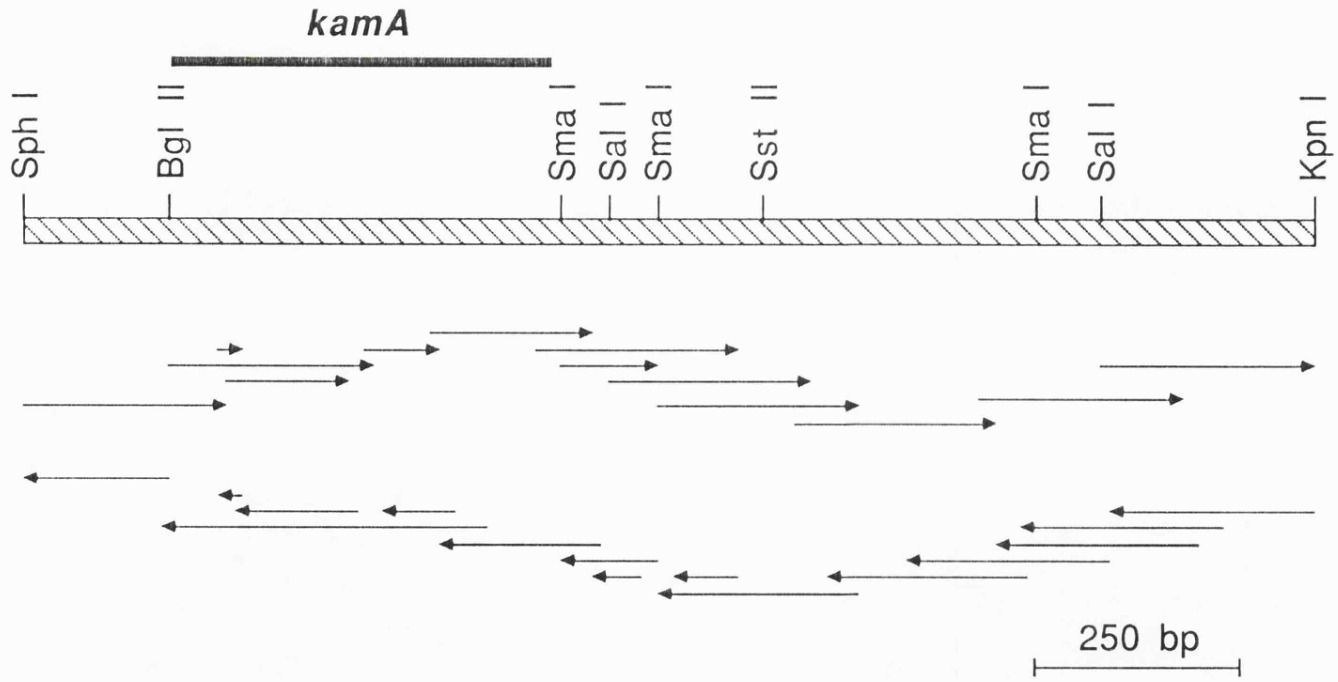


Fig. 5.4 Strategy for Determining the Nucleotide Sequence of *kamA*

Sequence was determined through all restriction sites used to generate subclones.

SphI
GCATGCGCCGCGTGGTGGGCAAGCGGGTCCAGGAGTTCTCCGACGCCGAGTTCGAGCAGC
CGTACGCGGCGCACCACCCGTTCGCCCAGGTCCTCAAGAGGCTGCGGCTCAAGCTCGTCG

TACGGAGTCAGTACGACGACGTGGTGCTCGACGTCGGCACCGGCGACGGGAAGCATCCGT
ATGCCTCAGTCATGCTGCTGCACCACGAGCTGCAGCCGTGGCCGCTGCCCTTCGTAGGCA

RBS
ACAAGGTCGCCCGCCAGAACCCCTCCCGGCTGGTGGTGGCGCTCGACGCCGACAAGAGCC
TGTTCCAGCGGGCGGTCTTGGGGAGGGCCGACCACCACCGCGAGCTGCGGCTGTTCTCGG

BglII
GGATGGAGAAGATCTCGGCGAAGGCGGCGGCCAAGCCCGCGAAGGCCCTGCCCAACCTGC
CCTACCTCTTCTAGAGCCGCTTCCGCCGCCGGTTCGGGCGCTTCCGGGACGGGTTGGACG

M E K I S A K A A A K P A K A L P N L L -

TGTACCTGTGGGCCACCGCCGAGCGGCTCCCCCCGTTGTCGGGGGTGGGCGAGCTGCACG
ACATGGACACCCGGTGGCGGCTCGCCGAGGGGGGCAACAGCCCCCACCCGCTCGACGTGC

Y L W A T A E R L P P L S G V G E L H V -

TCCTCATGCCGTGGGGCAGCCTGCTGCGCGGGGTCCTCGGCTCCTCGCCGGAGATGCTGC
AGGAGTACGGCACCCCGTCGGACGACGCGCCCCAGGAGCCGAGGAGCGGCCTCTACGACG

L M P W G S L L R G V L G S S P E M L R -

GCGGGATGGCGGCGGTGTGCCGGCCGGGCGCGTCCTTCCTGGTCGAGTGCAACCGCAGCT
CGCCCTACCGCCGCCACACGGCCGGCCCGCGCAGGAAGGACCAGCTCACGTTGGCGTCGA

G M A A V C R P G A S F L V E C N R S C -

ApaI
GCCGTGGGCCCCCGGTGCCGGAGGTGGGCGAGCACCCCGAGCCCACCCCGGACTCCGCCG
CGGCACCCGGGGGCCACGGCCTCCACCCGCTCGTGGGGCTCGGGTGGGGCCTGAGGCGGC

R G P P V P E V G E H P E P T P D S A D -

```
ACGAGTGGCTGGCGCCCCGCTACGCCGAGGCCGGGTGGAAGCTCGCCGACTGCCGCTACC
---------+---------+---------+---------+---------+---------+ 540
TGCTCACCGACCGCGGGGCGATGCGGCTCCGGCCCACCTTCGAGCGGCTGACGGCGATGG

  E  W  L  A  P  R  Y  A  E  A  G  W  K  L  A  D  C  R  Y  L  -

TGGAGCCGGAGGAGGTGGCGGGTCTGGAGACCTCCTGGACCCGCCGTCTGCACTCCTCCC
---------+---------+---------+---------+---------+---------+ 600
ACCTCGGCCTCCTCCACCGCCCAGACCTCTGGAGGACCTGGGCGGCAGACGTGAGGAGGG

  E  P  E  E  V  A  G  L  E  T  S  W  T  R  R  L  H  S  S  R  -

                                    (Sau3A)                SmaI
GCGACCGGTTCGACGTGCTCGCGCTCACCGGCACGATCAGTCCGTGACGCCGGACCCCGG
---------+---------+---------+---------+---------+---------+ 660
CGCTGGCCAAGCTGCACGAGCGCGAGTGGCCGTGCTAGTCAGGCACTGCGGCCTGGGGCC

  D  R  F  D  V  L  A  L  T  G  T  I  S  P  *

                                                       SalI
GCGTCGGGCGGGCGTGAACGTCCGCCGCGCGCCGGGGGCCAGCTCGACGCGCGTGTCGTC
---------+---------+---------+---------+---------+---------+ 720
CGCAGCCCGCCCGCACTTGCAGGCGGCGCGGCCCCCGGTCGAGCTGCGCGCACAGCAG

                                                    SmaI
GACCCGGACCGGCAGCGGCACGCGTTGTCCGCGTCGGCCAGCAACTCCACGGTCCCCGGG
---------+---------+---------+---------+---------+---------+ 780
CTGGGCCTGGCCGTCGCCGTGCGCAACAGGCGCAGCCGGTCGTTGAGGTGCCAGGGGCCC

                                    ↓   ↓↓
GTGAGCACCACGGTCAACCCCCGGCCACGCCAACCGGCAGCGGAACCCGCAAGCCGGGTC
---------+---------+---------+---------+---------+---------+ 840
CACTCGTGGTGCCAGTTGGGGGCCGGTGCGGTTGGCCGTCGCCTTGGGCGTTCGGCCCAG

ACGCCCTCCGGCAGCAGGGGGGCGAACCGCAGCTCCTCGCCGGCCGGGTCCGCGCCGAGG
---------+---------+---------+---------+---------+---------+ 900
TGCGGGAGGCCGTCGTCCCCCCGCTTGGCGTCGAGGAGCGGCCGGCCCAGGCGCGGCTCC

    SstII
AAGCCGCGGACAGCACCAGCCAGGCGCCGGCCGCCGCTCCGAGGTGGACGCCGTTGCGGA
---------+---------+---------+---------+---------+---------+ 960
TTCGGCGCCTGTCGTGGTCGGTCCGCGGCCGGCGGCGAGGCTCCACCTGCGGCAACGCCT

AGTTCTCCCTGGGCGCGAAGTCGAGGTTGTAGCGGGCGGACCCGGCGAGCAGCCGTGGCC
---------+---------+---------+---------+---------+---------+ 1020
TCAAGAGGGACCCGCGCTTCAGCTCCAACATCGCCCGCCTGGGCCGCTCGTCGGCACCGG
```

```
AGCGGGACGTCGCCGGCCCACGCCGCCGCGATGGCGTGCGGTCACTGGGGCTCAGCGACT
---------+---------+---------+---------+---------+---------+ 1080
TCGCCCTGCAGCGGCCGGGTGCGGCGGCGCTACCGCACGCCAGTGACCCCGAGTCGCTGA

                                                      (HinfI)
GCCGGTGCGCCGTGCGGGGCTCGTAGTACCGCAGGTTCCGTTCGATGTCGGAACGCGAGA
---------+---------+---------+---------+---------+---------+ 1140
CGGCCACGCGGCACGCCCCGAGCATCATGGCGTCCAAGGCAAGCTACAGCCTTGCGCTCT

ATCGTTCGGACAACAACACCATCTCTGGACGACGTCGGCCTGTTTGACCAGTCCGCCCGA
---------+---------+---------+---------+---------+---------+ 1200
TAGCAAGCCTGTTGTTGTGGTAGAGACCTGCTGCAGCCGGACAAACTGGTCAGGCGGGCT

                                                  SmaI
CCGGCGCTCGTCGTTCTCCCTTCACCGGGAGGGGGAGACTTCTGACCCCGGGTGGCGCTCG
---------+---------+---------+---------+---------+---------+ 1260
GGCCGCGAGCAGCAAGAGGGAAGTGGCCCTCCCCCTCTGAAGACTGGGCCCACCGCGAGC

CACGTCTCGTCGGGCAGTGAGAAGTACCCGTCGAAGACCGGGACCACACCCGTCCCTGGT
---------+---------+---------+---------+---------+---------+ 1320
GTGCAGAGCAGCCCGTCACTCTTCATGGGCAGCTTCTGGCCCTGGTGTGGGCAGGGACCA

                                  SalI
CGGCGGGAGATACACCTCCCCCGCCCAGTCGACCCGATTTCTGGAGAAGCGTCCCGATGA
---------+---------+---------+---------+---------+---------+ 1380
GCCGCCCTCTATGTGGAGGGGGCGGGTCAGCTGGGCTAAAGACCTCTTCGCAGGGCTACT

TTTCCCGCCGCTCCAGCGTCCGCCAGGACCGACGGATTCCGGGCGGACAGGTCGAGAAGG
---------+---------+---------+---------+---------+---------+ 1440
AAAGGGCGGCGAGGTCGCAGGCGGTCCTGGCTGCCTAAGGCCCGCCTGTCCAGCTCTTCC

GTTTCCGCCGCCCAGGACAGGCCCCAGCCCGCGAGGAGGTTGGTGCCGAAATGGTGATCG
---------+---------+---------+---------+---------+---------+ 1500
CAAAGGCGGCGGGTCCTGTCCGGGGTCGGGCGCTCCTCCAACCACGGCTTTACCACTAGC

ACGTCGTAATGGAACTCGTCGAACCCCATCACGCCACGCACGGTCCGCCCGCCCACCTCG
---------+---------+---------+---------+---------+---------+ 1560
TGCAGCATTACCTTGAGCAGCTTGGGGTAGTGCGGTGCGTGCCAGGCGGGCGGGTGGAGC
```

```
                                KpnI
GGGTCGAACCGGAGCAGCCCGGCGAGGGTACC
---------+---------+---------+-- 1592
CCCAGCTTGGCCTCGTCGGGCCGCTCCCATGG
```

Fig. 5.5 Nucleotide sequence of DNA embracing kamA together with the predicted amino acid sequence (below) and the positions of key restriction sites (above). The putative ribosome binding site (RBS) is overlined and the transcription initiation sites are indicated (O). The Sau3A and HinfI sites used to generate a probe in order to determine the 3′ end of the transcript are shown in brackets because these are not the only sites for those enzymes. The 3′ end of the transcript is located using a downward arrow (↓).

Table 5.1. Codon usage of *kamA* .

aa	codon		aa	codon		aa	codon		aa	codon	
Phe	TTT	0	Ser	TCT	0	Tyr	TAT	0	Cys	TGT	0
	TTC	2		TCC	6		TAC	3		TGC	4
	TTA	0		TCA	0	Term	TAA	0	Term	TGA	1
Leu	TTG	1		TCG	3		TAG	0	Trp	TGG	5
Leu	CTT	0	Pro	CCT	0	His	CAT	0	Arg	CGT	2
	CTC	6		CCC	7		CAC	3		CGC	7
	CTA	0		CCA	0	Gln	CAA	0		CGA	0
	CTG	13		CCG	9		CAG	0		CGG	3
Ile	ATT	0	Thr	ACT	0	Asn	AAT	0	Ser	AGT	1
	ATC	2		ACC	5		AAC	2		AGC	2
	ATA	0		ACA	0	Lys	AAA	0	Arg	AGA	0
Met	ATG	4		ACG	1		AAG	5		AGG	0
Val	GTT	0	Ala	GCT	0	Asp	GAT	0	Gly	GGT	1
	GTC	3		GCC	8		GAC	5		GGC	6
	GTA	0		GCA	0	Glu	GAA	0		GGA	0
	GTG	6		GCG	10		GAG	14		GGG	5

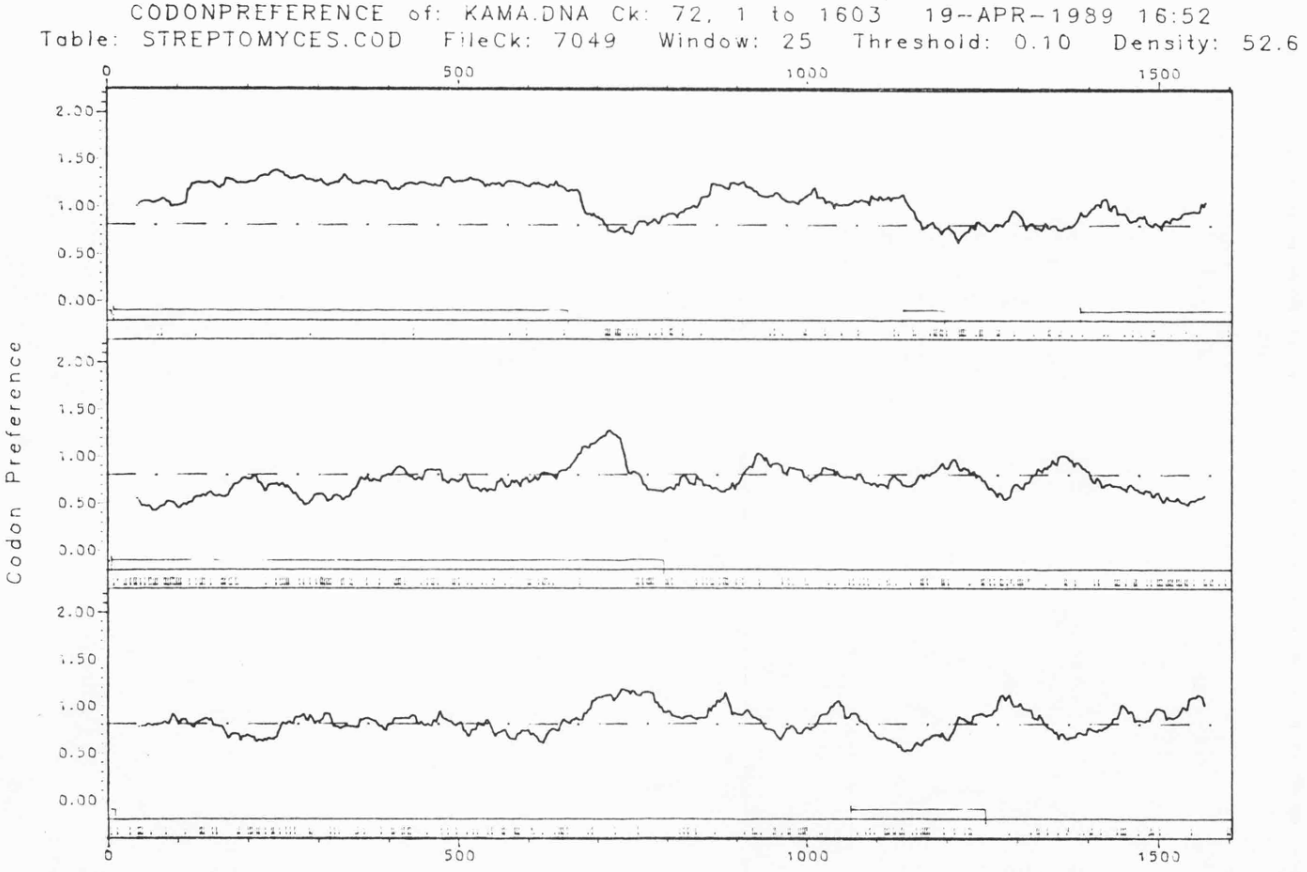


Fig. 5.6 Analysis of the sequence of *kamA* using the UWGCG programme CODONPREFERENCE

The analysis used the default settings (i.e. a moving window of 25 codons) and a codon usage table based on a number of *Streptomyces* gene sequences. Open reading frames are shown as boxes below their respective reading frames. "Rare" codons are marked below the reading frame plot. Open reading frames showing the biased codon usage typical of *Streptomyces* is revealed by a codonpreference index of › 0.8.

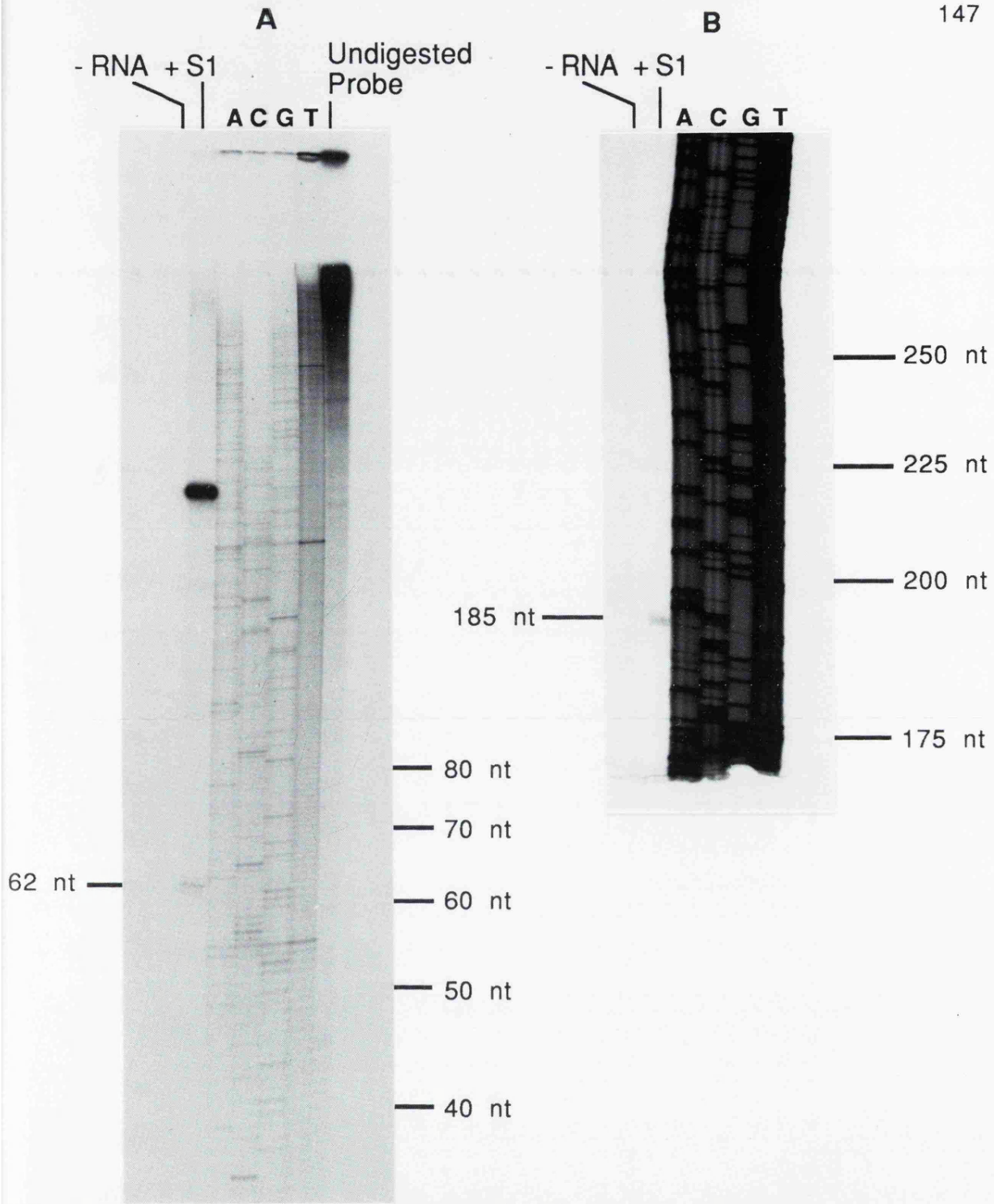


Fig. 5.7 Determination of the 5' end of the *kamA* transcript.

(A) The *Bgl* II *Sph* I fragment of pLST413 (see Fig. 5.2) labelled at the *Bgl* II site was hybridised with RNA from *S. lividans* TSK41. The products of digestion were run on a urea-polyacrylamide gel together with a sequencing ladder as marker (pUC18).

(B) A second promoter was identified by hybridising the *Bgl* II - *Pst* I fragment of pLST41, labelled at the *Bgl* II site, with RNA from *S. lividans* TSK41 and subsequent digestion with nuclease S1. The products were run on a urea-polyacrylamide sequencing gel together with a sequencing gel as marker (pUC9).

720 and 1350 in Fig. 5.5 lie outside the proposed coding sequence of *kamA* (see also Fig. 5.2). Accordingly, this fragment was deleted from pLST413 and the reduced insert was transferred to *S. lividans* as an *Sph*I-*Sst*I fragment in pIJ702. Predictably, this construct, pLST417, was able to confer kanamycin and apramycin resistance to *Streptomyces*.

5.2. Transcript analysis of *kamA*.

The transcriptional start site was determined by high resolution S1 nuclease mapping, using as probe the *Bgl*II-*Sph*I fragment of pLST413 (Fig. 5.2) radiolabelled with [γ-^{32}P] ATP at the *Bgl*II site together with RNA from *S. lividans* TSK41. The product of digestion was subjected to electrophoresis on a 7M urea, 6% polyacrylamide gel together with a sequencing ladder (pUC18) as a size marker and with suitable controls. The results show a band corresponding to 62 nucleotides in length in addition to a band equal in size to the intact probe (Fig. 5.7). The latter was presumably due either to the presence of an additional promoter upstream of the *Sph*I site or to reannealing of the probe under the hybridisation conditions. In order to test this, a second, larger, probe was used, specifically the *Bgl*II-*Pst*I piece of pLST41 (Fig. 5.1) again labelled at the *Bgl*II site with [γ-^{32}P] ATP. When this was hybridised with RNA from *S. lividans* TSK41, and treated with S1 nuclease and the products analysed by gel electrophoresis, a band of 185 nucleotides was observed. So a second promoter exists (*kamA*P2), 176 bp upstream of the translational start site the sequence of which, determined on one strand only, is shown together with that of *kamA*P1 in Fig. 5.8.

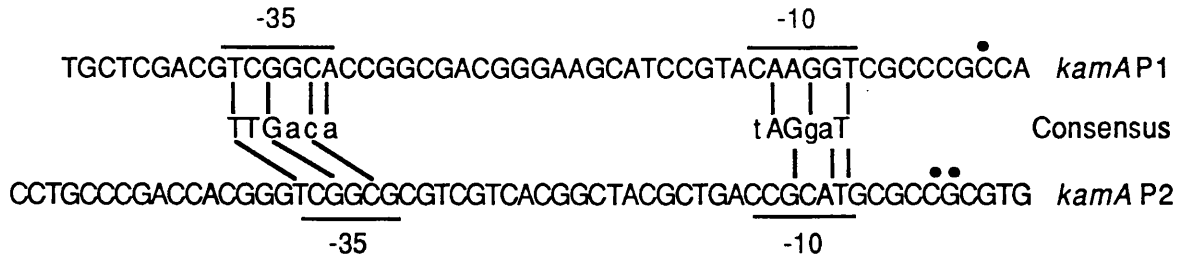


Fig. 5.8 The sequences of the promoters that direct transcription of *kamA*.

The putative "-35" and "-10" regions are indicated. They are separated by 22 bp in the case of *kamA*P1 and 20 bp in that of *kamA*P2.

Given that the transcript initiating at *kamA*P1 is only 62 bp upstream of the *Bgl*II site shown to be within the structural gene (coding sequence), the translational start site must lie between these two points. Only 3 possible initiation codons exist two of which are GTG codons (at positions 153 and 156 in Fig. 5.5), the other an ATG (position 183 in Fig. 5.5). Neither of the GTG codons have a Shine-Dalgarno sequence preceding them and although the ATG codon does have a sequence upstream which is complementary to the 3' end of 16S rRNA, the interaction would not be a particularly strong one.

The transcriptional termination site was determined using, as probe, the 395 bp *Sau*3A-*Hinf*I fragment of pLST413 labelled at the *Sau*3A site using [α-^{32}P] dATP and purified from a polyacrylamide gel, together with RNA from *S. lividans* TSK41. After treatment with S1 nuclease and analysis on a 7 M urea, 6% polyacrylamide gel a band of approximately 175 bp was observed (Fig. 5.9).

The UWGCG programme, TERMINATOR (Brendel and

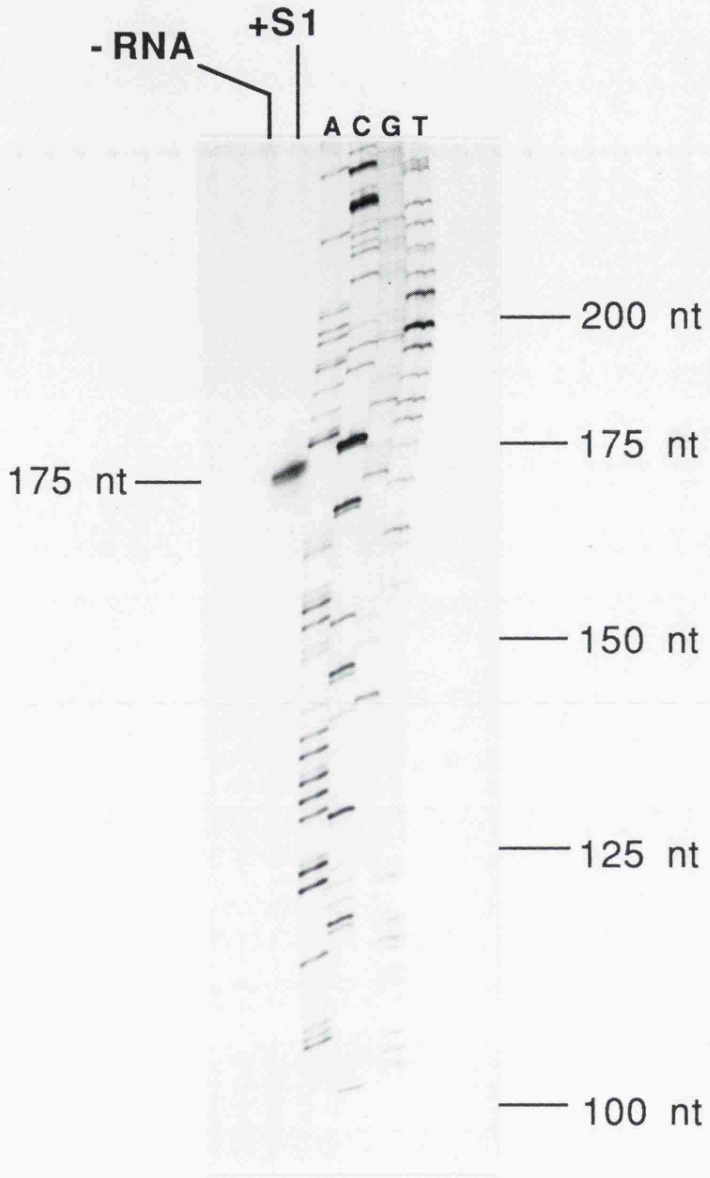


Fig. 5.9 Determination of the site of termination of transcription

A *Sau* 3A - *Hin* fl fragment of pLST413 (see text) was labelled at the *Sau* 3A site using [α ^{32}P] dATP and hybridised with RNA from *S. lividans* TSK41. the products of digestion with S1 were run on a urea-polyacrylamide gel together a sequencing ladder as size marker.

Trifonov, 1984), predicted a transcript termination site at position 675 in Fig. 5.5, after the large inverted repeat immediately following the transcriptional stop codon. However, the 3' end of the transcript maps about 149 bp downstream of this site. The mRNA between the translational and transcriptional termination points can be folded into a series of secondary structures (Fig. 5.10), which may explain why the transcript terminates at that point.

5.3. Expression of *kamA* in *E. coli.*

Preliminary attempts to express *kamA* in *E. coli* were (as described earlier) unsuccessful. With knowledge of the sequence of the *Sph*I-*Kpn*I fragment the possible reasons for this failure can be suggested. Evidently, the gene cannot be expressed from its own promoter, so that expression might only be achieved by using *E. coli* signals. Since the gene is read in the direction *Sph*I → *Kpn*I, a transcriptional fusion could be achieved by insertion of this fragment in pUC19. However, as described earlier, such a construct did not confer kanamycin resistance on *E. coli* possibly because RNA polymerase initiating at *lacZ* failed to transcribe the entire gene. A construct which resulted in translational fusion was possible since the *Bgl*II site in *kamA* is located at the N-terminus (cutting within the third codon in fact). The *Bgl*II-*Kpn*I fusion in pUC19 (pLST414) could not confer resistance because, in this construct, *kamA* is not in frame with *lacZ* (Fig. 5.11). Therefore, in an attempt to produce a fusion protein, the *Bgl*II-*Kpn*I fragment was ligated with pUC9 restricted with *Bam*HI and *Kpn*I in which *kamA* was in frame with *lacZ* (Fig. 5.11). The fusion was verified by DNA sequencing, but *E. coli* strains harbouring such a

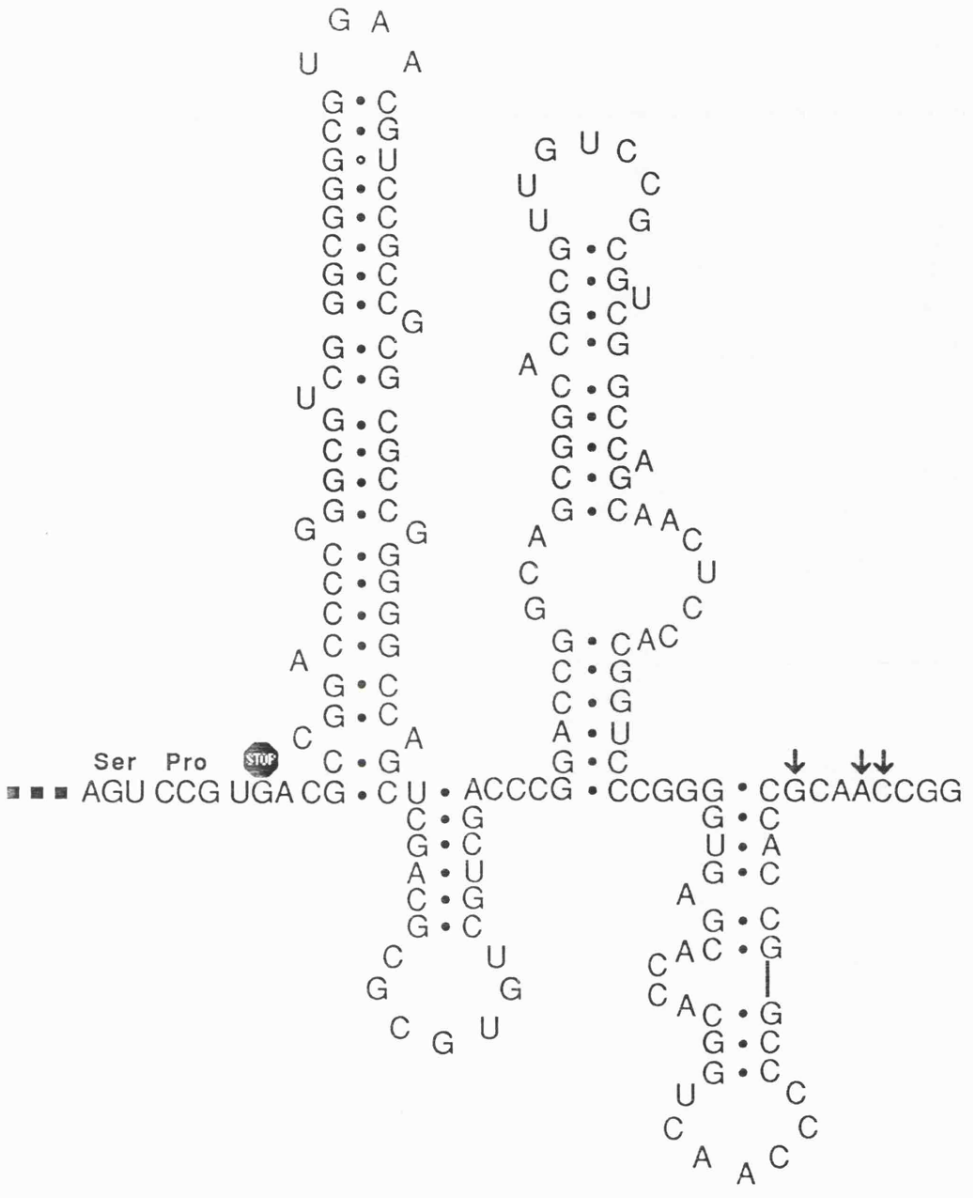
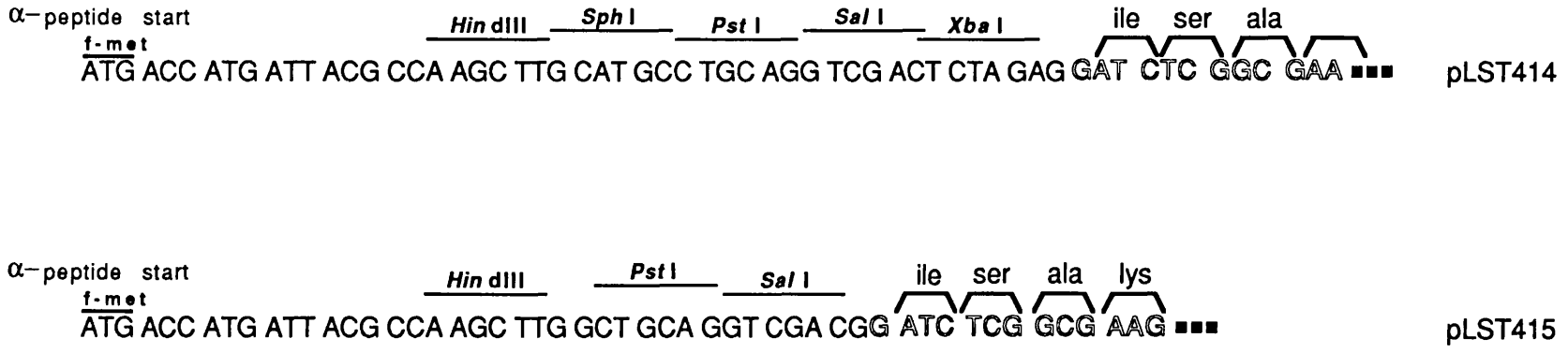


Fig 5.10 A possible secondary structure for the untranslated 3' end of the *kamA* mRNA.

The sites of termination of transcription, as determined by S1 mapping, are indicated by downward arrows.



GATC pUC sequence

GATC *kamA* sequence

Fig. 5.11 Nucleotide sequence of the translational fusions generated in pLST414 and pLST415.

In an attempt to express *kamA* in *E. coli* a plasmid was constructed in which the gene was fused to the *lacZ'* of pUC9. A fusion at the *Bgl* II site of *kamA* removes only 3 amino acids from the methylase and adds 13 aminoacids to the N-terminal. ⌐¬ indicates the reading frame of *kamA.*

plasmid were not resistant to kanamycin even in the presence of an inducer of *lacZ* (i.e. IPTG).

Without other evidence, the most obvious suggestions are that either the first three amino acids are essential for an active methylase or that the addition of the 13 N-terminal amino acids of *lacZ* are deleterious to enzyme function.

5.4. An explanation for pLST412.

Plasmid pLST412 was discussed in the introduction at a superficial level but with the publication of the sequence of pIJ101 (Kendall and Cohen, 1988) and knowledge of the sequence of pLST413, more precise comments can be made concerning the nature of pLST412.

The first suggestion that plasmid pLST412 might not be what it seemed came from the observation that it possessed an extra *Pst*I site which presumably was derived from pIJ702. According to the published sequence, the *Pst*I site of pIJ702 lies on a 253 bp *Sau*3 A fragment, which, together with the estimated size of the insert in pLST412 (0.8-0.9 kb) allows the construction of a restriction map for pLST412 (Fig. 5.12).

Interestingly, the construction requires the ligation of the "extra, pIJ702 derived, *Sau*3A fragment" to the *Sau*3A site at position 634 (Fig. 5.5) which is within the coding sequence of *kamA*. As a consequence, the methylase encoded by pLST412 is the result of a

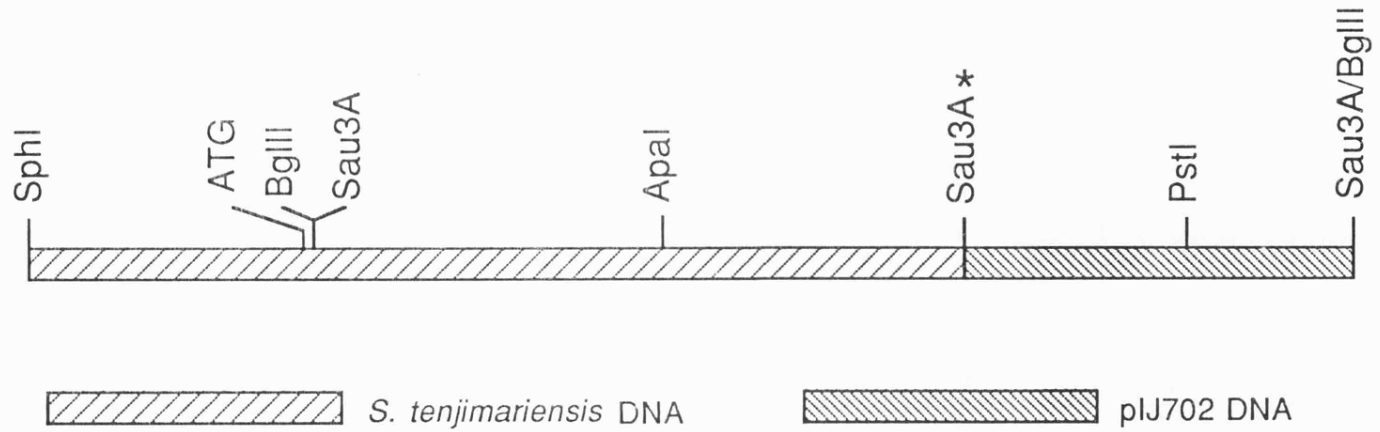


Fig. 5.12 Restriction map of pLST412

This plasmid is the result of the ligation of at three DNA molecules. These include pIJ702, a fragment of *S. tenjimariensis* DNA and the 257bp *Sau* 3A fragment of pIJ702 that contains a *Pst* I site. Surprisingly, the aminoglycoside resistance methylase encoded by this plasmid is probably the result of a translational fusion since the *Sau* 3A site (*) lies 10 bp inside the *kamA* coding sequence.

translational fusion between all of the *kamA* protein, except the 2 carboxyterminal amino acids, and additional amino acids encoded by the "extra *Sau*3A fragment".

5.5. Another member of the *kam* family.

A gene has been isolated from *Saccharopolyspora hirsuta*, the fortimicin producer, which apparently confers identical aminoglycoside resistance pattern as *kamA* (G. Tiraby, personal communication). Moreover, this gene has been subcloned and expressed in *E. coli* as well as its sequence determined. The nucleotide sequence has been kindly made available by Prof. Tiraby, prior to publication. Without further data the only available method for determining the open reading frame coding for the resistance gene would be to employ a computer to detect codon bias in each of the six frames. Since the gene from *Sacc. hirsuta* is expressed in *E. coli* utilising transcriptional signals from the Gram-negative host, computer analysis identified three possible translational start codons.

The restriction maps of the *S. tenjimariensis* DNA embracing *kamA* and that of *Sacc. hirsuta* were compared but they appeared to be totally different (Fig. 5.13). Be that as it may the nucleotide sequence of *kamA* was compared with that of the resistance gene from *Sacc. hirsuta* using the UWGCG programme COMPARE (Maizel and Lenk, 1981) and represented diagrammatically using the programme DOTPLOT (Fig. 5.14). The analysis revealed a high degree of similarity between the region of *S. tenjimariensis* DNA which codes for *kamA* and that which is purported to encode the *Sacc. hirsuta* gene,

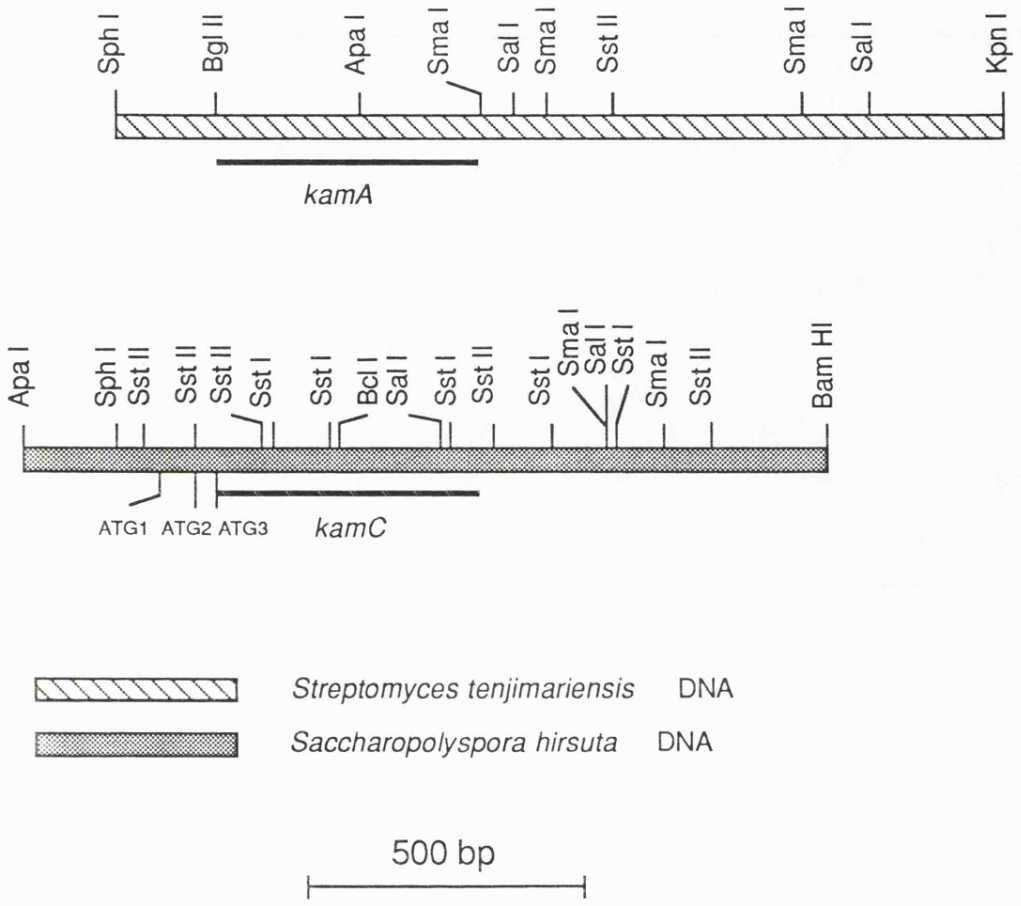
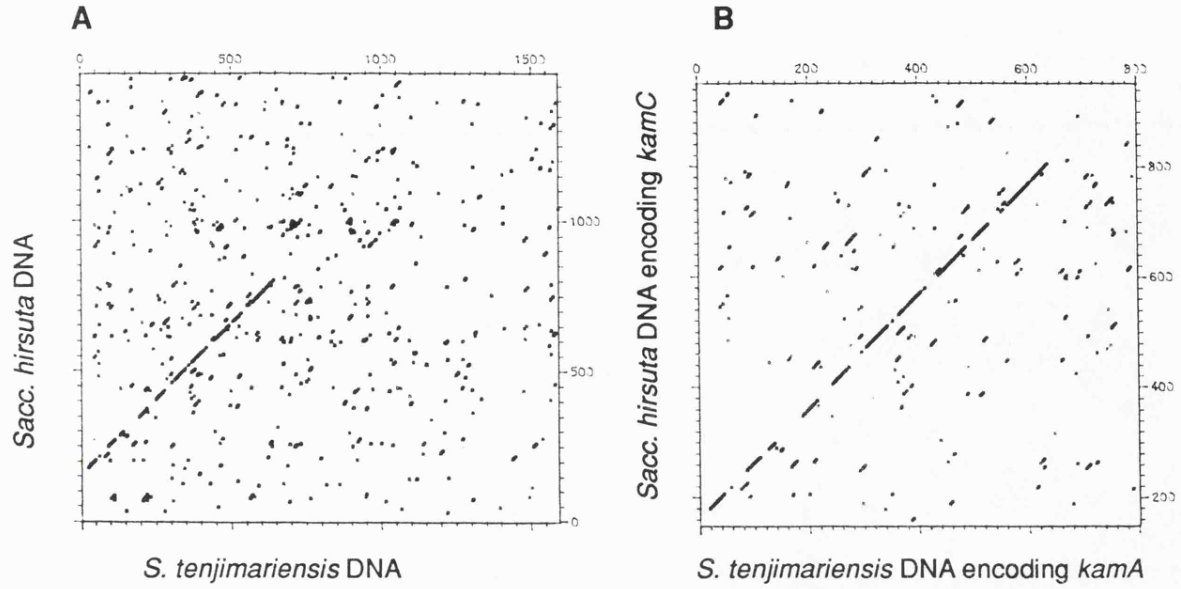


Fig. 5.13 Restriction map of DNA fragments containing kanamycin-apramycin resistance determinants.

A resistance determinant that confers the same resistance pattern as does *kamA* was isolated from *Sacc. hirsuta.* However, comparison of the restriction maps showed that there was no similarity, at least at this crude level of analysis. On the other hand, more detailed study revealed that the two resistance genes were very likely to encode methylases with identical modes of action (see Figs. 5.14 & 5.15). The positions of three possible translational initiation sites for *kamC* are indicated (ATG1-3).



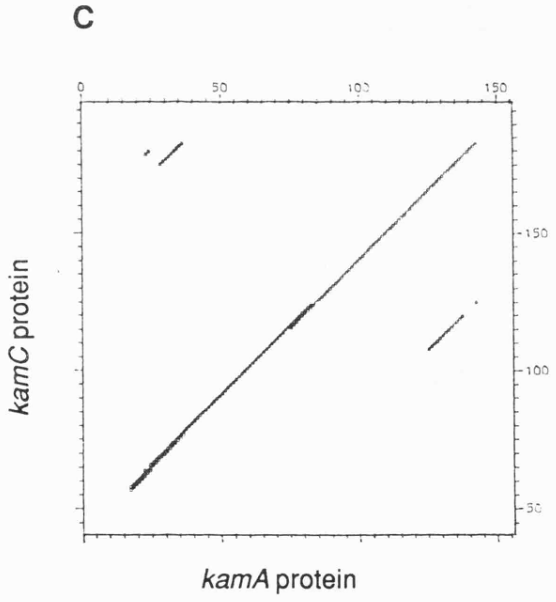


Fig. 5.14 Comparison of the DNA sequences of *kamA* and *kamC*, and the proteins they encode using the UWGCG programmes COMPARE and DOTPLOT.

(A) Comparison of the total *S. tenjimariensis* and *Sacc. hirsuta* DNA sequenced.
(B) Comparison of the *kamA*, *kamC* and upstream sequences.
(C) Comparison of the predicted amino acid sequences.
All comparisons used the default settings, i.e. for (A) and (B) window size = 21, stringency = 14 and for (C) window size = 30, stringency = 8

subsequently designated *kamC*. The similarity ranged from nucleotide 1 to approximately 650 following the structural gene where the identity stops.

Similar analysis of the amino acid sequence obviously revealed a strong case for suggesting that these two proteins are homologous (Fig. 5.14 and Fig. 5.15). However, if the upstream sequences of both genes are used to make a protein translation, in frame with the respective genes, the similarity between these two putative peptide sequences is poor. So, while the DNA sequences preceding the putative translation start sites of *kamA* and *kamC* are very similar, the proteins that they might encode are utterly different.

Indeed, using such analysis, the most likely candidate for the initiation codon of *kamC* was identified as ATG3 (Fig. 5.13).

While the comparison of the sequences strongly suggests that *kamC* is a methylase which modifies ribosomal RNA, it also raises a question over the expression of the genes in *E. coli*, and that is, if *kamC* can be expressed as a transcriptional fusion in *E. coli*, why can *kamA* not be similarly expressed? At least three explanations exist.

The first, and most obvious is to suggest that transcripts initiating at *lacZ* in pLST413 terminate prior to the translation initiation codon of *kamA*, while perhaps no transcription terminator exists upstream of *kamC*.

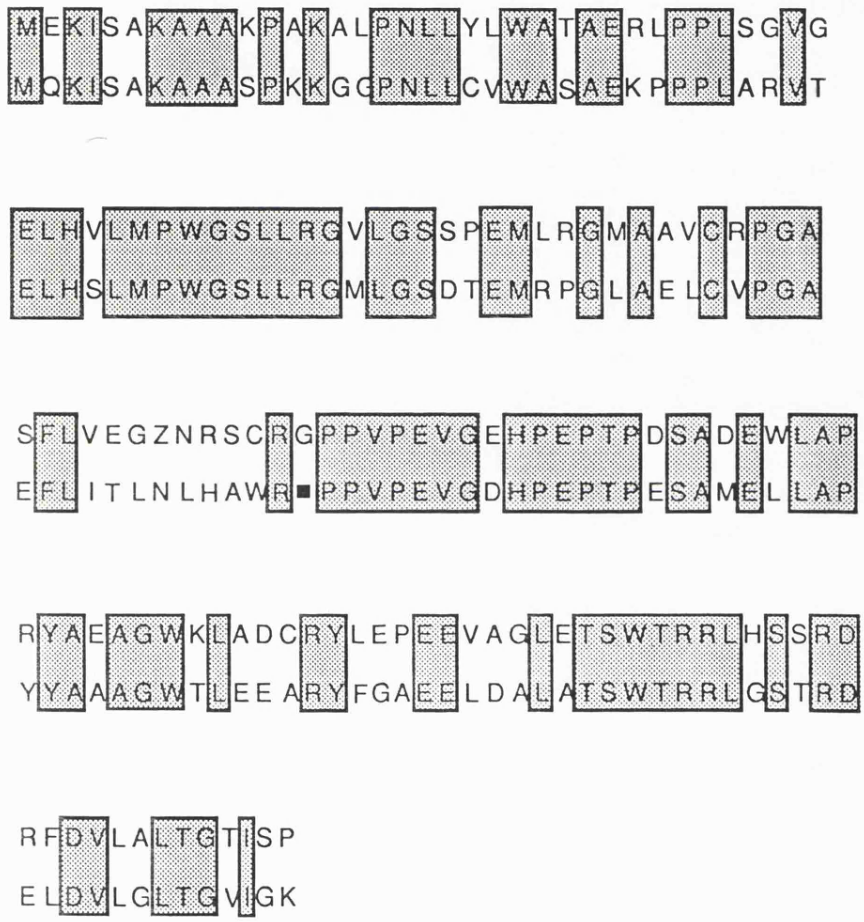


Fig. 5.15 Alignment of the predicted amino acid sequences of *kamA* and *kamC*

Although the restriction maps of the DNA fragments containing *kamA* and *kamC* bear little resemblance, the proteins that they encode must surely have the same mode of action.

Secondly, the ribosome binding site of *kamC* allows a more stable interaction than does that preceding the *S. tenjimariensis* gene and could, conceivably, result in *kamC* being expressed while *kamA* is not.

Finally, when transcription initiates at *lacZ* in pLST413, ribosomes can initiate at the α-fragment ATG, and will not terminate before the *kamA* coding sequence, which is not in frame. Ribosomes actively translating the upstream sequence of the *S. tenjimariensis* DNA might preclude ribosomes from initiating at the *kamA* ATG.

The results from Chapter 3 suggested that genes similar to *kgmA* might be present in a number of *Micromonospora* species that produce gentamicin-like antibiotics as well as in *S. tenebrarius*. It can now be seen that genes similar to the *kamA* gene of *S. tenjimariensis* exist in at least two other actinomycetes, namely *S. tenebrarius* and *Sacc. hirsuta*.

Since methylase genes are responsible for aminoglycoside resistance in a number of producing organisms it is now quite reasonable to suppose that, if as predicted, these are the origin of clinical resistance determinants, then it ought to be possible to demonstrate their existance in clinical isolates.

CHAPTER 6

DISCUSSION

6.1. Autoregulation.

It would be plausible to suppose that the level of resistance conferred by antibiotic inactivating enzymes might be related to gene dosage, in the sense that, more gene copies might mean more enzyme molecules which in turn could inactivate more substrate. However, such arguments do not apply to ribosome modifying enzymes, such as the methylases described in Chapters 4 and 5, since the amount of substrate (i.e. the mycelial content of rRNA) is such that relatively few enzyme molecules ought to be needed for complete modification. Regulation of the level of expression of such a gene would therefore be useful and a precedent exists for a novel mechanism of autoregulation by RNA binding proteins. Synthesis of most, if not all, ribosomal proteins in *E. coli* is coordinately and stoichiometrically balanced with the assembly of mature ribosomes. It has been established that a translational feedback mechanism exists and plays an important role in this regulation. One system extensively studied in this regard is the L11 r-protein operon, which is composed of the genes for L11 and L1. Experiments identified L1 as a translational repressor regulating the synthesis of both L11 and L1 (Yates *et al.*, 1980; Dean and Nomura, 1980; Brot *et al.*, 1981). The mRNA target site for this repression has been localised to a region at or near the translation initiation site of the first cistron, L11. The inhibition of translation of the second cistron, L1, is the direct result of the coupling of its translation to the preceding L11 cistron. The model for translational feedback regulation proposes that the mRNA target site structurally resembles the L1 binding site on 23S rRNA, although the affinity of L1 for the mRNA site is weaker that for the binding site on 23S rRNA. In this way, repression of translation

takes place only under conditions in which L1 synthesis exceeds the available 23S rRNA and the concentration of free L1 in the cellular pool increases.

In addition, there is evidence for a mechanism of autoregulation of the staphylococcal *ermC* gene, which encodes an rRNA methylase conferring resistance to the MLS antibiotics. The control of expression of the erythromycin-inducible gene is complex. A translational attenuation system allows "activation" of the mRNA by erythromycin causing the ribosomes to stall while translating a leader peptide, resulting in the transcript adopting an alternative secondary structure and releasing the previously sequestered ribosome binding site of the methylase gene. When enough ribosomes have been modified, they no longer stall in the leader sequence with the result that ribosome binding site of *ermC* become sequestered in secondary structure again. Consistent with the existence of this methylation mediated feedback loop is the observation that most of the erythromycin sensitive *ermC* mutations, caused by deletions in the structural gene, lead to the overproduction of inactive methylase, in the presence of inducing concentrations of erythromycin. However, one particular point mutation in the *ermC* coding sequence, which abolishes methylase activity does not result in overproduction of the altered protein. The hypothesis to explain this breakdown in the feedback regulation, is to suggest a role for autoregulation (Denoya *et al.*, 1986). The sequence at the *ermC* ribosome binding site is identical to that surrounding and including the site of action of the methylase in 23S rRNA to which it must bind. The possibility exists then that the inactive

point mutant, while lacking enzyme activity, retains a regulatory function by binding to its own mRNA, thereby precluding ribosomes from initiating there (Gryzcan *et al.*, 1980; Shivakumar *et al.*, 1980; Horinouchi and Weisblum, 1981).

This very attractive model of autoregulation might also be proposed to regulate *kgmA* expression. Within the *kgmA* leader there exists a sequence CGUCA, close to the initiation codon, which is also found embracing the site of action (residue G-1405) of the *kgmA* product in 16S rRNA (Fig. 6.1). This pentameric sequence occurs only twice in the total DNA sequenced, the other occurrence being on the non-coding strand at positions 855-851 in Fig. 4.5 of Chapter 4, but it is the location of this sequence close to the putative ribosome binding site of *kgmA* that is so suggestive. Perhaps it too can be recognised by the *kgmA* protein.

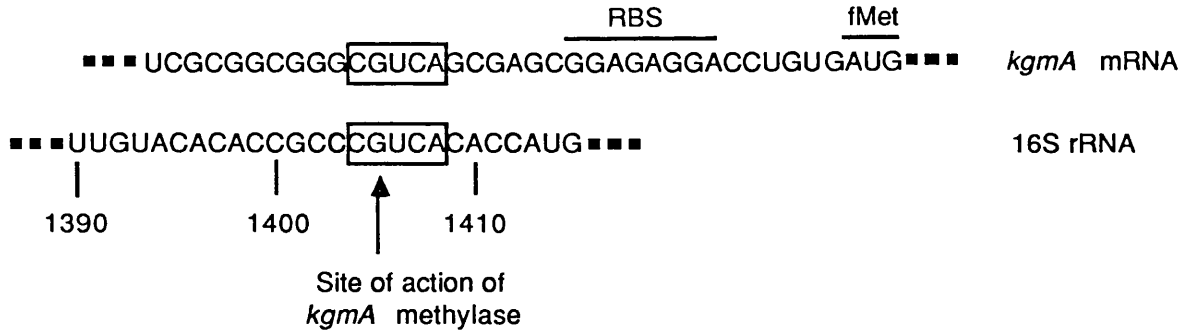


Fig. 6.1 Autoregulation of *kgmA*

There exists, in the *kgmA* mRNA leader, a sequence that is identical to that embracing the site of action of the methylase in 16S rRNA. The possibility exists then, that the methylase could bind to its own transcript in an autoregulatory manner.

6.2. Recombination: A model for the generation of pLST143.

The deletion event that resulted in the generation of pLST143 is reminiscent of similar events that occur when linear plasmid molecules are used to transform *E. coli*. A series of experiments involving the transformation of *E. coli* with linearised pBR322 and sequence analysis of the rearranged constructs, (Conley *et al.*, 1986a,b) revealed that many transformants contained recircularised plasmids bearing deletions and other rearrangements. Homologous recombination between directly repeated sequences ranging from 4 to 10 bp in length was proposed as having a role in the recircularisation process. According to the strategy for subcloning *kgmA* from pLST14, all of the sites in the polylinker between *Sst*I and *Pst*I ought to have been removed, but in pLST143 restriction mapping and sequence analysis revealed that this had not been the case.

When polylinker sequences such as those in pUC18 are restricted, simultaneously, with two endonucleases, a proportion of the DNA molecules are sometimes cut by only one enzyme due to the close proximity of the recognised sequences. If a pUC18 molecule, linearised with *Pst*I was involved in ligation with the relevant *Sst*I-*Pst*I fragment of pLST14, the result would be necessarily a linear molecule. Recombination of this type of construct *in vivo*, subsequent to transformation of *E. coli*, could conceivably result in the generation of pLST143 (Fig. 6.2). The nature of the sequence at the site of deletion also supported this theory since there is 5 base identity between pUC18 and *kgmA* at the point where the putative recombination would

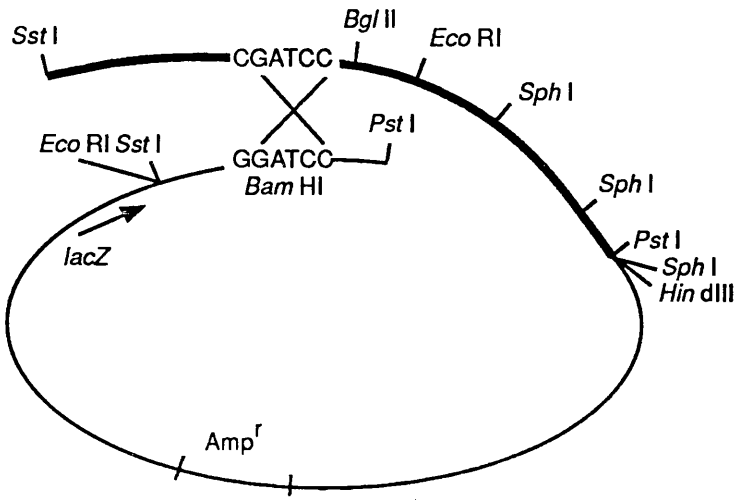


Fig. 6.2 A model for the generation of pLST143

Plasmid pLST143 was the result of an *in vivo* rearrangement that deleted about 1.4 kb of DNA, but allowed the expression of *kgmA* in *E. coli*. A mechanism proposed by Conley and co-workers to explain the rearrangements observed when *E. coli* is transformed with linear plasmid might be applied to those involved in the construction of pLST143. If the relevent fragment of pLST14 was ligated to a molecule of pUC18, restricted only by *Pst*I, then the result would be that shown above. Transformation of *E. coli* with what is necessarily a linear molecule could, conceivably, result in *in vivo* recircularisation similar to that described by others. Indeed, examination of the sequence at the point of deletion reveals five base (GATCC) identity between the *M. purpurea* DNA and pUC18.

have taken place (Fig. 6.2). While there is no direct evidence in favour of such a hypothesis, the observations certainly fulfil the criteria of the model proposed by Conley and co-workers.

6.3. Is *kgmA* encoded on an extrachromosomal element?

The occurrence of circular extrachromosomal DNA in *Streptomyces* is a firmly establish observation and has been discussed earlier (Chapter 1). However, the existence of giant linear plasmids in many of the *Streptomyces* tested has been a recent and unexpected discovery. Furthermore, the involvement of these plasmids in antibiotic biosynthesis and resistance has been demonstrated, at least in the case of methylenomycin.

The possibility exists, of course, that genes responsible for aminoglycoside biosynthesis and resistance are plasmid borne although no such elements have been discovered in the three aminoglycoside producing streptomycetes tested: *S. bikiniensis* (streptomycin), *S. kanamyceticus* (kanamycin) and *S. fradiae* (neomycin). Certainly, a property of *kgmA* deserving comment, is the observation that, in Southern analysis, the copy number of *kgmA* in *M. purpurea* seems to be remarkably low. While, approximately equal amounts of chromosomal DNA were required from *S. tenebrarius, M. inyoensis* and *M. rhodorangea*, to get equivalent signals using a *kgmA* derived radiolabelled probe, about 5 times the quantity of *M. purpurea* chromosomal DNA was needed to achieve a similar signal. The alternative explanations for this observation are that either *S. tenebrarius, M. rhodorangea* and *M. inyoensis* have multiple copies of

the kanamycin-gentamicin methylase gene or that *M. purpurea* has less than 1 copy of *kgmA* per chromosome. The implications of the latter conclusion are not only that *kgmA* should be extrachromosomally encoded, but that this element should only exist in a limited number of mycelial units.

A growing actinomycete colony is composed of different cell types, including basal or substrate mycelium, some of which will be lysing or degenerating in a mature colony, as well as cells developing from mycelium to hyphae, and mature hyphae forming spores. Could it be then that some plasmids are only maintained in certain cell types?

6.4. Do aminoglycosides cause misreading in *E. coli* DH143?

The occurrence of multiple binding sites for aminoglycosides on ribosomes has previously been invoked to explain their multiple effects, including inhibition of protein synthesis and mistranslation. Now, we see that the inhibitory effects of kanamycin and gentamicin can be blocked by modification of a single site within 16S rRNA. It should therefore be possible, with the generation of plasmid constructs which allow the expression of *kgmA* in *E. coli*, to determine whether ribosomes already modified by the methylase still respond to the potential misreading effects of aminoglycosides.

Attempts to answer this question have been unsuccessful. The experiments examined the ability of aminoglycosides to suppress nonsense mutations in bacteriophage λ. Some mutations were in enzymic functions, which might be easily suppressed and others were

mutations in structural proteins which ought to be relatively difficult to suppress (Table 6.1). It was proposed that, in the presence of certain aminoglycoside antibiotics nonsense mutations would be suppressed leading to plaque formation following the transfection of a suppressor free host. However, the inhibitory effect of both kanamycin and gentamicin on the growth of the suppressor free host was greater than the ability to suppress the nonsense mutations.

Perhaps the experiments could be more easily performed using *E. coli* strains harbouring nonsense mutations in auxotrophic markers.

When the right conditions for misreading can be demonstrated in *E. coli* then the effect of rRNA methylation on mistranslation of codons can be examined, thereby resolving the question of whether this group of antibiotics have multiple active binding sites on the ribosome.

6.5. Do rRNA methylases have a common ancestor?

Self-protection in organisms which produce antibiotics that inhibit ribosomal function is often mediated via specific methylation of ribosomal RNA (see Table 6.2). Computer assisted comparisons of the DNA or protein sequences reveal little about possible RNA or S-adenosylmethionine binding sites. Comparisons have shown that *ermE* and *ermSF (tlrA)* products possess a high degree of similarity not only to each other but to methylases encoded by other members of the *erm* family of genes, which is perhaps not surprising since the product of action of these enzymes is identical.

Table 6.1. Bacteriophage λ Amber mutants.

λ bacteriophage	Function of mutated gene product
λ *Pam* 80	DNA synthesis
λ *Pam* 3	DNA synthesis
λ i^4 *Oam* 8	DNA synthesis
λ *Sam* 7	Host lysis
λ *Nam* 7 *Nam* 53 i^{434} cI^-	Antitermination factor
λ imm^{21} *Qam*	Late gene regulation
λ *Jam* 6	Tail component
λ *Aam Bam*	DNA packaging and head component

Table 6.2 Methylation of ribosomal RNA associated with antibiotic-resistance determinants isolated from actinomycetes.

Organism	Resistance gene	Resistance character	Site of RNA methylation	Methylated residue	References
S. azureus	*tsr*	Thiostrepton	23S (1067)	Am	Thompson *et al.*, 1982
Sacc. erythraea	*ermE*	MLS	23S (2058)	m_2^6 A	Skinner *et al.*, 1983
S. fradiae	*tlrA(ermSF)*	MLS	23S (2058)	m_2^6 A	Zalacain & Cundliffe submitted
S. caelestis	*clr*	Lincosamides	23S (2058)	m^6 A	Calcutt & Cundliffe in preparation
S. lividans	*lrm*	Lincosamides	23S (2058)	m^6 A	Jenkins *et al*, submitted
M. purpurea	*kgmA*	Kanamycin Gentamicin	16S (1405)	m^7 G	Beauclerk &Cundliffe, 1987
S. tenjimariensis	*kamA*	Kanamycin Apramycin	16S (1408)	m^1 A	Beauclerk & Cundliffe, 1987
S. tenebrarius	*kgmB*	Aminoglycosides	16S (1405)	m^7 G	A. Beauclerk,
	kamB		(1408)	m^1 A	personnal communication.
S. kanamyceticus	*kan*	Kanamycin Gentamicin	16S (1405)	m^7 G	A. Beauclerk, personnal communication
Sacc. hirsuta	*kamC*	Kanamycin Apramycin	16S (1408)?	m^1 A ?	This study

The protein encoded by a carbomycin resistance gene (*carB*) isolated from *S. thermotolerans* has also been shown to be similar to the *ermE* methylase but it remains to be established whether *carB* encodes an enzyme with a similar mode of action.

It was quite reasonable to expect that, since *kgmA* and *kamA* acted at separate sites but which were close to each other in 16S rRNA, some similarity might be observed between either their DNA sequences, or that of the proteins encoded by them. The observation that the two proteins were so different in size seemed to remove any possibility of adequate comparative data, which appeared to be confirmed when computer analysis failed to identify similarity at the level of either the DNA or protein.

The identification of similarity between isofunctional enzymes from different bacterial genera, such as that between the *erm* genes, has led to speculation that they are homologous (that is, derived from a common ancestor).

One such study (Yeh and Ornston, 1980) centred on the enzymes involved in the β-ketoadipate pathway from *Pseudomonas* and *Acinetobacter*. Comparisons of the sequences of two enzymes which mediate similar intralactonic rearrangements, suggests that the enzymes have diverged from a common ancestor, despite having only 20% identity after introduction of gaps into the sequence. It would seem that the significance of sequence similarity is in the eye of the beholder.

With this level of protein sequence similarity (can anyone be sure of true homology) in mind, the *kgmA* and *kamA* predicted protein sequences were aligned and positions of identity and conservative substitutions indicated (Fig. 6.3).

Although the level of sequence similarity is quite considerable, the actual significance is obviously limited. However, with so many methylase sequences now available, it might be possible to start making some conclusions concerning substrate binding domains.

The resolution of questions in science invariably results in the proposition of further enquiries. This study is no exception and some of the more obvious new queries have been outlined in this discussion chapter. They range from rather wide questions about the origin of methylases and their possible role in clinical resistance to specific ones concerning the regulation of *kgmA* expression. Arguably, the most obvious experiments for consideration are those which might resolve the long-standing controversy of whether aminoglycosides have multiple sites of action on the ribosome.

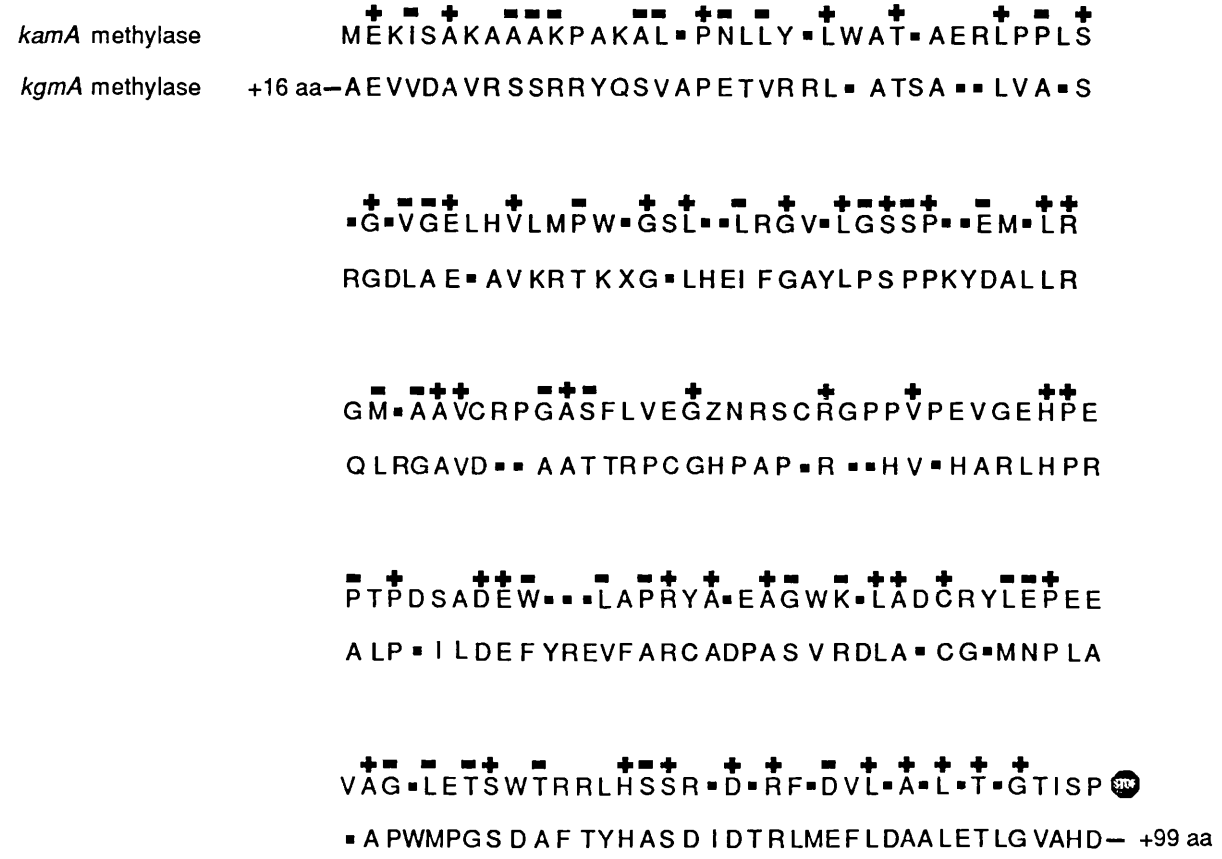


Fig. 6.3 Comparison of the amino acid sequences of the methylases encoded by *kgmA* and *kamA*.

The protein sequence of the *kamA* methylase is shown together with part of that of the *kgmA* methylase. Positions of identity are indicated (✚), as well as those of conservative amino acid substitution (▬). The latter were based on the following groups.

1. P,A,G,S,T (neutral, weakly hydrophobic)
2. Q,N,E,D,B,Z (hydrophilic, acid amine)
3. H,K,R (hydrophilic, basic)
4. L,I,V,M (hydrophobic)
5. F,Y,M (hydrophobic, aromatic)
6. C (cross-link forming)

REFERENCES

Adams, C.W., Fornwald, J.A., Schmidt, F.J., Rosenberg, M. and Brawner, ME. (1988). *J. Bacteriol.*, **170**, 203-212.

Ahmed, Z.U. and Vining, L.C. (1983). *J. Bacteriol.*, **154**, 239-244.

Altenbuchner, J., and Cullum, J. (1985). *Mol. Gen. Genet.*, **201**, 192-197.

Antonov, P.P., Ivanov, I.G., Markov, G.G. and Benigni, R. (1978). *Studia Biophysica Berlin*, **69**, 67-74.

Ausubel, F.M. (1984). *Cell*, **37**, 5-6.

Baldacci, E. and Locci, R. (1961). *Annu. Microbiol. Enzymol.*, **11**, 19-30.

Beauclerk, A.D.D., Hummel, H., Holmes, D.J., Böck, A., and Cundliffe, E. (1985). *Eur. J. Biochem.*, **151**, 245-255.

Beauclerk, A.A.D. and Cundliffe, E. (1987). *J. Mol. Biol.*, **193**, 661-671.

Benveniste, R. and Davies, J. (1973). *Proc. Natl. Acad. Sci. USA.*, **70**, 2276-2280.

Bibb, M.J., Freeman, R.F. and Hopwood, D.A. (1977). *Mol. Gen. Genet.*, **154**, 155-166.

Bibb, M.J. Ward, J.M. and Hopwood, D.A. (1978). *Nature*, **274**, 398-400.

Bibb, M.J., Ward, J.M., Kieser, T., Cohen, S.N. and Hopwood, D.A. (1981). *Mol. Gen. Genet.*, **184**, 230-240.

Bibb, M.J. and Cohen, S.N. (1982). *Mol. Gen. Genet.*, **187**, 265-277.

Bibb, M.J., Findlay, P.R. and Johnson, M.W. (1984). *Gene*, **30**, 157-166.

Bibb, M.J., Janssen, G.R. and Ward, J.M. (1985). *Gene*, **38**, 215-226.

Bibb, M.J., Jones, G.H., Joseph, R., Buttner, M.J. and Ward, J.M. (1987). *J. Gen. Microbiol.*, **133**, 2089-2096.

Birmingham, V.A., Cox, K.L., Larson, J.L., Fishman, S.E., Hershberger, C.L. and Seno, E.T. (1986). *Mol. Gen. Genet.*, **204**, 532-539.

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

Blanco, M.G., Hardisson, C. and Salas, J.A. (1984). *J. Gen. Microbiol.*, **130**, 2883-2891.

Brawner, M., Ingram, C., Fornwald, J., Lichenstem, H. and Westpheling, J. (1988). In: *Molecular Biology of Actinomycetes '88.* (Okami, Y., Beppu, T and Ogawara, H., ed.). pp. 35-40.

Brendel, V. and Trifonov, E.N. (1984). *Nucleic Acids Res.*, **12**, 4411-4427.

Brot, N., Caldwell, P. and Weissbach, H. (1981). *Arch. Biochem. Biophys.*, **206**, 51-53.

Buckel, P., Buchberger, A., Böck, A. and Wittmann, H.G. (1977). *Mol. Gen. Genet.*, **158**, 47-54.

Buttner, M.J. and Brown, N.L. (1985). *J. Mol. Biol.*, **185**, 177-188.

Buttner, M.J., Fearnley, I.M. and Bibb, M.J. (1987). *Mol. Gen. Genet.*, **209**, 101-109.

Buttner, M.J., Smith, A.M. and Bibb, M.J. (1988). *Cell*, **52**, 599-607.

Caltrider, P.G. (1967) In: *Antibiotics. Vol I.* (Gottlieb, D. and Shaw, P.D., ed.). pp. 677-680.

Champness, W.C. (1988). *J. Bacteriol.*, **170**, 1168-1174.

Chan, S.J., Weiss, J., Konrad, M., White, T., Bahl, C., Yu, S.D., Mark, D. and Steiner, D.F. (1981). *Proc. Natl. Acad. Sci. USA.*, **78**, 5401-5405.

Chater, K.F. (1972). *J. Gen. Microbiol.*, **72**, 9-28.

Chater, K.F. and Cooper, C.S. (1975). *Annual Report of the John Innes Institute*, **66**, 83.

Chater, K.F. and Bruton, C.J. (1983). *Gene*, **26**, 67-78.

Chater, K.F. and Bruton, C.J. (1985). *EMBO J.*, **4**, 1893-1987.

Chater, K.F., Lawlor, E.J., Mendez, C., Bruton, C.J., Davis, N.K., Plaskitt, K., Guthrie, E.P., Daly, B.L., Baylis, H.A. and Trong, K. VU. (1988) In: *Biology of Actinomycetes '88* (Okami, Y., Beppu, T and Ogawara, H., ed.). pp. 64-70.

Clark-Curtiss, J.E., Jacobs, W.R., Docharty, M.A., Ritchie, L.R. and Curtiss III, R. (1985). *J. Bacteriol.*, **161**, 1093-1102.

Conley, E.C., Saunders, V.A. and Saunders, J.R. (1986a). *Nucleic Acid Res.*, **14**, 8905-8917.

Conley, E.C., Saunders, V.A., Jackson, V. and Saunders, J.R. (1986b). *Nucleic Acids Res.*, **14**, 8919-8932.

Courvalin, P., Weisblum, B. and Davies, J. (1977). *Proc. Natl. acad. Sci. USA.*, **74**, 999-1003.

Cowing, D.W., Bardwell, J.C.A., Craig, E.A., Woolford, C., Hendrix, R.W. and Gross, C.A. (1985). *Proc. Natl. Acad. Sci. USA.*, **82**, 2679-2683.

Crameri, R. and Davies, J.E. (1986). *J. Antibiotics*, **39**, 128-135.

Curran, J. and Kolakofsky, D. (1988). *EMBO J.*, **7**, 245-251.

Daniels, D.L., De Wet, J.R. and Blattner, F.R. (1980). *J. Virology*, **33**, 390-400.

Davies, J., Gorini, L. and Davis, B.D. (1965). *Mol. Pharmacol*, **1**, 93-106.

Dehottay, P., Dusart, J., De Meester, F., Joris, B., van Beeumen, J., Erpicum, T., Frère, J-M. and Ghuysen, J-M. (1987). *Eur. J. Biochem.*, **166**, 345-350.

Denoya, C.D., Bechnofer, D.H. and Dubnau, D. (1986). *J. Bacteriol.*, **168**, 1133-1141.

Deuschle, U., Kammerer, W., Gentz, R. and Bujard, H. (1986). *EMBO J.*, **5**, 2987-2994.

Dietz, A. (1988). In: *Biology ofActinomycetes'88* (Okami, Y., Beppu, T. and Ogawara, E., ed.). pp. 203-209.

Distler, J., Ebert, A., Mansouri, K., Pissowotzki, K., Stockman, M. and Piepersberg, W. (1987). *Nucleic Acids Res.*, **15**, 8041-8056.

Drake, J.W. and Baltz, R.H. (1976). *Annu. Rev. Biochem.*, **45**, 11-37.

Dyson, P., and Schrempf, H. (1987). *J. Bacteriol.*, **169**, 4796-4803.

Epp, J.K., Burgett, S.G. and Schoner, B.E. (1987). *Gene*, **53**, 73-83.

Favarolo, J., Treisman, R. and Kamen, R. (1980) *Methods in Enzymol.*, **65**, 718-749.

Feinberg, A.P. and Vogelstein, B. (1984). *Anal. Biochem.*, **137**, 266-267.

Feitelson, J.S. and Hopwood, D.A. (1983). *Mol. Gen. Genet.*, **90**, 394-398.

Feitelson, J.S., Malpartida, F. and Hopwood, D.A. (1985). *J. Gen. Microbiol.*, **131**, 2431-2441.

Fornwald, J.A., Schmidt, F.J., Adams, C.W., Rosenberg, M. and Brawner, M.E. (1987). *Proc. Natl. Acad. Sci. USA.*, **84**, 2130-2134.

Gellert, M. (1981). *Annu. Rev. Biochem.*, **50**, 879-910.

Gentz, R. and Bujard, H. (1985). *J. Bacteriol.*, **164**, 70-77.

Gesteland, R. F. (1966) *J. Mol. Biol.*, **16**, 67-84.

Gil, J.A. and Hopwood, D.A. (1983). *Gene*, **25**, 119-132.

Gil, J.A., Kieser, H.M. and Hopwood, D.A. (1985). *Gene*, **38**, 1-8.

Gitt, M., Wang, M. and Doi, R. (1985). *J. Mol. Biol.*, **260**, 7178-7185.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. and Stormo, G. (1981). *Annu. Rev. Microbiol.*, **35**, 365-403.

Gorini, L. (1974). In: *Ribosomes* (Chambliss, G., Craven, G., Davies, J., Davis, K., Kahan, L. and Nomura, M., ed.) pp. 791-803.

Gough, J.A. and Murray, N.E. (1983). *J. Mol. Biol.*, **166**, 1-19.

Gregory, K.F. and Huang, J.C.C. (1964). *J. Bacteriol.*, **87**, 1287-1294.

Gribskov, M., Devereux , J. and Burgess, R.R. (1984). *Nucleic Acids Res.*, **12**, 539-549.

Grossman, A.D., Erickson, J.W. and Gross, C.A. (1984). *Cell*, **3**, 383-390.

Gryzcan, T.J., Contente, S. and Dubnau, D. (1978). *J. Bacteriol.*, **134**, 318-329.

Gryzcan, T.J., Grandi, G., Hahn, J., Grandi, R., and Dubnau, D. (1980). *Nucleic Acids Res.* **8**, 722-735.

Gryczan, T.J., Israeli-Reches, M., del Bue, M. and Dubnau, D. (1984). *Mol. Gen. Genet.*, **194**, 349-356.

Guzman, P., Westpheling, J. and Youngman, P. (1988). *J. Bacteriol.*, **170**, 1598-1609.

Harris, J.E., Chater, K.F., Bruton, C.J. and Piret, J.M. (1983). *Gene*, **22**, 167-174.

Hawley, D.K. and McClure, W.R. (1983). *Nucleic Acids Res.*, **11**, 2237-2255.

Hay, R.E. and Moran, C.P.Jr. (1985). In: *Molecular Biology of Differentiation* (Hoch, J.A. and Setlow, P., ed.). pp. 170-175.

Hercomb, J., Thierbach, G., Baumberg, S. and Parish, J.H. (1987) *J. Gen. Microbiol.*, **133**, 317-322

Higgens, C.E. and Kastner, R.E. (1968). *Antimicrob. Agents and Chemother.* 1967. 324-331.

Hirschman, J., Wong, P-K., Sei, K., Keener, J. and Kustu, S. (1985). *Proc. Natl. Acad. Sci. USA.*, **82**, 7525-7529.

Hodgson, D.A. and Chater, K.F. (1981). *J. Gen. Microbiol.*, **124**, 339-348.

Hopwood, D.A., Kieser, T., Wright, H.M. and Bibb, M.J. (1983). *J. Gen. Microbiol.*, **129**, 2257-2269.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., ward, J.M. and Schrempf, H. (1985). In: *Genetic Manipulation of Streptomyces : A Laboratory Manual.* John Innes Foundation, Norwich, England.

Hopwood, D.A. (1988). *Proc. R. Soc. Lond.*, **235**, 121-138.

Horinouchi, S., Byeon, W-H. and Weisblum, B. (1983). *J. Bacteriol.*, **154**, 1251-1262.

Horinouchi, S., and Weisblum, B. (1981). *Mol. Gen. Genet.*, **182**, 341-348.

Hunt, T.P. and Magasanik, B. (1985). *Proc. Natl. Acad. Sci. USA.*, **82**, 8453-8457.

Ikeda, H., Tanaka, H. and Omura, S. (1982a) *J. Antibiot.*, **4**, 497-506.

Ikeda, H., Tanaka, H. and Omura, S. (1982b) *J. Antibiot.*, **4**, 507-516.

Igo, M.M. and Losick, R. (1986). *J. Mol. Biol.*, **191**, 615-624.

Janssen, G.R., Bibb, M.J., Smith, C.P., Ward, J.M., Kieser, T. and Bibb, M.J. (1985). *Microbiology-1985*. (Lieve, L., ed.). pp. 392-396.

Jaurin, B., Grundstrom, T., Edlund, T. and Normark, S. (1981). *Nature*, **290**, 221-225.

Jaurin, B., Grundstrom, T., and Normark, S. (1982). *EMBO J.*, **1**, 875-881.

Jaurin, B. and Normark, S. (1983). *Cell*, **32**, 809-816.

Jaurin, B. and Cohen, S.N. (1984). *Gene*, **28**, 83-91.

Jaurin, B. and Cohen, S.N. (1985). *Gene*, **39**, 191-201.

Jelenc, P.C. and Kurland, C.G. (1979). *Proc. Natl. Acad. Sci. USA.*, **76**, 3174-3178.

Jenkins, J.R., and Cooper, R.A. (1988). *J. Bacteriol.*, **170**, 5317-5324.

Jones, G.H. (1979). *Arch. Biochem. Biophys.*, **198**, 195-204.

Kamimiya, S. and Weisblum, B. (1988). *J. Bacteriol.*, **170**, 1800-1811.

Kammerer, W., Deuschle, U., Gentz, R. and Bujard, H. (1986). *EMBO J.*, **5**, 2995-3000.

Kassavetis, G.A. and Geiduschek, E.P. (1984). *Proc. Natl. Acad. Sci. USA.*, **81**, 5101-5105.

Katz, E., Thompson, C.J. and Hopwood, D.A. (1982). *J. Gen. Microbiol.*, **129**, 2703-2714.

Katz, L., Brown, D., Boris, K., and Tuan, J. (1987). *Gene*, **55**, 319-325

Kendall, K. and Cullum, J. (1984). *Gene*, **29**, 315-321.

Kendall, K. and Cohen, S.N. (1988). *J. Bacteriol.*, **170**, 4634-4651.

Kenney, T.J. and Moran, C.P.Jr. (1987). *J. Bacteriol.*, **169**, 3329-3339.

Khoklov, A.S. (1982). In: *Overproduction of Microbial Products* (Krumphanzi, V., Sikyta, B. and Vanek, Z., ed.). pp. 97-109.

Kieser, T., Hopwood, D.A., Wright, H.M. and Thompson, C.J. (1982). *Mol. Gen. Genet.*, **185**, 223-238.

Kieser, T. (1984). *Plasmid*, **12**, 19-36.

Kieser, T., Moss, M.T., Dale, J.W. and Hopwood, D.A. (1986). *J. Bacteriol.*, **168**, 72-80.

Kinashi, H., Shimaji, M. and Sakai, A. (1987). *Nature*, **328**, 454-456.

Laemmli, U.K. (1970). *Nature*, **227**, 680-685.

Landlick, R., Vaugh, V., Lau, E.T., van Bogelen, R.A., Erickson, J.W. and Niedhart, F.C. (1984). *Cell*, **38**, 175-182.

Langridge, J., Langridge, P. and Berquist, P.L. (1980). *Analyt. Biochem.*, **103**, 264-271.

Lawlor, K.J., Baylis, H.A. and Chater, K.F. (1987). *Genes Develop.*, **1**, 1305-1310.

Long, C.M., Virolle, M-J., Chang, S-Y., Chang, S. and Bibb, M.J. (1987). *J. Bacteriol.*, **169**, 5745-5754.

Luedemann, G.M. and Brodsky, B.C. (1964). *Antimicrobial Agents and Chemother.*, 1963. 116-124.

Lydiate,D.J., Ikeda, H., and Hopwood, D.A. (1986). *Mol. Gen. Genet.*, **203**, 79-88.

Maizel, J.V.Jr. and Lenk, R.P. (1981). *Proc. Natl. Acad. Sci. USA.*, **78**, 7665-7669.

Malik, V.S. (1977). *J. Antibiot.*, **30**, 897-899.

Malpartida, F. and Hopwood, D.A. (1984). *Nature*, **309**, 462-464.

Malpartida, F. and Hopwood, D.A. (1986). *Mol. Gen. Genet.*, **205**, 66-75.

Mandel, M. and Higa, A. (1970). *J. Mol. Biol.*, **53**, 159-168.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory. Cold Spring Harbor, New York.

Martin, J.F. and Demain, A.L. (1980). *Microbiol. Rev.*, **44**, 230-251.

Matkovic, B., Piendl, W. and Böck, A. (1984). *FEMS Microbiol. Lett.*, **24**, 273-276.

McLaughlin, J.R., Murray, C.L. and Rabinowitz, J.C. (1984). *J. Biol. Chem.*, **256**, 11283-11291.

Méndez, C. and Chater, K.F. (1987). *J. Bacteriol.*, **169**, 5715-5720.

Merrick, M.J. (1976). *J. Gen. Microbiol.*, **96**, 299-315.

Miller, A. and Walker, J.B. (1969). *J. Bacteriol.*, **99**, 401-405.

Min, K.T., Kim, M.H. and Lee, D-S. (1988). *Nucleic Acids Res.*, **16**, 5075-5088.

Moazed, D., Stern, S. and Noller, H. (1986). *J. Mol. Biol.*, **187**, 399-416.

Monod, M., Mohan, S. and Dubnau, D. (1987). *J. Bacteriol.*, **169**, 340-350.

Moran, C.P.Jr., Long, N. and Losick, R. (1981). *Nucleic Acids Res.*, **9**, 5979-5990.

Moran, C.P.Jr., Johnson, W.C. and Losick, R. (1982). *J. Mol. Biol.*, **162**, 709-713.

Murakami, T., Holt, T.J. and Thompson, C.J. (1989). *J. Bacteriol.*, **171**, 1459-1466.

Murphy, E. (1985). *J. Bacteriol.*, **162**, 633-640.

Nakano, M.M., Mashiko, H. and Ogawara, H. (1984). *J. Bacteriol.*, **157**, 79-83.

Neal, R.J. and Chater, K.F. (1987). *Gene*, **58**, 229-241.

Nimi, O., Ito, G., Sueda, S. and Nomi, R. (1971). *Agric. Biol. Chem.*, **35**, 848-855.

Noriji, C., Watabe, H., Katsumata, K. and Yamada, Y. (1980). *J. Antibiot.*, **33**, 118-121.

Ochi, K. (1988). In: *Biology of Actinomycetes '88.* (Okami, Y., Beppu, T. and Ogawara, H., ed.). pp.330-337.

Ohnuki, T., Imanaka, T. and Aiba, S. (1985). *J. Bacteriol.*, **164**, 85-94.

Okanishi, M., Manome, T. and Umezawa, H. (1980). *J. Antibiot., 33, 88-91.*

Omura, S., Ikeda, H. and Tanaka, H. (1981). *J. Antibiot.*, **34**, 478-482.

Pestka, S. (1977). In: *Molecular Mechanisms of Protein Biosynthesis* (Weissbach, H. and Pestka, S., ed.). pp. 467-553.

Piendl, W. and Böck, A. (1982). *Antimicrob. Agents and Chemother.*, **22**, 231-236.

Piendl, W., Böck, A. and Cundliffe, E. (1984). *Mol. Gen. Genet.*, **197**, 24-29.

Piepersberg, W., Distler, J., Ebert, A., Heinzel, P., Mansouri, K., Mayer, G. and Pissowotzki, K. (1988). In: *Biology of Actinomycetes '88.* (Okami, Y., Beppu, T. and Ogawara, H., ed.). pp. 86-91.

Ponnambalam, S., Webster, C., Bingham, A. and Busby, S. (1986). *J. Biol. Chem.*, **261**, 16043-16048.

Prince, J.B., Taylor, B.H., Thurlow, D.L., Ofegand, J. and Zimmerman, R.A. (1982). *Proc. Natl. Acad. Sci. USA.*, **79**, 5450-5454.

Puglia, A.M., Alberti, M., Bucca, G., Granozzi, C., Passantino, R. and Misuracea, F. (1986). In: *Book of Abstracts; Fifth International Symposium on the Genetics of Microorganisms.* (Alacevic, M., Hranueli, D. and Toman, Z., ed.). pp. 102.

Reitzer, L. and Magasanik, B. (1987). In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology.* (Neidhardt, F., Ingraham, J., Brooks, L., Magasanik, B., Schaecter, M. and Umbarger, H., ed.). pp. 302-320.

Rhodes, P.M., Hunter, I.S., Friend, E.J. and Warren, M. (1984). *Biochem. Soc. Trans.*, **12**, 586-587.

Rodicio, M.R., Bruton, C.J. and Chater, K.F. (1985). *Gene*, **34**, 283-292.

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

Schmitz, A. and Galas, D.J. (1979). *Nucleic Acids Res.*, **6**, 111-132.

Schrempf, H. (1983). *Mol. Gen. Genet.*, **189**, 501-505.

Schrempf, H., Bujard, H. , Hopwood, D.A. and Goebel, W. (1975). *J. Bacteriol.*, **121**, 416-421.

Seno, E.T., Bruton, C.J. and Chater, K.F. (1984). *Mol. Gen. Genet.*, **193**, 119-218.

Ser, A. (1957). *J. Antibiot.*, **10**, 181-188.

Shine, J. and Dalgarno, L. (1974). *Proc. Natl. Acad. Sci. USA.*, **71**, 1342-1346.

Shivakumar, A.G., Hahn, J., Grundi, G., Kozlov, Y., and Dubnau, D. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3903-3907.

Siebenlist, U., Simpson, R.B. and Gilbert, W. (1986). *Cell*, **20**, 269-281.

Skeggs, P.A., Thompson, J. and Cundliffe, E. (1985). *Mol. Gen. Genet.*, **200**, 415-421.

Skeggs, P.A. (1986). PhD. Thesis. University of Leicester.

Skeggs, P.A., Holmes, D.J. and Cundliffe, E. (1987). *J. Gen. Microbiol.*, **133**, 915-923.

Skinner, R., Cundliffe, E. and Schmidt, F.J. (1983). *J. Biol. Chem.*, **258**, 12702-12706.

Smith, C.P. and Chater, K.F. (1987). In: *Proceedings of the Fifth International Symposium on the Genetics of Microorganisms. Part B.* (Alacevic, M., Hranueli, D. and Toman, Z., ed.). pp. 9-15.

Smith, C.P. and Chater, K.F. (1988). *Mol. Gen. Genet.*, **211**, 129-137.

Stanzak, R., Matsushima, P., Baltz, R.H. and Rao, R.N. (1986). *Biotechnology*, **4**, 229-232.

Stragler, P., Parsot, C. and Bouvier, J. (1985). *FEBS Lett.*, **187**, 11-15.

Streicher, S.L. and Tyler, B. (1981). *Proc. Natl. Acad. Sci. USA.*, **78**, 229-233.

Tai, P-C. and Davis, B.D. (1979). *Biochemistry*, **15**, 193-198.

Takashaki, H., Tanaka, K. and Shiina, T. (1988). In: *Biology of Actinomycetes '88*. (Okami, Y., Beppu, T. and Ogawara, H., ed.) pp. 58-63.

Talkington, C. and Pero, J. (1979). *Proc. Natl. Acad. Sci. USA.*, **76**, 5465-5469.

Talmadge, K. and Gilbert, W. (1982). *Proc. Natl. Acad Sci. USA.*, **79**, 1830-1833.

Tatti, K.M. and Moran, C.P.Jr. (1984). *J. Mol. Biol.*, **175**, 285-297.

Thiara, A.S. and Cundliffe, E. (1988). *EMBO J.*, **7**, 2255-2259.

Thiara, A.S. and Cundliffe, E. (1989). *Gene*, in press.

Thole, J.R., Dauwerse, H.G., Daus, P.K., Groothius, D.G., Schouls, L.M. and van Embden, J.D.A. (1985). *Infect. Immun.*, **50**, 800-806.

Thompson, C.J., Ward, J.M. and Hopwood, D.A. (1980). *Nature*, **286**, 525-527.

Thompson, C.J., Ward, J.M. and Hopwood, D.A. (1982a). *J. Bacteriol.*, **151**, 668-677.

Thompson, C.J., Skinner, R.H., Thompson, J., Ward, J.M., Hopwood, D.A. and Cundliffe, E. (1982b). *J. Bacteriol.*, **151**, 678-685.

Thompson, C.J. and Gray, G.S. (1983). Proc. Natl. Acad. Sci. USA., **80**, 5190-5194.

Thompson, J., Schmidt, F.J. and Cundliffe, E. (1982). *J. Biol. Chem.*, **257**, 7915-7917.

Thompson, J., Rae, S. and Cundliffe, E. (1984). *Mol. Gen. Genet.*, **195**, 39-43.

Thompson, J., Skeggs, P.A. and Cundliffe, E. (1985). *Mol. Gen. Genet.*, **201**, 168-173.

Thompson, J., Cundliffe, E. and Dahlberg, A.E. (1988). *J. Mol. Biol.*, **203**, 453-465.

Tinoco, Jr. I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973). *Nature*, **246**, 40-41.

Tohyama, H., Okami, Y. and Umezawa, H. (1987). *Nucleic Acids Res.*, **14**,

Vara, J., Malpartida, F., Hopwood, D.A. and Jiménez, A. (1985). *Gene*, **33**, 197-206.

Virolle, M-J. and Bibb, M.J. (1988). *J. Mol. Microbiol.*, **2**, 197-208.

Virolle, M-J., Long, C.M., Chang, S. and Bibb, M.J. (1988). *Gene*, **74**, 321-334.

Vivian, A. (1971). *J. Gen. Microbiol.*, **69**, 353-364.

Vögtli, M. and Hütter, R. (1987). *Mol. Gen. Genet.*, **208**, 195-203.

Wagman, G.H. and Weinstein, M.J. (1980). *Annu. Rev. Microbiol.*, **34**, 537-557.

Walker, M.S. and Walker, J.B. (1970). *J. Biol. Chem.*, **245**, 6683-6689.

Wang, J.C. (1985). *Annu. Rev. Biochem.*, **54**, 665-697.

Warburton, N., Boseley, P.G. and Porter, A.G. (1983) *Nucleic Acids Res.*, **11**, 5837-5854.

Ward, J.M., Janssen, G.R., Kieser, T., Bibb, M.J., Buttner, M.J. and Bibb, M.J. (1986). *Mol. Gen. Genet.*, **203**, 468-478.

Watanabe, T. (1971). *Ann. N.Y. Acad. Sci.*, **182**, 126-140.

Wawrousek, E.F., Narashima, M. and Hansen, J.N. (1984). *J. Biol. Chem.*, **259**, 3694-3702.

Westpheling, J., Ranes, M. and Losick, R. (1985). *Nature*, **313**, 22-27.

Westpheling, J. and Brawner, M. (1989). *J. Bacteriol.*, **171**, 1355-1361.

Witkin, E.M. (1969). *Annu. Rev. Microbiol.*, **23**, 487-514.

Witkin, E.M. (1976). *Bacteriol Rev.*, **40**, 869-907.

Wittman, H.G. and Wittmann-Liebold, B. (1974). In: *Ribosomes* (Chambliss, G., Craven, G., Davies, J., Davis, K., Kahan, L. and Nomura, M., ed.). pp. 115-140.

Xue, Y., Zhuang, Z., Zhu, Y., Xu, Y. and Dong, K. (1981). *J. Bacteriol.*, **146**, 412-414.

Yamaizumi, Z., Kuchino, Y., Harada, F., Nishimura, S. and McCloskey, J.A. (1980). *J. Biol. Chem.*, **255**, 2220-2225.

Yamamoto, H., Hotta, K., Okami, Y. and Umezawa, H. (1982). *J. Antibiotics*, **35**, 1020-1025.

Yanish-Perron, C., Vieira, J. and Messing, J. (1985). *Gene*, **33**, 103-119.

Yates, J.L., Arfsten, A.E. and Nomura, M. (1980). *Proc. Natl. Acad. Sci.* USA., **77**, 3590-3594.

Yeh, W.K. and Ornsten, L.N. (1980). *Proc. Natl. Acad. Sci. USA.*, **77**, 5365-5369.

Young, R.A., Bloom, B.R., Grosskinsky, M., Ivanyi, J., Thomas, D. and Davies, R.W. (1985a). *Proc. Natl. Acad. Sci. USA.*, **82**, 2583-2587.

Young, R.A., Mehra, V., Sweetser, D., Buchanan, T., Clark-Curtiss, J., Davis, R.W. and Bloom, B.R. (1985b). *Nature*, **316**, 450-452.

Zalacain, M., González, A., Guerrero, M.C., Mattaliano, R.J., Malpartida, F. and Jiménez, A. (1986). *Nucleic Acids Res.* **14**, 1565-1581.

Zalacain, M., Malpartida, F., Pulido, D. and Jiménez, A. (1987). *Eur. J. Biochem.*, **162**, 413-418.

Zuber, P. and Losick, R. (1983). *Cell*, **35**, 275-283.

Zuker, M. and Stiegler, P. (1981). *Nucleic Acids Res.*, **9**, 133- 148.

Abstract

NUCLEOTIDE ANALYSIS OF TWO ACTINOMYCETE AMINOGLYCOSIDE RESISTANCE DETERMINANTS

David J. Holmes

Resistance to aminoglycosides in the organisms that produce them is often ascribed to the well characterised and clinically important antibiotic modifying enzymes. However, at least three aminoglycoside producing actinomycetes, namely *Micromonospora purpurea, Streptomyces tenjimariensis,* and *Streptomyces tenebrarius* possess ribosomes that are refractory to some members of this class of drugs. In these cases, resistance is due to methylation of rRNA of the small ribosomal subunit. This study supports the possibility that this mechanism might be more widespread than hitherto suspected.

Two of the methylase genes have been analysed at the nucleotide level and their transcripts mapped. The gentamicin resistance methylase gene (*kgmA*) from *M. purpurea* codes for a 36 kDa protein consisting of 249 amino acids. Like most actinomycete genes, *kgmA* is not expressed in *E. coli* from its own promoter, although the determinant was expressed in this Gram-positive host as a result of DNA rearrangement. Sequence analysis of the mutated plasmid suggested that the methylase was expressed as a translational fusion with the *lacZ'* gene of pUC18, a view that was latter confirmed. Transcript mapping revealed that *kgmA* is probably read from a single promoter but that it might be part of a polycistron.

The second gene examined confers resistance to kanamycin and apramycin, and originated in *S. tenjimariensis*. This determinant (*kamA*) was shown to encode a predicted protein of 155 amino acids with a molecular weight of 19 kDa. Unlike *kgmA*, this gene could not be expressed as either a transcriptional or translational fusion in *E. coli*. Transcription of *kamA* is directed by tandem promoters and is a monocistron since the the transcript terminates only 160 bp downstream of the stop codon.