STUDIES ON ANTIBIOTIC RESISTANCE IN PRODUCING MICROORGANISMS

A thesis submitted to the
University of Southampton
for the degree of
Doctor of Philosophy
by

Christopher J. Herbert

Department of Biochemistry,
University of Southampton

September, 1983

ACKNOWLEDGEMENTS

I should like to thank my Supervisor, Professor M. Akhtar, FRS for his trust, Dr. W.T. Drabble for his patience and Dr. D.P. Bloxham for his encouragement; also Dr. J. Walker of the MRC Laboratories in Cambridge for teaching me the sequencing methods and Dr. I.G. Giles for all his work in running the computer programmes and teaching me how to use them.

I should also like to thank Joyce for turning my manuscript into a thesis and Vince for helping with the proof-reading. This work could not have been finished without the help and co-operation of Pete and Doug in stores, but might have been finished sooner without the endless hours of diverting entertainment provided by my colleagues in Lab. 423.

Financial support came from the SERC and Pfizer Central Research, Sandwich, Kent.

Finally, I should like to thank my Mother for her support and encouragement in everything I've done.

BEING BUT MEN

Being but men, we walked into the trees
Afraid, letting our syllables be soft
For fear of waking the rooks,
For fear of coming
Noiselessly into a world of wings and cries.

If we were children we might climb,
Catch the rooks sleeping, and break no twig,
And, after the soft ascent,
Thrust out our heads above the branches
To wonder at the unfailing stars.

Out of confusion, as the way is,
And the wonder that man knows,
Out of the chaos would come bliss.

That, then, is loveliness, we said,
Children in wonder watching the stars,
Is the aim and the end.

Being but men, we walked into the trees.

Dylan Thomas

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE

BIOCHEMISTRY

Doctor of Philosophy

STUDIES ON ANTIBIOTIC RESISTANCE IN PRODUCING MICROORGANISMS

by Christopher James Herbert

The aminoglycoside phosphotransferase (*APH*) gene of *B. circulans* has been cloned in *E. coli* (CC1105) using the vector pBR322 and has been found to confer antibiotic resistance. The 2.7 kb Sal I fragment containing the *APH* gene was ligated into the plasmid SLP1.2 and transformed into *S. lividans* 66 (CL1023), where it was also able to confer antibiotic resistance. When the phosphotransferase activity was assayed in crude extracts of *B. circulans*, *E. coli* CC1105 and *S. lividans* CL1023, using a variety of aminoglycosides as phosphate acceptors, a similar spectrum of activity was found in each strain.

Using the methods of Sanger and Messing, the sequence of the 2.7 kb Sal I fragment was determined. An examination of the DNA sequence allowed the protein sequence to be deduced. When the deduced protein sequence was compared with those from *S. fradiae*, TN5 and TN903, significant homology was found, indicating that all the phosphotransferases may have a common origin.

An examination of the sequence either side of the coding region allowed the identification of possible promotors, a ribosome binding site and a terminator. One of the promotors was similar to the σ^{37} promotors of *B. subtilis*, showing that the *B. circulans APH* gene may be under developmental control.

ABBREVIATIONS

APH	Aminoglycoside phosphotransferase
ATP	Adenosine triphosphate
BCIG	5-Bromo-4-chloro-3-indodyl-β-D-galactopyranoside
bp	Base pairs
CDTA	1,2-Diaminocyclohexane-N,N,N',N'-tetra-acetic acid
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dTTP	Deoxythymidine triphosphate
ddATP	Dideoxyadenosine triphosphate
ddCTP	Dideoxycytidine triphosphate
ddGTP	Dideoxyguanosine triphosphate
ddTTP	Dideoxythymidine triphosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid
h	Hours
IPTG	Isoproply-β-D-thiogalactopyranoside
kb	Kilobase
mRNA	Messenger ribonucleic acid
min	Minutes
O.D.	Optical density

PEG	Polyethylene glycol
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulphate
TES	(N-tris[hydroxymethyl]methyl-2-aminoethane sulphonic acid)
Tris	Tris(hydroxymethyl)aminoethane
u.v.	Ultra violet

CONTENTS

CHAPTER 1

THE GENETICS OF ANTIBIOTIC METABOLISM

1.1 Introduction

Microorganisms produce several thousand different antibiotics. These are part of a larger group of compounds called secondary metabolites but they have been the subject of more intensive study because of their economic and social importance. The value of these compounds to the producing organism is often unclear; typically they are produced at specific stages in the life of a colony or culture, and so their metabolism is an aspect of differential gene expression, and hence differentiation (Demain, 1974). They are called secondary metabolites to distinguish them from the primary metabolites of the cell, such as amino acids, nucleotides and co-factors, which have to be produced throughout the life of a culture. Although the antibiotics are often considered as one group of compounds, it should be remembered that chemically they consist of a wide variety of structures : β lactams, aminoglycosides, polyethers and peptides, whilst biologically they act at many sites within the cell : protein synthesis, cell wall synthesis, transcription and DNA replication (Hopwood, 1981).

1.2 Antibiotic resistance determinants

As all antibiotic producing organisms have to avoid being killed by their own antibiotics, they are a rich source of antibiotic resistance determinants. Many give rise to the same enzymic activities as those found in clinical isolates (reviewed in Davies & Smith, 1978), and this has led to the hypothesis that antibiotic-producing organisms were the original source of clinically significant antibiotic resistance (Benveniste & Davies, 1973). This hypothesis has been tested using a variety of immunological and nucleic acid hybridisation techniques, but as yet it has not been possible to produce evidence to support it (Davies & Smith, 1978; Courvalin *et al.*, 1978).

1.3 Possible methods of genetic determination of antibiotic metabolism

One of the interesting aspects of antibiotic metabolism is the range of antibiotics produced by similar strains and, conversely, the variety of strains that produce the same antibiotic (Hopwood, 1978). Thus, the aminoglycoside neomycin is produced by *S. fradiae*, *S. albogriseolus* and *M. chalcae*, whereas different isolates of *S. gresius* will produce streptomycin, grisin, candicidin, cephamycin, nonactin or cyclohexamide. This apparent 'looseness' of the connection between an antibiotic and a strain can be used to argue two things: firstly, that the production of a particular antibiotic is of no special significance to the organism (however, that an antibiotic is produced may be important); secondly, that there might be something unusual in the genetic determination of antibiotic production which makes it more movable, or facilitates its transfer from one strain to another. The simplest explanation of this would be that the genes involved in antibiotic metabolism are encoded on plasmids, and this has been suggested in several cases (reviewed in Hopwood, 1978). In general, three systems have been proposed for the genetic determination of antibiotic biosynthesis and resistance : (i) all structural and regulatory genes are chromosomal; (ii) all (or most) of the structural and regulatory genes are plasmid encoded, and, finally (iii) plasmid encoded genes control the expression (or high level expression) of chromosomal structural genes (Hopwood, 1979).

Any study of the genetics of antibiotic metabolism is greatly facilitated by a system for gene exchange and a well-defined linkage map. For this reason, the most reliable information has come from *S. coelicolor*, which is genetically the best characterised strain; unfortunately, none of the antibiotics which it produces is commercially important but it provides several useful examples.

1.4 Clusters of chromosomal genes in antibiotic biosynthesis

S. coelicolor produces several antibiotics. One of these is a binapthoquinone called actinorhodin, which is also an acid-base indicator. Normally, this accumulates in the mycelia as a red

pigment but, when the mycelia are exposed to ammonia vapour, the pigment turns blue. This provides a visual means of detecting mutants unable to synthesise actinorhodin (*act*) and allowed the isolation of 76 mutants specifically blocked in its synthesis (Rudd & Hopwood, 1979). These mutants could be placed into a series of seven phenotypic classes, based on their ability to cosynthesise actinorhodin in pairwise combinations, the accumulation of pigmented precursors or shunt metabolites and their antibiotic activity. By examining the polarity of the cosynthetic reactions, six of the classes of mutants could be placed in a linear sequence of biosynthetic blocks. One representative of each class was mapped and all the mutations were allocated to a short segment of the chromosome between *gua*A and *his*D. This strongly suggests that the genes for actinorhodin biosynthesis form an uninterrupted cluster on the chromosome, and possibly an operon. Although actinorhodin itself is not an important antibiotic, it is a good example of the large class of polyketide antibiotics.

The isolation of mutants unable to synthesise actinorhodin allowed the identification of a second red pigment antibiotic in the mycelia of *S. coelicolor*; this appeared to be a prodigiosin-like compound. In studies similar to those outlined above, 37 mutants unable to produce the red pigment were isolated, assigned to five phenotypic classes and mapped (Rudd & Hopwood, 1980). Again, all mapped to a short segment of the chromosome indicating that an uninterrupted gene cluster, or operon, was involved. Genetic evidence showed that there was no overlap between the biosynthetic pathways of the red pigment and actinorhodin.

In *S. rimosus* mutations in six genes involved in the biosynthesis of oxytetracycline have been mapped to two clusters on opposite sides of the circular linkage map (Rhodes *et al.*, 1981).

1.5 Plasmid encoded genes in antibiotic biosynthesis

Methylenomycin A is a third antibiotic produced by *S. coelicolor*; all of the mutations that specifically abolish methylenomycin biosynthesis are linked to the plasmid SCP1. Genetically SCP1 is very well characterised but physically it is only poorly characterised. This is because it is very difficult to isolate. The exact reason for this is not known but it could be due to its large size, approximately 150 kb.

Strains which lack SCP1 do not produce methylenomycin and when SCP1 is transferred to closely-related strains, such as *S. lividans* 66, methylenomycin production and resistance are always transferred (Hornemann & Hopwood, 1980). This clearly demonstrated that all the genes for methylenomycin resistance and production are located on the plasmid SCP1. Several chromosomal mutations that abolish methylenomycin synthesis have been isolated, but these also have other effects, such as the loss of sporulation and actinorhodin biosynthesis. These pleiotrophic mutations probably cause a general disturbance of secondary metabolism or differentiation.

1.6 Plasmid control of chromosomal genes in antibiotic biosynthesis

The most often quoted case of plasmid genes controlling the expression of chromosomal genes is in the production of chloramphenicol by *S. venezuelae*. Initially, a set of auxotrophic markers was used to construct a circular linkage map. Mutants unable to produce chloramphenicol (*cpp*) were isolated at a frequency of 2-5% after treatment with the 'curing agent' acriflavin. These could not be assigned to the linkage map and so were assumed to be plasmid linked mutations or, more probably, reflect the loss of a plasmid (Akagawa *et al.*, 1975). Later, it was found that these *cpp* strains did in fact produce approximately 10% of the wild-type levels of chloramphenicol. The isolation of a larger collection of *cpp* strains after u.v. irradiation, nitrosoguanidine mutagenesis, acriflavin treatment or incubation at high temperature did give mutants that produced no chloramphenicol. These were isolated at frequency of 0.1-0.2% and all mapped to the chromosome between *met* and *ilv*; again, the acriflavin and high temperature treatments yield some strains that produced ~ 10% of the normal levels of chloramphenicol and could not be mapped to the chromosome (Akagawa *et al.*, 1979). When strains which were presumed to have lost plasmids were crossed with each other, no recombinants that produced the wild-type level of chloramphenicol were found. However, when these presumed plasmidless strains were crossed with strains bearing chromosomal *cpp* mutations, a high frequency of recombinants that produced the wild-type level of chloramphenicol was obtained. This was taken to indicate that the plasmid genes controlled

the high level expression of the chromosomal genes. Some physical evidence for a plasmid in the wild-type but not the 'cured' strains was later obtained (Okanishi & Umezawa, 1978).

More recently, Ahmed and Vining isolated a plasmid from a chloramphenicol-producing strain of *S. venezuelae*. This plasmid, pUC3, had previously been characterised by restriction endonuclease digestion (Malik & Reusser, 1979). When treated with curing agents, such as acriflavin, chloramphenicol non-producing mutants could be isolated at high frequency. However, when these strains were examined, they still contained a plasmid which was indistinguishable from pUC3 on restriction endonuclease digestion. Plasmidless derivatives were isolated by forming and regenerating protoplasts, a method known to cause the loss of some plasmids in *Streptomyces*. The absence of the plasmid was confirmed by the inability of plasmid DNA to hybridise to digested total DNA in Southern Blot experiments. The plasmidless derivatives all produced wild-type levels of chloramphenicol, showing that the plasmid pUC3 is not required for chloramphenicol biosynthesis (Ahmed & Vining, 1983).

When the digests of total DNA were subjected to gel electrophoresis, bright bands could be seen in the background smear. These were thought to represent reiterated sequences in the DNA. The pattern produced by digestion of DNA from the wild-type and plasmidless strains was identical but DNA from one of the acriflavin-treated strains gave a different pattern, indicating that re-arrangements of the chromosomal DNA had occurred. These re-arrangements in the acriflavin-treated strain may be the cause of the reduced chloramphenicol production and could also explain why it is not possible to map these mutations. In the light of these results, we can say that pUC3 is not involved in chloramphenicol biosynthesis. As no connection has been established between pUC3 and the plasmid isolated by Okanishi, it is still possible to propose a tentative role for a plasmid in chloramphenicol biosynthesis. However, the detection of DNA re-arrangements in an acriflavin-treated strain argues against a simplistic interpretation of the earlier results.

S. reticuli harbours a 48 M dalton plasmid and plasmid genes have been postulated to be involved in a variety of functions in this strain (Schrempf & Goebel, 1979). Treatment with curing agents caused the loss of leukomycin production, melanin production and the ability to sporulate, in various combinations. These phenotypes have been correlated with changes in the plasmid (insertions and deletions) and plasmid loss. Recently, it has been demonstrated that in some of the cured derivatives the plasmid has been lost, but chromosomal re-arrangements had also occurred (Schrempf, 1982). In this case, no causal relationship between changes in the plasmid and phenotypic changes has been established. Until this can be done, the role of the plasmid in leukomycin production, melanin production and sporulation is open to considerable doubt.

The work on *S. venezuelae* and *S. reticuli* introduces an important aspect of much of the research in this area. Almost any antibiotic-producing streptomycete will yield antibiotic non-producing variants when treated with curing agents; but the loss of antibiotic production is almost always accompanied by a disruption of differentiation, such as the inability to sporulate. This is frequently taken as an indication of plasmid involvement in antibiotic production, resistance and sometimes sporulation. Often these observations are not further investigated but, in the case of neomycin biosynthesis by *S. fradiae*, the initial observations were pursued and this work provides a suitable cautionary tale.

Neomycin belongs to a family of related aminoglycoside antibiotics (Pearce & Rinehart, 1981) and plasmid genes have been implicated in the metabolism of several of them (listed in Hopwood, 1979). In their first experiments, Yagisawa *et al.* were able to isolate mutants of *S. fradiae* which were unable to synthesise neomycin and sensitive to neomycin, after cultures had been treated with the curing agent acridine orange (Yagisawa *et al.*, 1978). Two plasmids of M.W. 16 M daltons and 22 M daltons could be isolated from the wild-type cells but not the cured derivatives. This led to the hypothesis that the plasmids were involved in neomycin production

and resistance. In later experiments, the same group were unable to isolate the two plasmids from the wild-type strain, and could only isolate a single plasmid of 46 M daltons; a similar plasmid, but containing an insertion, was also present in the 'cured' strain. A mutant, which was more resistant to neomycin and appeared to be a high copy number mutant, also carried a similar but not identical insertion in the plasmid. Two strains that were sensitive to neomycin, unable to produce neomycin, and did not contain a plasmid were isolated after treatment with novobiocin or protoplast regeneration. The neomycin resistance genes had been cloned and it was possible to demonstrate that they were not located on the plasmid, so, in the case of neomycin resistance at least, the plasmid was not directly involved. The second hypothesis developed was that the plasmid controlled the expression of chromosomal genes (Komatsu *et al.*, 1981). Although not described in the published data, many of the strains used in these experiments were morphologically abnormal (J. Davies, personal communication).

Recently, Chung was able to isolate a plasmid from *S. fradiae*. Restriction enzyme digestion patterns indicate that it was the same as the large plasmid isolated by Komatsu *et al.* (Chung, 1982). Also, it was possible to demonstrate that the plasmid was the prophage form of the actinophage ϕSF1. During the course of this work, strains that did not contain the plasmid, and were morphologically normal, were isolated : neomycin production and resistance in these strains was the same as in the wild-type (Shiau-Ta Chung, personal communication). This seems conclusive evidence that the plasmid is not involved in neomycin resistance or production.

1.7 Conclusion

Several general conclusions can be drawn from this brief survey of the genetics of antibiotic production and resistance. Firstly, experiments that rely solely on the loss of various phenotypes, as a result of treatment with 'curing agents' to demonstrate plasmid involvement, are at best uninterpretable and at worst misleading.

Secondly, in all the cases where it has been possible to obtain definitive information, the changes in antibiotic production and resistance could be examined in the absence of other phenotypic changes, such as abnormal morphology. Thirdly, the main effect of 'curing agents' seems to be to cause re-arrangements of chromosomal and plasmid DNA, leading to a general disruption of secondary metabolism and differentiation. At the moment, the study of these strains does not seem very fruitful but, as our understanding of the developmental biology of these organisms increases, they may be helpful in elucidating the link between secondary metabolism and differentiation. Finally, the production of methylenomycin A by *S. coelicolor* still remains the only well-established case of a plasmid being involved in antibiotic biosynthesis; *a priori* it would seem unlikely that no other cases exist. That they have not been found is possibly more due to the methods that have been used to search for them.

1.8 Aims of this work

The aim of this project is to try and clarify some of the points mentioned above, by cloning and sequencing the aminoglycoside phosphotransferase (*APH*) gene of the butirosin-producing *B. circulans*. Butirosin is a member of the neomycin family of aminoglycoside antibiotics, and it is the *APH* gene which is responsible for the strain's resistance to its own antibiotic. In the genus *Bacillus*, much information is now available about the expression of developmentally regulated genes. An important feature of this is the production of different RNA polymerase σ factors at different stages of the developmental cycle. These bind to the RNA polymerase and enable it to recognise unusual promotor sequences; thus, genes with these promotor sequences are transcribed at different stages of the developmental cycle (Losick & Pero, 1981). The regulatory sequences of the *APH* gene from *B. circulans* may fit into this pattern and show how the antibiotic resistance is linked into the developmental cycle.

The *APH* genes from two transposons have already been sequenced (Beck *et al.*, 1982; Oka *et al.*, 1981). A comparison of the *B. circulans APH* gene sequence with the other sequences would be the most direct test of homology, and hence the hypothesis that the antibiotic-producing organisms are the source of the resistance determinants. Finally, the cloned gene would be useful as a probe in hybridisation experiments to monitor DNA re-arrangements involving the *APH* gene that might be induced by 'curing agents'.

While this work was in progress, similar studies were carried out on the *APH* gene from *S. fradiae* (Thompson & Gray, 1983).

CHAPTER 2

MATERIALS AND METHODS

2.1 Chemicals

All chemicals were obtained from BDH Ltd., Poole, Dorset or Sigma London Chemical Company Ltd., Kingston-upon-Thames, Surrey, with the exception of those listed below :

Aldrich Chemical Co. Ltd., Gillingham, Dorset

- Acrylamide

Amersham International Ltd., Amersham, Bucks

- $\gamma[^{32}P]$ adenosine triphosphate
- $\alpha[^{32}P]$ deoxyadenosine triphosphate (400 Ci/mmol PB 164)

BCL - The Boehringer Corporation (London) Ltd., Lewes, East Sussex

- Adenosine triphosphate
- Deoxynucleoside triphosphates
- DNA-polymerase Klenow fragment

Bethesda Research Laboratories (UK) Ltd., Cambridge

- Agarose
- Restriction endonucleases
- T_4DNA ligase

Difco Laboratories, Detroit, Michigan, USA

- Bacto-agar
- Bacto-inorganic salts starch agar
- Bacto-peptone
- Bacto-tryptone
- Bacto-yeast extract
- Nutrient agar
- Nutrient broth

Hopkins and Williams, Chadwell Heath, Essex

- Repelcote

Kodak Ltd., Bristol, Avon

DX-80 Developer

Indicator stop solution

Rapid acid hardener

Unifix powder

Miles Laboratories Ltd., Stoke Poges, Slough

λ ($C_{1}s_{57}$) DNA

Oxoid Ltd., Basingstoke, Hants.

Dorset egg slopes

Malt extract

Potato dextrose agar

Tryptic soya broth

Tryptic soyatone agar

P.L. Biochemicals Ltd., Northampton

Dideoxynucleoside triphosphates

M13 15 base universal primer

(100 µl of the primer is sufficient for approximately 100 reactions)

Table 2.1. List of strains used in this work

Escherichia coli JM103 - Δ *lac pro*, *sup* E, *thi*, *str* A, *end A*, *sbc* B15, *hsd* R4
F' *tra* D36, *pro* AB, *lac* I, *lac* ZΔM15

Escherichia coli V517 - Marcina *et al.*, 1978

Escherichia coli HB101 - *pro*, *leu*, *thi*, *lac* Y, *hsd* R, Rec A, *rps* L20, *ara* 14, *gal* K2, *xyl* 5, *mtl* 1, *sup* E44

CC1100	-	HB101 harbouring pBR322
CC1101	-	HB101 harbouring pCH1
CC1103	-	HB101 harbouring pCH3
CC1104	-	HB101 harbouring pCH4
CC1105	-	HB101 harbouring pCH5
CC1106	-	HB101 harbouring pCH6
CC1107	-	HB101 harbouring pCH7
CC1108	-	HB101 harbouring pCH8
CC1109	-	HB101 harbouring pCH9
CC1110	-	HB101 harbouring pCH10
CC1111	-	HB101 harbouring pCH11

Bacillus circulans 3312 - produces Butirosins

Streptomyces fradiae ATCC10745 - produces Neomycins

Streptomyces ribosdificus ATCC21294 - produces Ribostamycin

Streptomyces rimosus forma *paromomycinus* ATCC14827 - produces Paromomycins

Streptomyces lividus ATCC21178 - produces Lividomycins

Streptomyces lividans 66 :

M180	-	*S. lividans* 66 harbouring the plasmid SLP1.2
CL1023	-	*S. lividans* 66 harbouring pCLI
CL1024	-	*S. lividans* 66 harbouring pCL2

E. coli JM103 was obtained from Bethesda Research Laboratories (UK) Ltd.

E. coli V517 was a gift from E.M. Lederberg, Plasmid Reference Centre, Stanford University, California.

E. coli HB101 was a gift from M.J. Pocklington, Department of Biology, University of Southampton.

B. circulans 3312 was a gift from Dr. J. Davies, BIOGEN SA.

S. lividans 66 and M180 were gifts from Professor D.A. Hopwood, John Innes Institute, Norwich.

All aminoglycoside producing *Streptomyces* were obtained from the American Type Culture Collection, Maryland, USA.

2.2 Media

LB Broth :

per dm^3	Bacto-tryptone	10 g
	Bacto-yeast extract	5 g
	NaCl	5 g

LB Agar :

LB Broth set with 1.5% Bacto-agar

2TY :

per dm^3	Bacto-tryptone	16 g
	Bacto-yeast extract	10 g
	NaCl	5 g

H Top agar :

per dm^3	Bacto-tryptone	10 g
	NaCl	8 g
	Bacto-agar	6 g

The above recipes are all taken from Miller (1972).

Minimal GT agar :

(Vogel & Bonner, 1956)

50 x salts	$MgSO_4.7H_2O$	10 g
per dm^3	Citric acid	100 g
	K_2HPO_4	500 g
	$NaNH_4HPO_4.4H_2O$	175 g
	Distilled water	670 ml

Thiamine 10 mg/ml in distilled water

per dm^3	Glucose	2 g
	Bacto-agar	15 g

After autoclaving sterile 50 x salts and thiamine (to a concentration of 10 µg/ml) were added.

YEME :

(Chater *et al.*, 1982)

per dm^3	Difco-yeast extract	3 g
	Difco-peptone	5 g
	Oxoid malt extract	3 g
	Glucose	10 g

R2YE :

(Chater *et al.*, 1982)

	Bacto-yeast extract	2 g
	$MgCl_2.6H_2O$	4 g
	K_2SO_4	100 mg
	1 M TES pH 7.2	10 ml
	Sucrose	41 g
	Glucose	4 g
	Bacto-agar	9 g
	Distilled water	350 ml
After autoclaving add :		
	KH_2PO_4 (1 g/100 ml)	2 ml
	$CaCl_2.2H_2O$ (23.6 g/80 ml)	4 ml
	L Proline (12 g/100 ml)	10 ml

Minimal agar :

per dm^3	Asparagine	0.5 g
	K_2HPO_4	0.5 g
	$MgSO_4.7H_2O$	0.2 g
	$FeSO_4.7H_2O$	0.01 g
	Bacto-agar	15.0 g

After autoclaving, sterile glucose solution is added to give a final glucose concentration of 1% (w/v).

2.3 Maintenance of cultures

All *E. coli* strains were maintained on Dorset egg slopes, stored at room temperature and subcultured once a year.

E. coli JM103 was always streaked out on a minimal GT agar plate prior to use to ensure that the F factor had been maintained.

B. circulans was maintained on Difco-nutrient agar slopes supplemented with 1% glucose, stored at 4°C and subcultured every six months.

Streptomyces strains were stored at 4°C, subcultured every six months and maintained on the following media :

S. fradiae ATCC10745 - Bacto-inorganic salts starch agar.

S. ribosdificus ATCC21294 - Potato dextrose agar.

S. rimosus forma *paromomycinus* ATCC14827 - Tryptic soyatone agar.

S. lividus ATCC21178, *S. lividans* 66 and its derivatives - R2YE

M13 bacteriophages mp8 and mp9 were maintained as plaques in H-Top agar overlays and stored at 4°C.

2.4 Sterilisation

Solutions were sterilised by autoclaving at 120°C and 15 lbs/sq. inch for 15 min. Velvet pads were wrapped in aluminium foil and autoclaved at 120°C and 15 lbs/sq. inch for 1 h. Antibiotic solutions were made by dissolving the antibiotic in sterile water.

2.5 Restriction endonuclease cleavage

Restriction endonuclease digests were carried out in low, medium or high salt buffers, according to the method of Davies *et al.*, 1980.

2.6 Alcohol precipitation of DNA

DNA was precipitated from solution by adding 0.1 volumes of 3 M sodium acetate pH 5.5, followed by 2.5 volumes of ethanol; the mixture was chilled in a solid CO_2/isopropanol bath for at least 20 min followed by centrifugation at 12,000 g in a microcentrifuge for 5 min. The supernatant was discarded and the pellet washed with ethanol and centrifuged as before. Finally, the pellet is drained, dried in a vacuum desiccator and dissolved in a suitable buffer.

2.7 Conditions for the use of T_4 DNA ligase

Typical ligations were carried out using a vector concentration of 20 µg/ml and a ratio of vector to insert DNA of 1 : 4. The reaction was carried out at 4°C in the following buffer : $MgSO_4.7H_2O$ 10 mM, DTT 10 mM, ATP 1 mM and Tris-HCl pH 7.6 20 mM. A sterile 10 x solution

of Tris-HCl and $MgSO_4.7H_2O$ could be stored at room temperature, with ATP and DTT added just prior to use.

2.8 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out in a home-made "submarine gel" tank using Tris-acetate buffer (Davis *et al.*, 1980).

50 x buffer per dm^3	Tris base	242 g
	Glacial acetic acid	57 ml
	Disodium EDTA	37.2 g
	pH 8.1	

The concentration of agarose used was 0.7-1.2%, depending on the size of the fragments. Normally, the electrophoresis was carried out for 18 h, using a voltage gradient of 1 volt/cm. Molecular weight standards for linear fragments were obtained by digesting λ DNA with Hind III (Daniels & Blattner, 1982) and for supercoiled plasmids the family plasmids isolated from *E. coli* V517 were used (Marcina *et al.*, 1978). Before loading, the samples were mixed with 0.2 volumes of 20% Ficoll and 0.01% Orange G dissolved in water. After electrophoresis, the gels were stained in 1 µg/ml ethidium bromide for 30 min and destained in water for 15 min.

2.9 Photography of gels

The stained gels were placed on a midrange u.v. transilluminator (Ultra-Violet Products Inc., California, USA) and photographed using a POLAROID CU-5 Land Camera with a Kodak Wrattan 23A filter. A 1 minute exposure at F 4.7 was usually sufficient. The film used was Polaroid 665 pack film; this produces a print and a negative. The film was processed according to the manufacturers' instructions.

2.10 Determination of DNA

For large amounts of DNA the OD_{260} was measured; an OD_{260} of 1 was assumed to be equivalent to 50 µg of DNA/ml. For small amounts of DNA, aliquots were run on an agarose gel with standards and the amount of DNA in the band was estimated visually.

2.11 Determination of protein

Protein was assayed by the method of Bradford (1976) using bovine γ-globulin as standard.

2.12 Determination of aminoglycoside phosphotransferase activity

Aminoglycoside phosphotransferase activity was measured in crude sonicated cell extracts using the radioactive phosphate transfer method of Haas & Dowding, 1975.

Note:

The phosphotransferase assays were carried out for 20 min in a reaction mixture which contained 7.5 nMoles of ATP and approximately 3 nMoles of antibiotic. The transfer of 10,000 dpm to the antibiotic is equivalent to the production of approximately 0.8 nMoles of phosphorylated antibiotic.

Table 2.2. Large scale preparation of plasmid DNA from *E. coli*

1. Grow plasmid bearing strains overnight at 37°C in 10 ml of LB broth containing appropriate antibiotic(s).

2. Inoculate 500 ml of LB broth containing appropriate antibiotic(s) with the 10 ml overnight culture, grow at 37°C on an orbital shaker until $OD_{600} > 1.0$ (For strains containing plasmids which do not chloramphenicol, amplify, e.g. *E. coli* V517, grow overnight at 37°C and proceed to stage 4).

3. Add chloramphenicol to 200 µg/ml and continue shaking at 37°C overnight.

4. Harvest cells by centrifugation at 7,500 g and 4°C for 10 min, decant the supernatant and drain the pellet thoroughly.

5. Resuspend the pellet in 10 ml of TE (50 mM Tris 10 mM EDTA pH 8.0) containing 25% (v/v) sucrose and transfer to a 100 ml conical flask.

6. Add 2 ml of freshly prepared lysozyme (10 mg/ml) in TE and incubate at 37°C for 15 min with gentle shaking. After two minutes, transfer the flask to an ice bath and incubate for a further 10 min.

7. Add 10 ml of 0.25 M EDTA, pH 8.0, mix gently and incubate on ice for 10 min.

8. Add 2 ml of 20% (w/v) SDS in TE and mix gently until the suspension clarifies. This may require warming in a 37°C water bath.

9. Add 6 ml of 5 M sodium chloride in TE, mix, transfer to a 50 ml centrifuge tube and store on ice for 2 h.

10. Centrifuge at 27,000 g for 1 h and decant the supernatant into a sterile 50 ml polycarbonate centrifuge tube.

11. Add a 0.5 volume of 30% (w/v) polyethyleneglycol 6000 in TE, mix and store overnight at 4°C.

Table 2.2 continued

12. Centrifuge at full speed in a bench centrifuge for 5 min, discard the supernatant and drain the pellet thoroughly.

13. Dissolve the pellet in 7.5 ml of sterile TE (using a 5 ml automatic pipette, if necessary).

14. Dissolve 8.6 g of caesium chloride in the DNA solution; this may now go opaque due to polyethyleneglycol precipitating.

15. Spin at 27,000 g, 20°C, for 20 min to sediment the polyethyleneglycol. Carefully decant the clear supernatant.

16. Add 1 ml of ethidium bromide (10 mg/ml) in TE and mix. Measure the refractive index and, if necessary, adjust to 1.3925 with TE or solid caesium chloride.

17. Transfer to 14 ml thin-walled polycarbonate centrifuge tubes. These must be full so overlay the DNA solution with liquid paraffin.

18. Centrifuge at 46,000 g and 20°C for 48-72 h.

19. The DNA bands can be visualised using longwave u.v.; the lower band (plasmid DNA) can be removed by inserting a needle and syringe just below it and carefully sucking the band into the syringe.

20. The ethidium bromide is removed by extracting four times with isopropanol (saturated with caesium chloride solution) and the caesium chloride is removed by dialysing against one change of 1.5 dm^3 of TE overnight at 4°C.

21. Plasmid solutions are stored in TE at +4°C or -20°C.

Table 2.3. Transformation of *E. coli* with plasmid DNA
(Based on Davis *et al.*, 1980)

1. Grow *E. coli* HB101 overnight in 10 ml of LB broth at 37°C.
2. Inoculate 20 ml of LB broth (in a 100 ml conical flask) with 0.2 ml of the overnight culture of *E. coli* HB101.
3. Incubate on an orbital shaker at 37°C until the OD_{600} ~ 0.6 (2-3 h).
4. Harvest the cells by centrifugation at full speed in a bench centrifuge for 5 min.
5. Resuspend in 10 ml of 50 mM $CaCl_2$ and incubate on ice for 1 h.
6. Sediment the cells by centrifugation at full speed in a bench centrifuge for 5 min.
7. Resuspend the cells in 1 ml of 50 mM $CaCl_2$, add 0.1 ml of the cell suspension to 0.1 ml of the DNA solution and incubate on ice for 10 min.
8. Heat treat the transformation mixtures at 45°C for 2 min.
9. Add 1 ml of L broth and incubate at 37°C for 1 h to allow expression of the transformed genes.
10. Spread 0.1 ml aliquots of the transformed cells onto plates supplemented with appropriate antibiotic(s). Incubate at 37°C. Colonies should be visible after 18-36 hours' growth.

Table 2.4. Rapid isolation of plasmids from *E. coli* strains growing on solid media

(Based on Kado & Liu, 1981)

1. Inoculate LB-agar plates with patches of the *E. coli* strains to be screened and incubate at 37°C overnight.
2. Using a wooden toothpick, scrape about 0.5 cm^2 of the plate and suspend the cells in 100 µl of lysing solution in a microcentrifuge tube.
3. Incubate at 55°C for 1 h.
4. Add 100 µl of phenol/chloroform and vortex to mix.
5. Centrifuge at 12,000 g in a microcentrifuge for 5 min to separate the phases.
6. Take 30 µl of the aqueous phase and mix with 10 µl of loading solution and electrophorese in an agarose gel in the usual way.

Lysing solution contains 3% SDS and 50 mM Tris pH 12.5. The pH of the solution is important for efficient removal of chromosomal DNA. It was adjusted with 2 M NaOH and measured with alkalite pH papers (BDH).

Phenol/chloroform (50/50) is prepared by melting phenol at 50°C and adding an equal volume of chloroform.

Table 2.5. Rapid isolation of plasmids suitable for restriction endonuclease digestion from *E. coli* strains

(ISH-Horowicz, 1982)

1. Inoculate 5 ml of LB broth containing an appropriate selecting antibiotic with the desired strain of *E. coli* and incubate at 37°C overnight.
2. Pipette 1.5 ml of the overnight culture into a microcentrifuge tube and centrifuge at 12,000 g in a microcentrifuge for 1 min. The remainder of the culture may be stored at 4°C and used later.
3. Discard the supernatant, briefly respin to remove liquid from the wall of the tube and remove the remaining medium by aspiration.
4. Resuspend the pellet by vortexing in 100 µl of lysozyme solution and store for 5 min at room temperature.
5. Add 200 µl of alkaline SDS, mix by rapidly inverting 2 or 3 times, but do not vortex. Store on ice for 5 min.
6. Add 150 µl of ice-cold potassium acetate (~ pH 4.8), close the cap of the tube and vortex in an inverted position to mix. Store on ice for 5 min.
7. Centrifuge at 12,000 g in a microcentrifuge for 5 min and then transfer the supernatant to a fresh tube.
8. Add 400 µl of phenol/chloroform, mix by vortexing and separate the phases by spinning at 12,000 g in a microcentrifuge.
9. Transfer the supernatant to a fresh tube and add 800 µl of ethanol at room temperature. Mix by vortexing and stand at room temperature for 2 min.
10. Centrifuge at 12,000 g for 5 min in a microcentrifuge; discard the supernatant and drain the pellet.
11. Add 1 ml of 70% ethanol, vortex briefly and centrifuge as before.
12. Discard the supernatant and dry the pellet in a vacuum desiccator.

Table 2.5 continued

13. Dissolve the pellet in 50 µl of TE (10 mM Tris 1 mM EDTA, pH 8.0) containing pancreatic RNase (20 µg/ml) which has been heated to 90°C for 10 min to destroy any DNase activity.

14. 10-15 µl of the above solution should be sufficient for digestion by restriction enzymes and gel electrophoresis.

Lysozyme solution contains 50 mM glucose, 10 mM EDTA, 25 mM Tris.HCl (pH 8.0) and 4 mg/ml lysozyme, which is added just before use.

Alkaline SDS contains 1% SDS and 0.2 M NaOH. This should not be kept for more than two weeks.

Potassium acetate (~ pH 4.8) is made by mixing 60 ml of 5 M potassium acetate, 11.5 ml of glacial acetic acid and 28.5 ml of distilled water.

2.13 Manipulation of *Streptomyces*

The methods described in this section for the manipulation of *Streptomyces* are all designed for *S. lividans* 66 and its derivatives. If other strains are used, incubation conditions may need to be changed. The problem most often encountered is difficulty in forming protoplasts when the mycelia are treated with lysozyme. This can usually be overcome in one of three ways : the concentration of lysozyme can be increased; the length of the incubation can be increased; or the concentration of glycine in the growth medium can be increased. As a guide, the concentration of glycine that inhibits the growth by 50% is usually sufficient to make the mycelia lysozyme-sensitive.

The practical and theoretical aspects of gene cloning in *Streptomyces* have recently been reviewed by Chater *et al.* (1982).

Table 2.6. Isolation of *Streptomyces* total DNA
(Based on Chater *et al.*, 1982)

1. Inoculate 50 ml of YEME medium containing 34% sucrose, 5 mM magnesium chloride and 0.5% glycine with 0.25 ml of a dense spore suspension and incubate on an orbital shaker at 30°C for 36-48 h.
2. Harvest the mycelium by centrifugation at 18,000 g and 4°C for 10 min.
3. Resuspend the mycelium in 4 ml of 25% sucrose, 50 mM Tris-HCl pH 8.0.
4. Add 0.6 ml of lysozyme (10 mg/ml) in TE and incubate at 30°C for 1 h, occasionally mixing by pipetting.
5. Add 1.2 ml of 0.5 M EDTA and 0.7 ml of predigested pronase (10 mg/ml) in TE (10 mM Tris 1 mM EDTA, pH 8.0). Incubate at 30°C for 5 min.
6. Add 3.6 ml of SDS (3.3%) in TE, mix gently (do not vortex) and incubate at 37°C for 2 h.
7. Shake with 6 ml of TE saturated phenol for 10 min.
8. Add 6 ml of chloroform and shake for 5 min, transfer to a 50 ml pyrex centrifuge tube and separate the phases by centrifugation at full speed in a bench centrifuge for 10 min.
9. Transfer the aqueous phase to a new tube and repeat 7 and 8.
10. Transfer the aqueous phase to a new tube and add RNase to 40 µg/ml (the RNase should be heated to 90°C for 10 min to eliminate DNase activity) and incubate at 37°C for 1 h
11. Add 0.2 volumes of 5 M sodium chloride, mix gently, add 0.5 volumes of 30% (w/v) PEG 6000 in TE, mix gently and leave at 4°C overnight.
12. Centrifuge at full speed on a bench centrifuge for 10 min to pellet the DNA.

Table 2.6 continued

13. Gently dissolve the pellet in 5 ml of TE containing 0.1 M sodium chloride, using a 5 ml automatic pipette, if necessary.

14. Add 0.5 ml of 3 M sodium acetate pH 4.5, and 14 ml of ethanol. Mix and leave overnight at -20°C.

15. Sediment the DNA by centrifugation at 22,000 g and 4°C for 20 min.

16. Wash the pellet twice with ice-cold ethanol.

17. Drain the pellet and dry in a vacuum desiccator. Dissolve the pellet in sterile TE and store at -20°C.

Table 2.7. Large scale preparation of *Streptomyces* plasmid DNA (Based on Bibb *et al.*, 1977)

1. Inoculate 500 ml of YEME medium containing 34% sucrose and 0.1% magnesium chloride with 1 ml of a dense spore suspension. Incubate on an orbital shaker at 30°C for 48 h.
2. Harvest the mycelium by centrifugation at 7,500 g and 4°C for 10 min. Wash with 10% glycerol and sediment the mycelium by a further centrifugation at 7,500 g and 4°C for 10 min.
3. Resuspend the mycelium in 50 ml of TE (10 mM Tris 1 mM EDTA, pH 8.0) containing 34% sucrose and transfer to a 1 dm^3 beaker in a water bath at 30°C.
4. Add 10 ml of 0.25 M EDTA pH 8.0 and mix.
5. Add 10 ml of lysozyme (25 mg/ml) in TE, mix and incubate for 15 min at 30°C.
6. Transfer the mixture to an ice bath and add 50 ml of ice-cold TE containing 34% sucrose, followed by 30 ml of ice-cold 0.25 M EDTA pH 8.0. Mix gently.
7. Dispense 30 ml aliquots into pre-cooled 50 ml centrifuge tubes.
8. Add 3.6 ml of SDS (10%) in TE to each tube and mix by inversion. This will lyse the cells and the suspension will become very viscous.
9. Add 7.2 ml of 5 M sodium chloride to each tube and mix by inversion. This will be difficult as the suspension is very viscous at this point, but it is important that it is not shaken vigorously.
10. Store at 4°C overnight.
11. Centrifuge at 27,000 g and 4°C for 30 min. Carefully decant the supernatant into a sterile conical flask and add 0.5 volumes of 30% (w/v) PEG 6000 in TE. Mix and store overnight at 4°C to precipitate the DNA.

Table 2.7 continued

12. Collect the precipitate by centrifugation at 1,500 g and 4°C for 10 min. Carefully decant the supernatant and allow the pellet to drain thoroughly.

13. Dissolve the pellet in 7.5 ml of sterile TE (using a 5 ml automatic pipette, if necessary).

14. Dissolve 8.6 g of caesium chloride in the DNA solution. This may now go opaque due to PEG coming out of solution.

15. Centrifuge at 27,000 g and 20°C for 20 min to sediment the PEG. Carefully decant the clear supernatant.

16. Add 1 ml of ethidium bromide (10 mg/ml) in TE and mix. Measure the refractive index and, if necessary, adjust to 1.3925.

17. Transfer to 14 ml thin-walled polycarbonate centrifuge tubes. These must be full so overlay the DNA solution with liquid paraffin.

18. Centrifuge at 46,000 g and 20°C for 48-72 h.

19. The DNA bands can be visualised using longwave u.v. The lower band (plasmid DNA) can be removed by inserting a needle and syringe just below it and carefully sucking the band into the syringe.

20. The ethidium bromide is removed by extracting four times with isopropanol (saturated with caesium chloride solution) and the caesium chloride is removed by dialysing against one change of 1.5 dm^3 of TE overnight.

21. Plasmid solutions are stored in TE at +4°C or -20°C.

2.14 Small scale preparation of *Streptomyces* plasmid DNA

The large scale preparation of *Streptomyces* plasmid DNA is an expensive and time-consuming process, taking eight days to complete. For screening recombinant clones, it is preferable to use a rapid plasmid preparation based on the alkaline lysis of protoplasts. Plasmid DNA prepared in this way still contains significant amounts of chromosomal DNA but is suitable for agarose gel electrophoresis, transformation and simple restriction enzyme digests.

Table 2.8. Rapid preparation of plasmid DNA from *Streptomyces* (Based on Chater *et al.*, 1982)

1. Inoculate 50 ml of tryptic soya broth with 1 ml of a dense spore suspension and heat to 50°C for 10 min to stimulate germination. Incubate on an orbital shaker at 30°C for 24 h.
2. Harvest the mycelium by centrifugation at 18,000 g and 4°C for 10 min. Resuspend the pellet in 20 ml of 10.3% sucrose and centrifuge as before.
3. Resuspend the mycelium in 2 ml of P medium containing lysozyme (1 mg/ml) and incubate at 30°C for 60 min, mixing by pipetting every 15 minutes.
4. Add 2 ml of hot lytic mixture and mix by pipetting up and down vigorously. Incubate at room temperature for 15 min, mixing occasionally.
5. Add 4 ml of phenol/chloroform and mix by vortexing for 30 sec.
6. Transfer to a glass centrifuge tube and separate the phases by centrifugation at full speed on a bench centrifuge for 10 min.
7. The upper phase can be loaded directly onto an agarose gel without the addition of loading solution.
8. For restriction enzyme digestion and transformation, the residual phenol and SDS must be removed by extracting three times with diethyl-ether, followed by repeated ethanol precipitation.

Lytic mixture contains : 10 ml of 5 M NaOH
20 ml of 10% SDS, 25 ml of 0.25 M CDTA and 45 ml of distilled water.

P medium contains : 41 g of sucrose, 100 mg of K_2SO_4, 810 mg of $MgCl_2.6H_2O$, 10 ml of 1 M TES pH 7.2 and 370 ml of distilled water. After autoclaving, 2 ml of KH_2PO_4 (1 g/100 ml) and 4 ml of $CaCl_2.2H_2O$ (29.5 g/80 ml) are added.

2.15 Transformation in *Streptomyces*

A prerequisite for the effective study of genes from *Streptomyces* is the ability to efficiently reintroduce isolated DNA into cells. It had been observed that early log phase cultures of *S. virginae* are competent for DNA uptake (Konvalinkova *et al.*, 1977) and this was used to transfect the mycelium with actinophage DNA. However, this was not developed into a generalised system for transfection. The ability of many *Streptomyces* strains to produce stable protoplasts that can regenerate to mycelial forms was well known (Okanishi *et al.*, 1974); also, it had been demonstrated that polyethylene glycol increased the efficiency of transfection of *E. coli* spheroplasts by bacteriophage DNA (Morita *et al.*, 1977). This lead to the development of more generalised systems for the transformation (Bibb *et al.*, 1978) and transfection (Suarez & Chater, 1980) of *Streptomyces* protoplasts in the presence of polyethylene glycol.

Table 2.9. Preparation, transformation and regeneration of *Streptomyces* protoplasts

(Based on Bibb *et al.*, 1978)

1. Inoculate 25 ml of YEME medium containing 34% sucrose, 5 mM magnesium chloride and 0.5% glycine with 0.1 ml of a dense spore suspension. Incubate on an orbital shaker at 30°C for ~ 30 h.

2. Harvest the mycelium by centrifugation at full speed in a bench centrifuge for 5 min. If this is not sufficient to pellet the mycelium, it indicates that the culture has not grown enough to be used for protoplast formation and transformation.

3. Wash the mycelium by resuspending the pellet in 10 ml of 10.3% sucrose. Centrifuge as before and repeat the wash.

4. Resuspend the mycelium in 4 ml of P medium containing lysozyme (1 mg/ml) and incubate at 30°C until the mycelia begin to break up and release protoplasts (15-60 min).

5. Pipette the suspension three times to help disrupt any mycelial clumps and release the protoplasts. Incubate for a further 15 min.

6. Add 5 ml of P medium and pipette as before.

7. Filter through cotton wool to remove mycelial fragments and sediment the protoplasts by centrifugation at full speed in a bench centrifuge for 5 min.

8. Wash the protoplasts by resuspending them in 4 ml of P medium and centrifuge as before.

9. Resuspend the protoplasts in 4 ml of P medium and dispense appropriate amounts to other tubes for transformation. Typically, 0.1 ml of this suspension is used for each regeneration plate.

10. Pellet the protoplasts as before and resuspend in the few drops of P medium remaining after the supernatant has been decanted.

Table 2.9 continued

11. Add the DNA in less than 20 µl of buffer. Immediately add 0.5 ml of PEG 1000 solution (2.5 g of PEG 1000 dissolved in 7.5 ml of P medium).

12. After 1 min, add 5 ml of P medium and sediment the protoplasts by centrifugation as before.

13. Resuspend in P medium (0.1 ml/plate) and spread on R2YE plates that have been dried at 42°C for 30 min.

14. Incubate at 28°C. Pocks can usually be seen after 2-3 days and regeneration is complete after 7-10 days.

The proportion of non-protoplasted cells in the final protoplast suspension can be calculated by making a series of parallel dilutions in P medium and distilled water. These are spread on R2YE plates and grown at 28°C for 4 days. As the protoplasts are osmotically unstable, only non-protoplasted cells will grow from the dilutions in water. These can be counted and compared with the total count from the dilutions in P medium.

2.16 M13 cloning and chain termination sequencing

Most of the methods in this section are based on the M13 cloning and sequencing manual produced by G. Winter and A.R. Coulson, M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge, England.

Table 2.10. Transfection of *E. coli* JM103 with M13

1. Inoculate 10 ml of 2TY with a colony of *E. coli* JM103 from a minimal GT plate and incubate at 37°C until the cells reach OD_{600} ~ 0.5. This will take about 4 h but will vary with the size of the inoculum.

2. Prepare dilutions of the phage solution. If it was produced from a chronically infected culture of *E. coli* JM103, dilutions ~ 10^{-9} will be required.

3. Prepare 3 ml aliquots of molten H top agar at 45°C by adding 25 µl of IPTG (24 mg/ml) in water and 25 µl of BCIG (25 mg/ml) in dimethyl formamide.

4. To the prepared H top agar, add 0.1 ml of phage dilution and 0.3 ml of *E. coli* JM103, mix and pour onto a warm minimal GT plate.

5. Leave at room temperature for 30 min to allow the H top agar to set, then incubate inverted at 37°C.

6. After approximately 18 hours' incubation, the plaques should be visible.

Table 2.11. Preparation of M13 replicative form DNA

1. Grow *E. coli* JM103 overnight in 20 ml of 2 TY at 37°C.
2. Inoculate 1.5 ml of 2 TY with a drop of the overnight culture of *E. coli* JM103.
3. Using a sterile toothpick, inoculate a blue plaque of the desired M13 strain into the 1.5 ml culture.
4. Incubate at 37°C on a rotary shaker for 5-6 hours.
5. Transfer the culture to a 1.9 ml microfuge tube and centrifuge for 5 min in a microfuge (12,000 g).
6. Pour the supernatant, which contains the phage, into a new microfuge tube and store on ice until needed for the bulk transfection.
7. Inoculate 500 ml of 2TY with 10 ml of the overnight culture of *E. coli* JM103 and grow on an orbital shaker at 37°C until the OD_{600} ~ 0.5 (3-4 h).
8. Add 0.5 ml of the freshly prepared phage solution to the 500 ml culture of *E. coli* JM103 and continue to incubate on an orbital shaker at 37°C overnight.
9. Harvest the cells by centrifugation at 7,500 g, 4°C, for 10 min.
10. Resuspend the cells in 200 ml of 150 mM NaCl 100 mM EDTA pH 8.0 and store on ice for 30 min.
11. Sediment the cells by a further centrifugation at 7,500 g, 4°C, for 10 min. The procedure is now the same as a large scale plasmid isolation from *E. coli* and continues from point 5 of Table 2.2.

2.17 Isolation of DNA from agarose gels

Several methods were used to isolate DNA from agarose gels during the course of this work. The first was to cut out the section of the gel containing the fragment, dissolve it in concentrated KI, add ethidium bromide and centrifuge to equilibrium. The DNA could be visualised under long-wave u.v., and extracted from the gradient in the normal way. However, this was an expensive and time-consuming procedure, taking a total of four days. The second method was to carry out the electrophoresis using low gelling temperature agarose (B.R.L.). The desired band was cut out and the agarose melted by heating to 65°C. The agarose could then be removed by repeated phenol extraction and the DNA isolated by ethanol precipitation. This was a much more convenient method but yields were often low (< 50%) and only a small proportion of the ends of the isolated DNA were left intact, as judged by the inability of fragment prepared in this way to undergo homopolymerisation. The last and most successful method used relies on the ability of the water in the gel to disrupt its structure when frozen (Table 2.12). This method proved to be a reliable way to isolate DNA of good quality and in high yield (~75%).

Table 2.12. Isolation of DNA from agarose gels

1. Prepare and run the gel in the normal way. Stain with ethidium bromide and visualise the DNA using a mid-range u.v. trans-illuminator.

2. Cut out the band containing the DNA to be isolated, dice into small pieces (~ 3 mm^3) and put into a disposable syringe (2 ml or 5 ml depending on the amount of material), which has two GF-C filters (Whatman) at the bottom of the barrel. Replace the plunger but do not compress the gel.

3. Cover the barrel with aluminium foil and place in a bath of solid CO_2/isopropanol for 30 min.

4. Remove the syringe from the solid CO_2/isopropanol bath and leave at room temperature for 20 min to thaw.

5. Remove the aluminium foil and squeeze the liquid out of the syringe into a siliconised glass tube, or a microfuge tube, taking care not to rupture the GF-C filters and allow particles of agarose through.

6. If the volume of liquid is too large at this stage, it can be reduced by extraction with iso-butanol.

7. Add half volume of TE (10 mM Tris 1 mM EDTA ph 8.0) saturated phenol and mix. Centrifuge for 2 min in a microfuge (12,000 g) to separate the phases.

8. Transfer the upper aqueous phase to a new microfuge tube and extract twice with 0.5 ml of di-ethyl ether.

9. Add 0.1 volume of 3 M NaAc pH 5.5 and 2.5 vol of ethanol. Mix and place in a solid CO_2/isopropanol bath for 15 min to precipitate the DNA.

10. Centrifuge for 5 min in a microfuge (12,000 g), drain the pellet thoroughly and wash with 1 ml of ethanol.

11. Centrifuge for 5 min in a microfuge (12,000 g), drain the pellet and dry in a vacuum desiccator.

12. Dissolve the pellet in an appropriate amount of TE and store frozen at -20°C.

When a small amount of a specific restriction fragment is needed for ligation into M13, it is often inconvenient to prepare a large amount by the previous method but it can be adapted to a small scale. A small section of the gel containing the fragment is cut out, put into a microfuge tube and frozen in a solid CO_2/isopropanol bath for 20 min. The gel is then allowed to thaw and centrifuged for 15 min in a microcentrifuge, the supernatant is removed with a pipette and made up to 200 µl with TE. This can be phenol extracted and ethanol precipitated in the usual way. All of the DNA produced by this method is used in a single ligation with M13 RF DNA.

2.18 Cloning of DNA fragments in M13 replicative form DNA

When DNA fragments are cloned in M13 replicative form DNA (RF), very small amounts of DNA are used in a 10 µl reaction volume. It has been found that the yield of plaques is higher when the ligations are done in sealed glass tubes rather than microfuge tubes. The tubes used are melting point tubes, one end of which has been heated and pulled out to a capillary. Typically, 1 µl of x 10 T_4 ligase buffer (Section 2.7), 3-10 ng of appropriately cut M13 RF DNA, 10-30 ng of insert DNA, ~ 0.1 units of T_4 DNA ligase, and sterile distilled water to a total of 10 µl are pipetted onto parafilm; these are sucked into the drawn out melting point tube. The non-capillary end is sealed in a low flame and allowed to cool; this draws the liquid into the tube so that the capillary end can also be sealed. These were incubated at 4°C for at least 12 h before being used for transfection. In each experiment, two controls were always performed : vector DNA alone to ensure that the cleavage of the RF form was complete, and vector DNA with ligase. This should yield only blue plaques unless an exonuclease has damaged the ends of the vector, in which case white plaques will be generated so that it is not possible to detect clones with inserts.

Table 2.13. Transfection of *E. coli* JM103 with M13 replicative form DNA

1. Grow *E. coli* JM103 overnight in 10 ml of 2TY at 37°C.
2. Inoculate 20 ml of 2TY (in a 100 ml conical flask) with 0.2 ml of the overnight culture of *E. coli* JM103.
3. Incubate on an orbital shaker at 37°C until the OD_{600} ~ 0.6 (2-3 h).
4. Harvest the cells by centrifugation at full speed in a bench centrifuge for 5 min.
5. Resuspend the pellet in 10 ml of 50 mM $CaCl_2$ and incubate on ice for 1 h.
6. Sediment the cells by centrifugation at full speed in a bench centrifuge for 5 min.
7. Resuspend the pellet in 2 ml of 50 mM $CaCl_2$, transfer 200 µl aliquots to microfuge tubes and add the ligation mixtures (usually 10 µl each). Incubate on ice for 1 h.
8. Heat treat the transformation mixtures for 2 min at 45°C. During this time prepare 3 ml aliquots of molten H top agar at 45°C by adding 25 µl of IPTG (24 mg/ml) in water and 25 µl of BCIG (25 mg/ml) in dimethyl formamide.
9. Add the transformation mixtures to the prepared H top agar, mix and pour onto a warm minimal GT plate.
10. Leave at room temperature for 30 min to allow the H top agar to set, then incubate, inverted, at 37°C.
11. After approximately 18 hours' incubation, the plaques should be clearly visible. A white plaque indicates that an insert has been introduced into the phage.

Table 2.14. Preparation of single stranded DNA template for sequencing

1. Grow *E. coli* JM103 overnight in 10 ml of 2TY at 37°C.
2. Inoculate 50 ml of 2TY with 1 ml of the overnight culture of *E. coli* JM103 and dispense 1.5 ml aliquots into sterile test tubes.
3. Using sterile toothpicks, inoculate individual white plaques into the seeded 1.5 ml cultures.
4. Incubate on an orbital shaker at 37°C for 5-6 h.
5. Transfer the cultures to 1.9 ml microfuge tubes and centrifuge for 5 min in a microfuge (12,000 g) to sediment the cells.
6. Pour the supernatant into another microfuge tube, making no effort to completely transfer the liquid as this might disturb some of the pellet.
7. Add 200 µl of 2.5 NaCl, 20% PEG 6000, mix by inversion and leave at room temperature for 15 min to allow the phage particles to precipitate.
8. Centrifuge for 5 min in a microfuge (12,000 g), discard the supernatant and spin again for a few seconds to bring the remaining liquid to the bottom of the tube. This can then be removed using a drawn out pasteur pipette.
9. Add 100 µl of TE (10 mM Tris, 1 mM EDTA pH 8.0) and vortex to resuspend the pellet. Add 50 µl of TE saturated phenol and vortex again. Leave at room temperature for 5 min, vortex again and centrifuge for 1 min in a microfuge (12,000 g) to separate the two phases.
10. Transfer the aqueous layer to another microfuge tube.
11. Add 10 µl of 3 M NaAc pH 5.5 and 250 µl of ethanol. Store at -20°C overnight to precipitate the DNA.

Table 2.14 continued

12. Centrifuge for 5 min in a microfuge (12,000 g), drain the pellet (not visible) thoroughly and wash with 1 ml of ethanol.

13. Centrifuge for 5 min in a microfuge (12,000 g), drain the pellet and dry in a vacuum desiccator.

14. Dissolve the pellet in 30 µl of TE and store frozen at -20°C. If necessary, 5 µl can be run on an agarose gel to check that the isolation has worked.

2.19 Stock solutions and buffers required for sequencing

Tris-Borate-EDTA

The sequencing gels are run in a continuous buffer system of Tris (133 mM),boric acid (44 mM) and EDTA (2.5 mM) pH 8.8, (TBE). It is convenient to make up a 10 x concentrated stock of the following composition : 162 g Tris base, 27.5 g boric acid and 9.5 g of disodium EDTA.$2H_2O$, made up to 1 dm^3 with distilled water. The pH of this stock should be approximately 8.8 and it should not be adjusted before use.

Acrylamide stock

Ths standard gel for sequencing is 7 M urea and 6% acrylamide. For this, an acrylamide urea stock is made in the following way : 288 g of urea (Analar), 34.2 g of acrylamide and 1.8 g of bis-acrylamide are made up to 500 ml with distilled water and dissolved with gentle heating (< 40°C). 20 g of washed Amberlite MB-1 resin is stirred with this solution for 30 min, the resin is removed by filtration, 60 ml of 10 x TBE is added and the stock is made up to 600 ml with distilled water. The stock is stable for several weeks when stored at 4°C. For longer sequence runs, it is sometimes desirable to use a 7 M urea 4% acrylamide gel. This can be made in a similar way, using 288 g of urea, 22.8 g of acrylamide and 1.2 g of Bis-acrylamide

10% ammonium persulphate

1 g of ammonium persulphate was dissolved in 10 ml of distilled water. This is stable for several months at 4°C.

N,N,N',N',-tetramethylethylenediamine

TMED was obtained from a commercial supplier and was kept at 4°C.

10 x annealing buffer

10 x annealing buffer consists of 100 mM Tris pH 8.5 and 100 mM magnesium chloride. This was stored frozen.

Termination mixes

The ratio of dideoxynucleoside triphosphate to deoxynucleoside triphosphate in the termination mixes is crucial if an easily readable sequence gel is to be produced. The final balancing of the termination mixes must be done using sequencing reactions but the following list of concentrations and dilutions gives a good initial approximation. The standard ddNTP solutions are 10 mM in TE; these are diluted to give working stocks according to the following protocol.

ddNTP	ddTTP	ddCTP	ddGTP	ddATP
Vol. of 10 mM ddNTP (µl)	42.0	5.0	10.0	6.7
Vol of TE (µl)	458	495	490	493
Final conc. of ddNTP (µM)	830	100	200	135

The dNTP's are stored at a concentration of 0.5 mM in TE. These are mixed according to the following protocol to give the dNTP°.

	dTTP°	dCTP°	dGTP°	dATP°
0.5 mM dTTP (µl)	5	100	100	100
0.5 mM dCTP (µl)	100	5	100	100
0.5 mM dGTP (µl)	100	100	5	100

The termination mixes are made by mixing equal volumes of the appropriate ddNTP working stock and dNTP°, so 200 µl of ddTTP working stock mixed with 200 µl of dTTP° would give 400 µl of T termination mix. All nucleotide solutions were stored frozen at -20°C.

dNTP chase

This is a solution in TE which is 0.5 mM with respect to each deoxynucleoside triphosphate and is stored frozen at -20°C.

Formamide dyes

100 ml of formamide (Analar) was stirred with 5 g of washed Amberlite MB-1 resin, the resin is removed by filtration, 30 mg of xylene cyanol FF and 30 mg of bromophenol blue were added. Finally, disodium EDTA is added to a concentration of 20 mM (0.75 g).

2.20 Preparation of sequencing gels

In order to obtain a sequence that can be read easily and without ambiguities, the gel must be poured without bubbles, as they cause distortion in the bands. Also, when the plates are separated at the end of the electrophoresis run, the gel must adhere to one plate only so that it is not stretched or torn. To give the gel the best chance of survival, the plates must be very clean and the inside of the notched plate is siliconised to encourage the gel to adhere to the unnotched plate when they are separated. The plates are scrubbed in detergent, rinsed and allowed to dry. One side of the notched plate is then treated with 'Repelcote' and allowed to dry. The plates are then washed twice with distilled water and finally twice with ethanol and allowed to dry. The well formers and side spacers are cut from 'Plastikard'; the side spacers are about 1 cm wide and the well formers contained 32 teeth approximately 4 mm wide. When new 'Plastikard' is coated with a substance which inhibits the polymerisation of acrylamide, this can be removed by scrubbing in undiluted detergent.

Just prior to pouring the gel, the plates are taped up and 40 ml of gel mix per gel is allowed to come to room temperature. The polymerisation is initiated by adding 300 µl of 10% ammonium persulphate and 50 µl of TMED to the gel mix. Using a 25 ml pipette, the gel mix is carefully poured down one side of the plates, then allowed to run across the bottom of the plates and, finally, the centre and other side of the plates are allowed to fill. If bubbles form, they can sometimes be fished out with a spare side spacer. The well former is now inserted and the gel clips put in place to compress the gel to the correct thickness (approximately 0.3 mm). The gel should be left for at least

an hour before it is used; when set up in the electrophoresis tank, the well slots should be flushed out just prior to loading to remove any urea that has leached out. The electrophoresis apparatus used was a 40 cm x 20 cm slab gel electrophoresis unit from Raven Scientific.

Table 2.15. Sequencing protocol. This procedure is designed for seven clones.

1. Mix 24 µl of single stranded universal primer, 11 µl of 10 x annealing buffer and 11 µl of distilled water.

2. Dispense 6 µl aliquots of annealing mix into 7 microfuge tubes and add 5 µl of single stranded DNA template. Place in a water-bath at 70°C, turn off the waterbath and allow to cool to < 37°C.

3. Dispense 4 x 9 µl (9 µCi) of $\alpha[^{32}P]$ dATP into siliconised glass tubes (1 x 5 cm) and dry down in a vacuum desiccator.

4. Arrange 28 capless microfuge tubes in a 7 x 4 matrix. Spin the cooled clone annealing mixes to bring the liquid to the bottom of the tubes; allocate each clone to a horizontal row of 4 tubes and dispense 2 µl of clone annealing mix to each tube in the row.

5. Resuspend the $\alpha[^{32}P]$ dATP from one tube in 18 µl of T termination mix, using a 100 µl Hamilton syringe fitted with a PB 600-1 dispenser. Dispense 2 µl of T termination mix to each tube in the first vertical row of 7 tubes.

6. Repeat for the other 3 vertical rows, using C, G and A termination mixes. Spin to mix.

7. Dilute 8 µl of DNA polymerase Klenow fragment (1 unit/µl) with 56 µl of distilled water; dispense 2 µl to each tube. Spin to mix and start the reactions. Incubate at room temperature for 15 min.

8. Dispense 2 µl of dNTP chase to each tube, spin to mix and incubate at room temperature for 15 min.

9. Dispense 4 µl of formamide dye mix to each tube, spin to mix and boil in a waterbath for 3 min to denature the samples.

10. Using a drawn out capillary tube, load half of each sample onto the first gel.

Table 2.15 continued

11. Boil the tubes for another minute and load the rest of the samples onto the second gel.

12. The gels are run at a constant power of 40 watts per gel, using an LKB 2197 power pack. Typically, the first gel is run for 1.5 h and the second for 4 h.

13. When a run is finished, the gel plates are separated and the gel should stay attached to the unnotched plate. This is placed in a tray of 10% acetic acid for 10 min to fix the DNA in the gel.

14. The gel is drained and transferred to a piece of 3 mm chromatography paper. The exposed surface is then covered with cling-film. A commercial gel drier is now used to dry the gel prior to autoradiography.

15. The cling-film is peeled from the surface of the dried gel and both the long and short gels are used to expose a sheet of X-ray film in an X-ray cassette.

16. After 12-48 hours of exposure at room temperature, the X-ray film is developed in the usual way.

2.21 Autoradiography

All manipulation of undeveloped film was done under a safelight with a Kodak 6B filter. Dried sequencing gels were exposed to Fuji RX X-ray film in an X-ray Okamoto express X-ray cassette at room temperature for 12-48 h. If a sequencing experiment has been successful, it is not necessary to use aids such as intensifying screens or pre-fogged film. These would also lead to a loss of resolution. The materials for developing the X-ray film were prepared as follows : developer was prepared by diluting 400 ml of Kodak DX-80 developer with 1600 ml of water; stop solution by diluting 80 ml of Kodak indicator stop bath with 1920 ml of water; and fixer by dissolving 260 g of Kodak unifix powder in 2000 ml of water. The film was developed by placing it in developer for 4 min, a stop bath for 1 min and fixer until the film had cleared (5-10 min). The film was then washed extensively with water and allowed to drain dry.

Table 2.16. Timetable for a sequencing experiment.

09.30	Pour two gels.
10.00	Anneal primer/templates.
10.15	Aliquot $\alpha[^{32}P]$ dATP and dry down.
11.15	Aliquot cooled primer/templates.
11.30	Add termination mixes to dried $\alpha[^{32}P]$ dATP, dissolve and add to primer/templates.
11.40	Dilute Klenow Enz and add to mixes to start the reactions.
11.55	Add chase.
12.00	Set up gels in apparatus.
12.10	Add formamide/dye to samples, boil for 3 min and flush out well slots.
12.15	Load first gel.
12.25	Boil samples, load second gel.
13.55	Fix and dry first gel.
16.30	Fix and dry second gel.
17.30	Set up autoradiograph.

CHAPTER 3

CLONING OF THE *B. circulans APH* GENE

3.1 Introduction

Since 1970, great advances have been made in our understanding of complex biological systems, such as the control of gene expression. The development of *in vitro* cloning techniques has been fundamental to many of these advances. The main object of *in vitro* gene cloning is to reduce the physical and genetic complexity of the genome by eliminating all genetic information that does not pertain to the biological property under investigation. In this respect, it can be considered a modern counterpart of the older *in vivo* techniques, such as the isolation of F-prime factors and transducing phages (Sherratt, 1981).

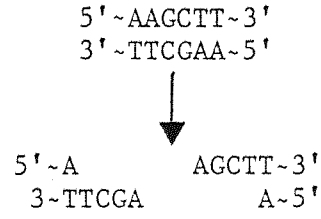
In general, there are five requirements for a successful gene cloning system : a method for generating DNA fragments, a vector to carry the fragments, a way of joining the vector and fragments together, a method for introducing the hybrid molecules into the host cell and a way of selecting the desired clones. Once cloned, the DNA fragment can be used in a variety of ways : as a template in the various gene expression systems (Zubay *et al.*, 1970; Roozen *et al.*, 1971; Sancar *et al.*, 1979), as a substrate for site directed mutagenesis (Taniguchi & Weissmann, 1978; Smith & Gillam, 1981), and as a starting point for sequence analysis (Sanger *et al.*, 1977a; Messing, 1982).

3.2 Generation of fragments

The DNA segment to be cloned will initially be present as part of a long DNA molecule, such as a large plasmid or chromosome. This must be fragmented in a way which allows the insertion of the fragments into a specific site in the vector. The simplest way to do this is to digest the DNA with a Class II restriction endonuclease, which produces a staggered break, and so leaves sticky ends (Fig. 3.1). These are short (usually 4 base) 3'- or 5'- single stranded termini that can

anneal with complementary termini produced by digestion with the same enzyme, and so can easily be inserted into appropriately cut vector.

Fig. 3.1. Recognition sequence for Hind III



If a section of DNA is 50% G+C rich, then enzymes that recognise an hexanucleotide sequence would be expected to cut the DNA once every 4,100 bases. Assuming that the average gene is 1-2 kb long, then there is a 25-50% chance that the enzyme will cut within the gene; the longer the segment to be cloned, the greater the chance of the enzyme cutting within the gene and the experiment being unsuccessful. This problem can be overcome by using a more random method to generate the fragments. A near random sample of fragments can be generated by partial digestion with 'frequent cutting' enzymes. These normally recognise a tetranucleotide sequence and are expected to cut the DNA approximately once every 200 bases. Random fragments can be produced by physically breaking the DNA; this can be done by controlled shearing or by sonnication (Britten *et al.*, 1974). Both of these methods produce fragments with single stranded termini, which must be filled-in (polished) to give blunt ends (using DNA polymerase Klenow fragment or T_4 DNA polymerase) before they can be inserted into the vector.

3.3 Vectors

Vectors are plasmid or phage replicons that can be maintained within a host cell and that can carry passenger DNA. The first vectors used for *in vitro* studies were natural plasmids that had one or more restriction sites in non-essential regions, such as the Eco RI site in the plasmid Col EI (Clarke & Carbon, 1975). These natural vectors were soon replaced by others specifically constructed

Fig. 3.2. Restriction map of pBR322

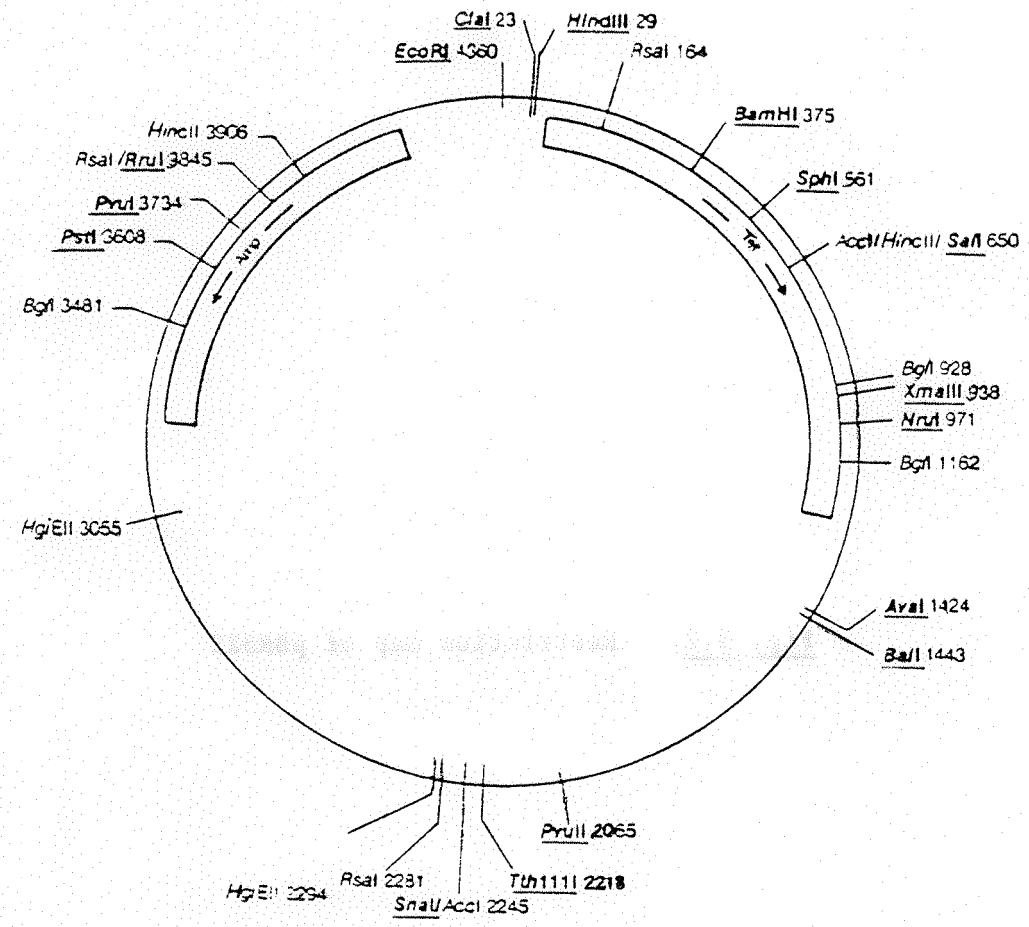


Fig. 3.2. Restriction map of pBR322.

Locations of the known restriction enzyme recognition sequences occurring three or fewer times in pBR322 DNA. Locations of the first bases within recognition sequences are expressed in nucleotide numbers counting from the centre of the Eco RI recognition sequence. The enzymes that cleave this DNA once only are underlined. (Sutcliffe, 1979).

for the *in vitro* cloning of DNA. The most well-known is pBR322 (Bolivar *et al.*, 1977), which possess many of the desirable features of a plasmid vector. The replicon used for pBR322 was a ColEI-like replicon which undergoes chloramphenicol amplification (Clewell, 1972); when chloramphenicol is added to late log-phase cultures, chromosomal DNA replication soon stops but plasmid replication continues increasing the relative proportion of plasmid DNA. This enables the plasmid to be purified in large amounts and consequently allows large amounts of the cloned DNA to be obtained. Using standard transformation procedures, pBR322 can easily be introduced into *E. coli*, and the antibiotic resistance genes which it carries allows the positive selection of transformants. There are unique sites for several Class II restriction endonucleases which can be used for inserting fragments. Some of these sites are within the antibiotic resistance genes (Fig. 3.2) and so, when the foreign DNA is inserted, the gene is inactivated. This can be used to select clones containing recombinant plasmids. Phage genomes have also been adapted and used as vectors (Murray & Murray, 1974). In any given experiment, the vector used will be determined by the needs of that experiment. Some vectors are specifically designed for certain experiments; the phage M13 vectors used in chain termination DNA sequencing are a good example of this (Messing, 1982).

3.4 Methods for joining DNA molecules

There are two main methods for joining DNA fragments to vectors: homopolymer tailing and ligation with DNA ligase. Homopolymer tailing is a convenient way of joining DNA fragments produced by physical disruption to vectors but it can also be used when the fragments are produced by restriction enzyme digestion. The 3'- ends of the fragment are extended with a polymer of one deoxynucleotide using deoxynucleotidyl terminal transferase. The 3' ends of the vector are then extended with a polymer of the complimentary deoxynucleotide. The two homopolymers are able to anneal and the gaps are filled in by the recipient cell after transformation (Lobban & Kaiser, 1973). If the restriction site in the vector and the nucleotide used for tailing are chosen carefully, the restriction site can be recreated, allowing the

Fig. 3.3. Insertion by homopolymer tailing

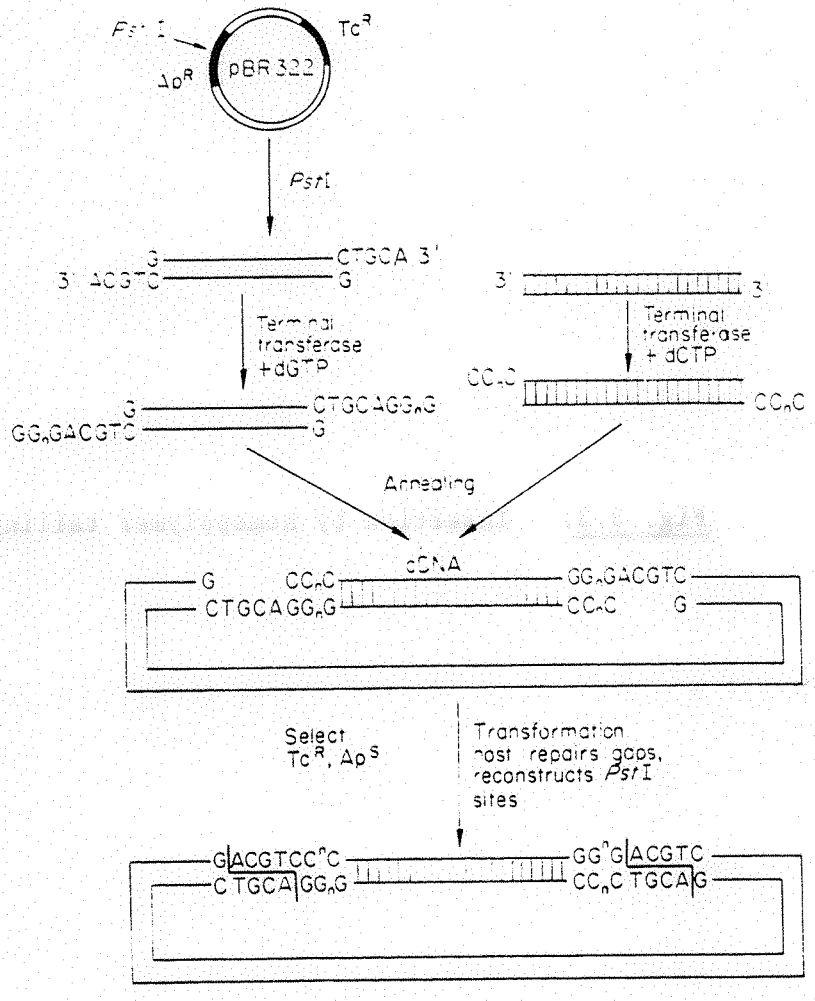


Fig. 3.3. Insertion by homopolymer tailing. Oligo (dC) DNA inserted into Pst I cut pBR322 so that the Pst I site is recreated.

inserted fragment to be excised (e.g. dG for Pst I cleaved vector molecules) (Fig. 3.3). One advantage of homopolymer tailing is that circularisation of the vector is not possible, so all transformants contain inserts. In most cloning experiments, fragments are joined to the vector using DNA ligase (Mertz & Davis, 1972). This is an enzyme produced by *E. coli* or phage T_4 infected *E. coli*, which is able to seal nicks between adjacent nucleotides in a duplex DNA chain (T_4 DNA ligase is normally used as it is able to ligate both sticky and blunt ends; the *E. coli* enzyme is only able to ligate sticky ends). In the simplest situation, the DNA to be cloned and the vector are cleaved with the same enzyme. However, some frequent cutters produce single stranded termini that can anneal with the termini produced by infrequent cutters (e.g. Sau 3A and Bam HI). This means that Sau 3A cut DNA can be ligated into Bam HI cut vector. A disadvantage of this method is that the Bam HI site is not necessarily recreated, so the fragments cannot easily be excised (Fig. 3.4).

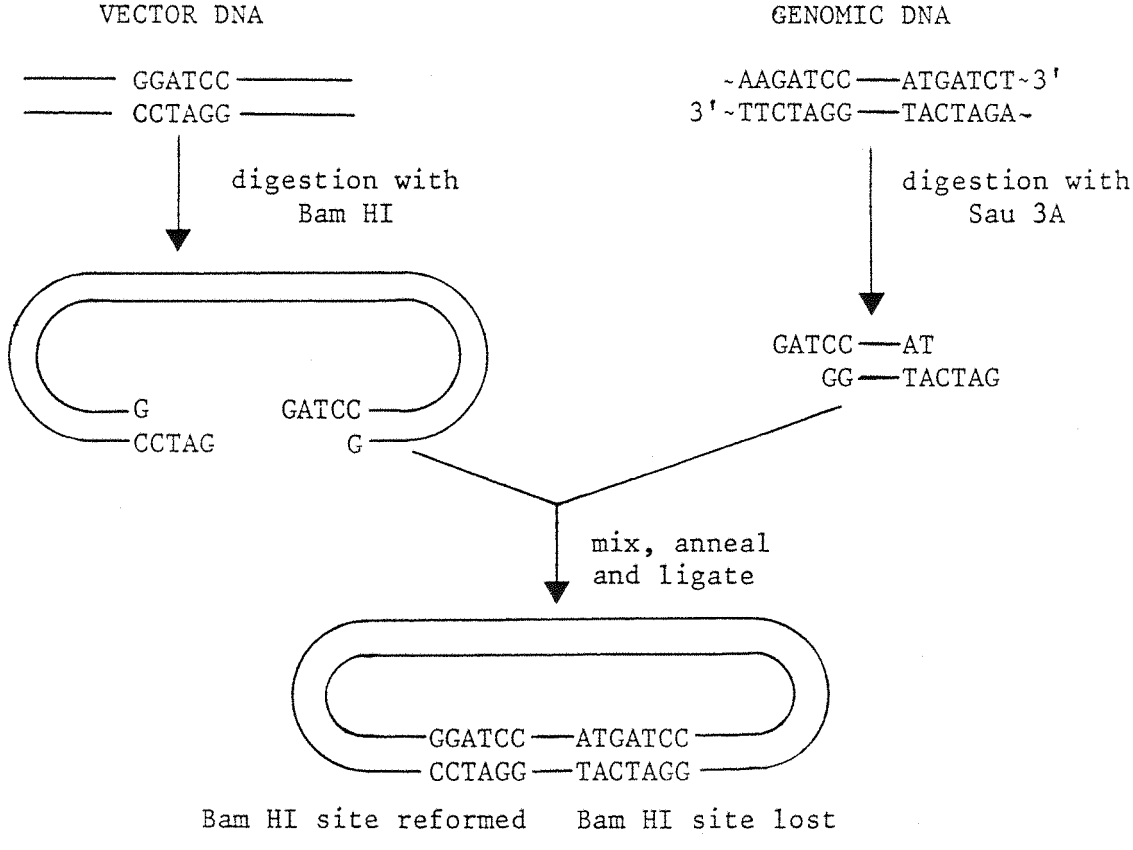


Fig. 3.4. Cloning Sau 3A fragments into a Bam HI site.

Fragments that have been produced by physical disruption and subsequently 'polished' can be inserted into vector cut to give blunt ends (e.g. Sma I). There are two disadvantages to this method : it is an inefficient process and again the fragments cannot easily be excised. These problems can be overcome by the use of linkers. Linkers are short oligodeoxynucleotides that contain the recognition sequence for one or more restriction enzymes (Greene *et al.*, 1975). The linkers are blunt end ligated to the insert DNA (this is quite efficient as a large molar excess of linker can be added) and the mixture is treated with an appropriate restriction enzyme to generate the termini required. The fragments can then be cloned in the normal way.

3.5 Introduction of hybrid molecules into the host

The introduction of hybrid molecules into the host bacteria is relatively simple if the host is *E. coli*, but more complicated for other bacteria (see Chapter 4). *E. coli* cells that have been treated with a cold solution of calcium chloride are able to take up plasmid DNA (transformation), or phage DNA (transfection), with a typical efficiency of 10^6 transformants per μg of purified plasmid or phage DNA (Mandel & Higa, 1970). For vectors based on phage λ, these frequencies have been improved by the development of methods for the *in vitro* packaging of phage particles (Hohn & Murray, 1977). The phage particles are much more infective and the relative yield of clones can be increased by one or two orders of magnitude.

3.6 Selection of clones

The final stage in a cloning experiment is the selection of the clone harbouring the desired hybrid molecule. If the cloned fragment produces a product for which there is a specific plate assay, or it can cause a change in the hosts phenotype (i.e. complement an auxotrophic marker), then this can be used in the selection (Chang *et al.*, 1978). When a polypeptide is produced, which has previously been purified, it is possible to raise antibodies to the purified material and detect the desired clone by immunological methods (Sanzey *et al.*, 1976). If a DNA or RNA probe homologous to the desired insert is available, then this can be radiolabelled and the clone detected by colony or plaque hybridisation (Grunstein & Hogness, 1975).

The different strategies and routes described here for *in vitro* cloning are summarised in Fig. 3.5. A more detailed description will be found in Timmis (1981) and Old and Primrose (1980).

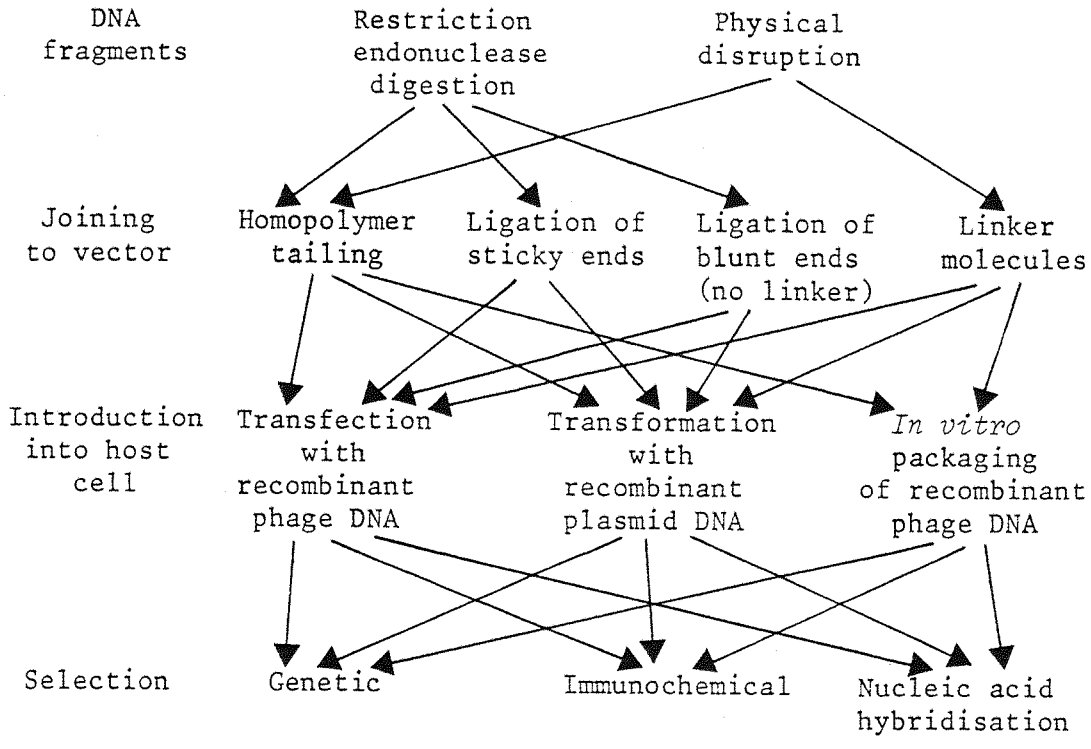


Fig. 3.5. Generalised scheme for DNA cloning. Favoured routes are shown by arrows.

3.7 Isolation of initial clones

The initial approach was to try and repeat the experiment of Courvalin *et al.* (1977). Total genomic DNA was isolated from 40 ml of late exponential phase *B. circulans* grown in Difco nutrient broth supplemented with 1% glucose (Table 2.6) and pBR322 was isolated from CC1100 (Table 2.2).

For the cloning experiments 1 µg of pBR322 and 2 µg of *B. circulans* DNA were digested with Eco RI, mixed and ethanol precipitated to stop the reaction. The pellet was resuspended in 100 µl of T_4 DNA ligase buffer (to give a vector concentration of 10 µg/ml) and 0.1 units of T_4 DNA ligase was added. This was incubated at 4°C for 18 h, diluted with 300 µl of TE (10 mM Tris 1 mM EDTA pH 8.0) and used to transform 400 µl of transformation competent *E. coli* HB101 (Table 2.3). After the heat shock, 5 ml of LB broth was added and the cells were incubated at 37°C for 1 h to allow expression of the transformed DNA. The cells were then harvested by centrifugation, resuspended in 1 ml of LB broth and spread on 10 LB agar plates supplemented with 20 µg/ml tetracycline and 20 µg/ml neomycin. After incubation at 37°C for 48 h, no colonies could be detected on any of the plates. This experiment was repeated twice but no clones were obtained. However, selection for Tet^r alone gave almost confluent growth, indicating that the ligation and transformation had been successful.

This caused a reappraisal of the experimental procedure. The most obvious anomaly was that in attempting to clone the *APH* gene from *B. circulans* we were selecting with 20 µg/ml neomycin, an antibiotic to which *B. circulans* has no significant resistance. A more rational approach would be to select with an aminoglycoside to which *B. circulans* is resistant. The logical choice was butirosin but at this time butirosin was not available, so it was decided to use ribostamycin at a concentration of 25 µg/ml. The cloning experiment was repeated as before, except that the transformation mixtures were spread onto 10 LB agar plates supplemented with 20 µg/ml tetracycline and 25 µg/ml ribostamycin. After 18 hours' incubation at 37°C, 11 colonies had grown. These were picked off, purified and tested for resistance to ampicillin (50 µg/ml), tetracycline (20 µg/ml), ribostamycin (25 µg/ml) and neomycin (20 µg/ml). All

Fig. 3.6. Restriction maps of pCH1, pCH2 and pCH3.

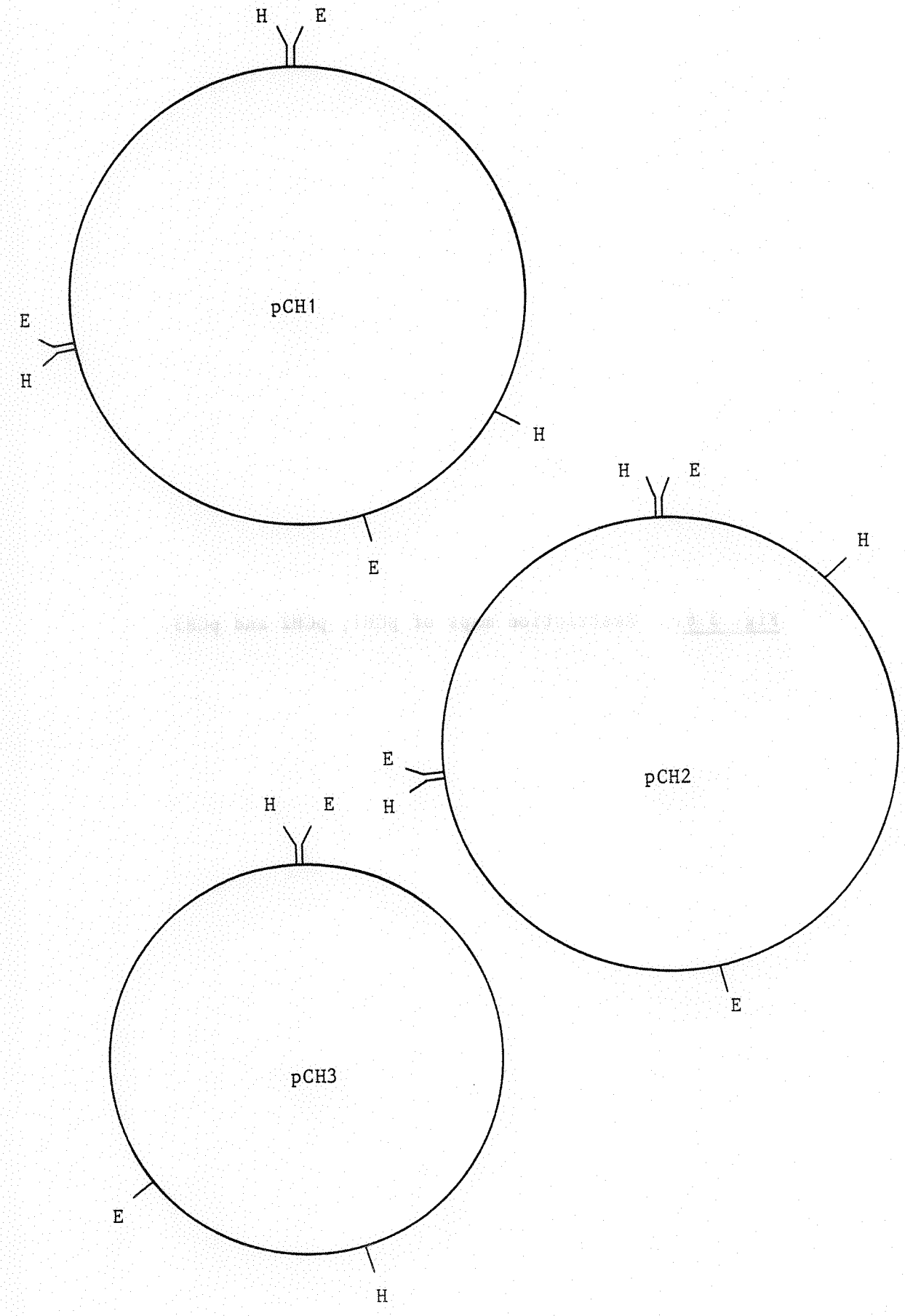


Fig. 3.6. Restriction maps of pCH1, pCH2 and pCH3 for Eco RI (E) and Hind III (H).

11 clones had the same phenotype being resistant to ampicillin, tetracycline and ribostamycin, but sensitive to neomycin. Agarose gel electrophoresis of rapid plasmid preparations from these strains (Table 2.4) showed that all 11 now harboured a plasmid; 2 harboured plasmids of ~16 kb and the other 9 harboured plasmids of 12 kb. Both of the 16 kb plasmids and 2 of the 12 kb plasmids were digested with Eco RI. Electrophoresis of the products revealed the same pattern for each plasmid: a band at 4.3 kb corresponding to pBR322, and a band at 7.5 kb. This can be explained if the larger plasmids contain 2 molecules of pBR322; digestion of these plasmids with Hind III indicated that this was the case. Both plasmids (pCH1 isolated from CC1101 and pCH2 isolated from CC1102) contained a direct repeat of pBR322, but the 7.5 kb insert was present in different orientations (Fig. 3.6). One strain (CC1103) harbouring one of the smaller plasmids with a single copy of pBR322 (pCH3) was also kept for further study.

Purified pCH1 was used to transform *E. coli* HB101 and transformants were selected for Tet^r only. > 500 of these transformants were replica plated onto LB agar containing 50 μg/ml ampicillin, LB agar containing 20 μg/ml tetracycline and LB agar containing 25 μg/ml ribostamycin. When the replicas were inspected > 98% of the original transformants had non-selectively become Amp^r and Rib^r; this demonstrates that ribostamycin, ampicillin and tetracycline resistance are now linked on pCH1 and that ribostamycin resistance is encoded in the 7.5 kb insert.

3.8 Subcloning the *APH* gene from pCH1

The phosphotransferase from *S. circulans* is reported to have a monomer M.W. of 28,500 daltons (Davies & Smith, 1978). This would require a coding region of some 800 bases and a total gene length of approximately 1,200 bases. The 7.5 kb Eco RI fragment was too large for sequencing experiments on a gene of this size, so the gene had to be subcloned and re-isolated on a smaller fragment. The isolation of pCH1 containing a direct repeat of pBR322 offered a simple way of doing this. When pCH1 is digested with an enzyme that has a single target site in pBR322, one of the fragments produced will be linearised pBR322.

If this enzyme also cuts within the insert, then its digest of pCH1 produces a mixture which can be used in ligations when attempting to subclone the *APH* gene. Restriction enzyme digestion of pCH3 showed that Hind III and Bam HI cut the insert once, while Sal I cut the insert three times. Hind III and Sal I were chosen as suitable enzymes to use in attempting to subclone the *APH* gene.

The plasmid pCH1 was purified from CC1101 (Table 2.2), 1 µg aliquots were digested with either Hind III or Sal I and the reactions were stopped by ethanol precipitating the DNA. The precipitates were resuspended in 100 µl of T_4 DNA ligase buffer and 0.1 units of T_4 DNA ligase was added. After incubation at 4°C overnight, the ligation mixtures were transformed into 100 µl each of transformation competent *E. coli* HB101 and spread on 5 LB agar plates containing 50 µg/ml ampicillin and 25 µg/ml ribostamycin. When the plates were examined after incubation for 18 h at 37°C, 101 colonies had grown on the plates from the experiment with Sal I and ~ 1,600 colonies had grown on the plates from the experiment with Hind III. From each experiment, 16 colonies were purified and tested for resistance to ribostamycin, ampicillin and tetracycline. In the Sal I experiment, 13 clones were Amp^r, Rib^r, Tet^s and 3 were Amp^r, Rib^r, Tet^r. In the Hind III experiment, 6 clones were Amp^r, Rib^r, Tet^s and 10 were Amp^r, Rib^r and Tet^r. As both the Sal I and Hind III sites in pBR322 lie within the tetracycline resistance gene, it would be expected that the subclones would be Amp^r, Rib^r, Tet^s; the appearance of Ampr, Rib^r, Tet^r clones could be due to some uncleaved pCH1 in the transformation mixture, or to the construction of recombinant plasmids that recreate the tetracycline resistance gene.

Agarose gel electrophoresis of rapid plasmid preparations from each of the Amp^r, Rib^r, Tet^s clones revealed the presence of plasmids larger than pBR322 but smaller than pCH1. The clone harbouring the smallest plasmid produced in each experiment was maintained for further study; from the Hind III experiment this was CC1104 harbouring pCH4, and from the Sal I experiment this was CC1105 harbouring pCH5. Restriction endonuclease digestion and agarose gel electrophoresis of pCH1, pCH3, pCH4 and pCH5 showed that they all carried an insert derived from the initial 7.5 kb Eco RI fragment (Fig. 3.7). The Sal I

insert in pCH5 is 2.7 kb and was a suitable size to use in further experiments. By comparing single and double restriction enzyme digests of pCH5, it was possible to deduce a restriction map of the plasmid and hence the insert (Figs. 3.8 and 3.9).

3.9 Assay of phosphotransferase activity in the clones

To confirm that the ribostamycin resistance was due to a phosphotransferase, the enzyme levels in crude extracts were measured using ribostamycin as the phosphate acceptor (Section 2.12). The ribostamycin resistant strains CC1101, CC1104 and CC1105 had approximately twice the phosphotransferase activity of the parent, *B. circulans*, whereas CC1100 harbouring the vector alone showed no detectable phosphotransferase activity (Table 3.1). This showed that the *APH* gene of *B. circulans* had been cloned and it was expressed in *E. coli*.

Table 3.1. Phosphotransferase activity in crude cell extracts.

See Methods section 2.12.

Strain	Protein Concentration (mg/ml)	Activity (dpm transferred/ 10 µl extract)	Relative Activity (dpm transferred/ µg protein)
B. circulans	2.0	2,740	137
CC1101	6.7	19,670	296
CC1104	6.1	17,850	293
CC1105	5.8	15,560	268
CC1100	4.5	230	5

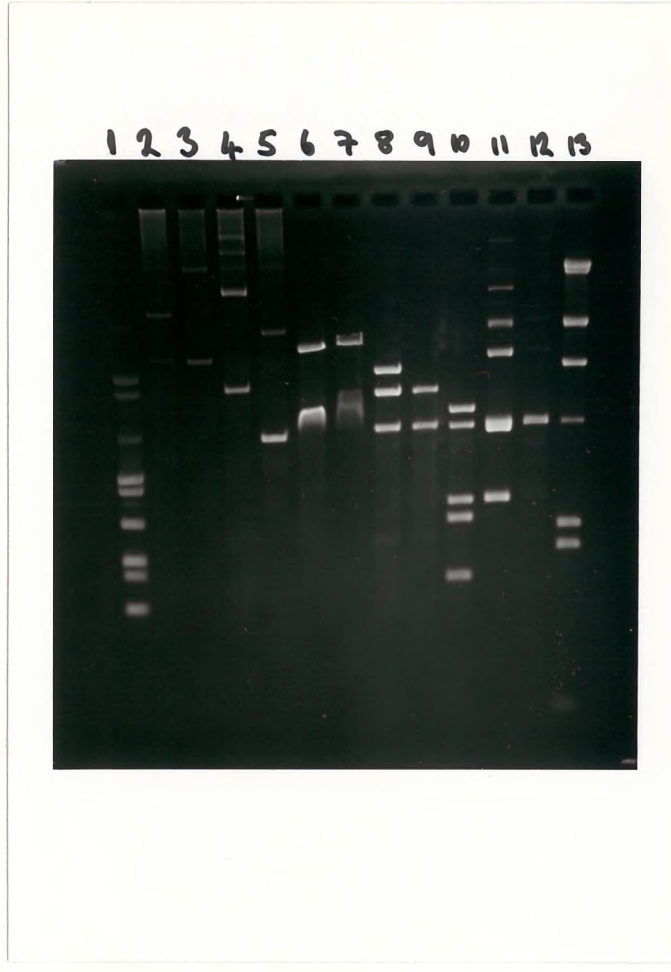


Fig. 3.7. Comparison of *APH* gene clones.

Track: 1) Plasmid standards purified from *E. coli* V517; 2) pCH1; 3) pCH3; 4) pCH4; 5) pCH5; 6) Eco RI digest of pCH1; 7) Eco RI digest of pCH3; 8) Hind III digest of pCH1; 9) Hind III digest of pCH4; 10) Sal I digest of pCH1; 11) Sal I digest (partial) of pCH5; 12) Sal I digest of pBR322; 13) Hind III digest of λ DNA.

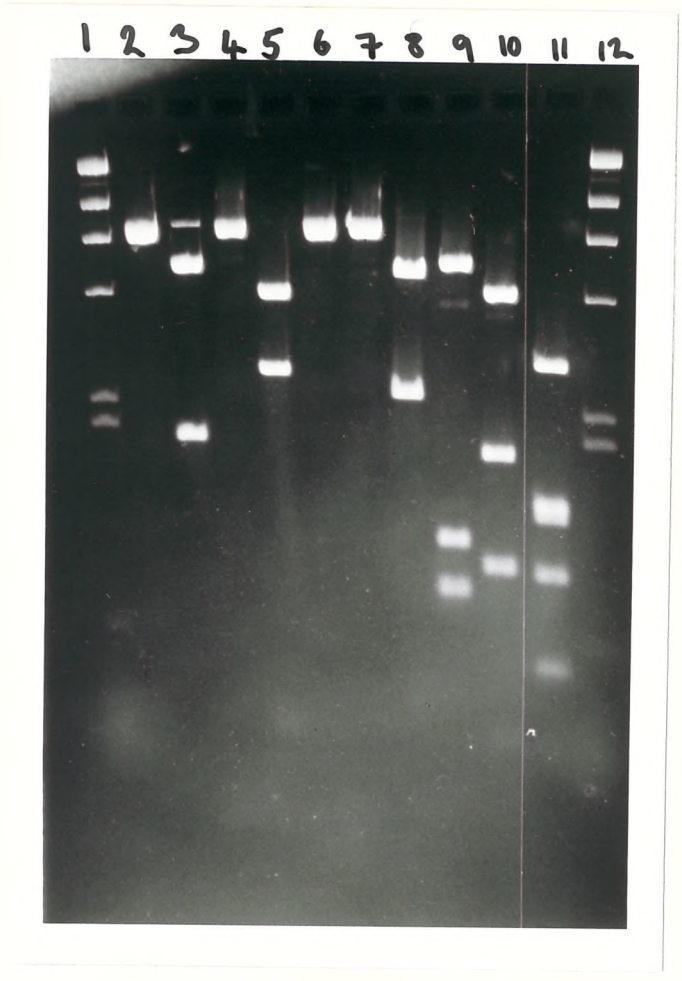


Fig. 3.8. Digests of pCH5.

Tracks 1) + 12) Hind III digest of λ DNA; all other tracks are digests of pCH5; 2) Eco RI; 3) Pst I; 4) Bam HI; 5) Sal I; 6) Bgl II; 7) Hind III; 8) Hind III + Bgl II; 9) Hind III + Pst I; 10) Sal I + Bgl II; 11) Sal I + Bgl II + Pst I.

This gel shows that the insert contains single Bgl II and Pst I sites (tracks 3 and 6) and that the Pst I site is closest to the Hind III site (tracks 8 and 9). The position of the sites and sizes of the fragments were confirmed by the Sal I + Bgl II and Sal I + Bgl II + Pst I digests (tracks 10 and 11).

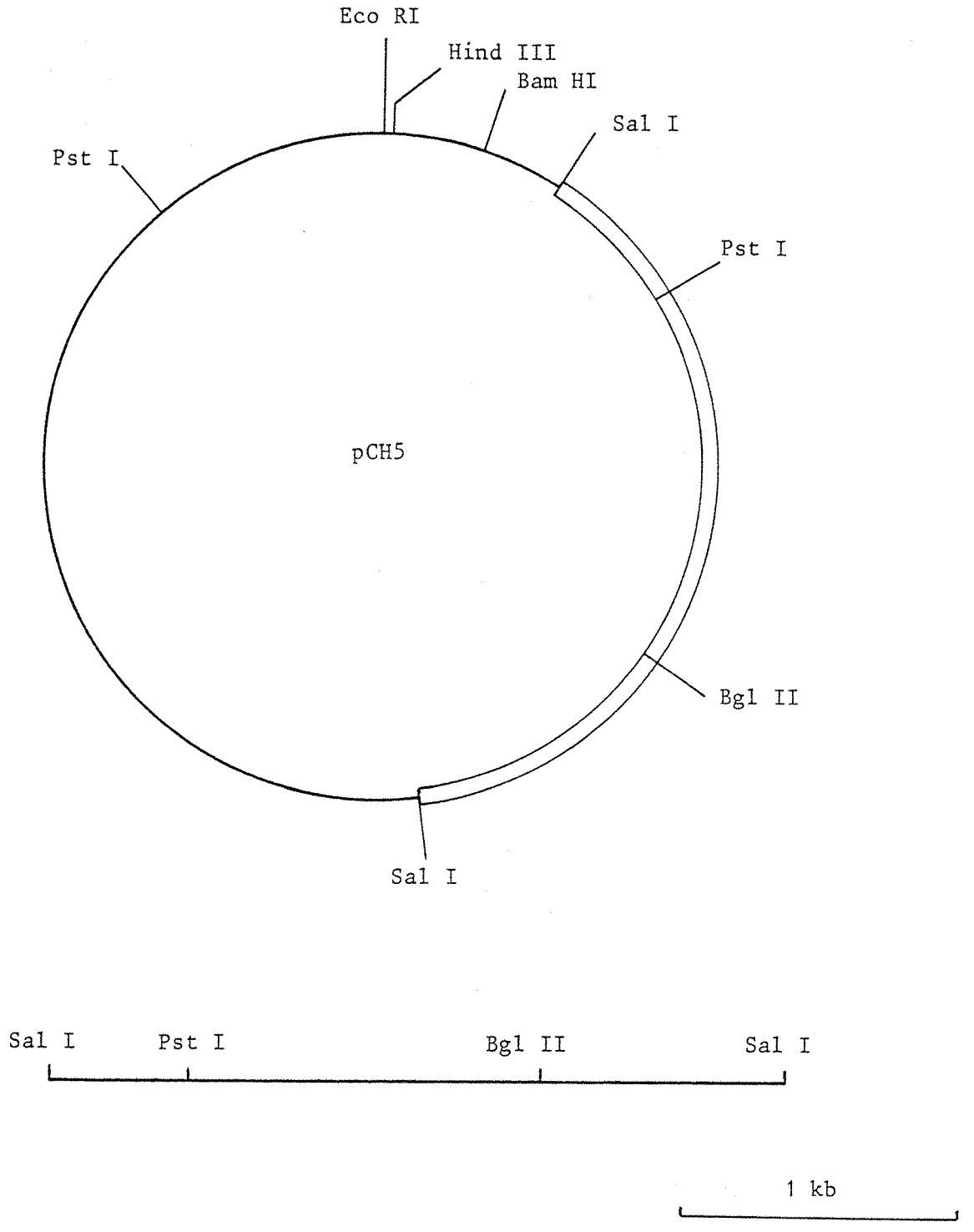


Fig. 3.9. Restriction map of pCH5 and the 2.7 kb Sal I insert.

3.10 Summary

The *APH* gene of *B. circulans* has been cloned as a 2.7 kb Sal I fragment that is a suitable starting point for other experiments. It has also been demonstrated that the *APH* gene does not have a target site for the restriction endonucleases Eco RI, Hind III or Sal I. One interesting difference between the clones isolated here and those isolated by Courvalin *et al.* (1977) is that CC1102, CC1103, CC1104 and CC1105, like *B. circulans*, are not resistant to neomycin, whereas those of Courvalin are. All other information, such as the size of the Eco RI fragment and the presence of a single Bam HI site within it, indicate that equivalent fragments have been cloned in each experiment. The difference might be caused by altered levels of expression from the different vectors and *E. coli* hosts used in the two studies.

CHAPTER 4

CLONING OF THE *B. circulans APH* GENE IN *S. lividans*

Introduction

4.1 In Chapter 3, a brief description of the various techniques available for *in vitro* gene cloning in *E. coli* was presented. Considering the relatively advanced level of this technology, it is perhaps necessary to explain why cloning systems need to be developed for other organisms. The most obvious reason is that genes from other organisms may not be expressed when transferred to *E. coli*. More importantly, other organisms (i.e. *Bacilli* and *Streptomyces*) possess interesting features, such as a developmental cycle and the production of secondary metabolites (Hopwood, 1981), which can be successfully studied only in hosts that are able to provide the appropriate internal environment (i.e. different RNA polymerase σ factors (Losick & Pero, 1981)). Finally, *E. coli* is not of great importance in the fermentation industry, so if *in vitro* cloning technology is to be used in strain improvement programmes, it is necessary to develop technologies appropriate to industrial strains. Five requirements for an *in vitro* cloning system were listed in Chapter 3, only two of which need to be changed when developing a system for a new organism : (i) the vectors used, and (ii) the method of introducing the hybrid molecules into the host cells.

4.2 Development of transformation in *Streptomyces*

Most of the methods used to introduce DNA into cells are based on transformation. The first observation of bacterial transformation by purified DNA was made by Avery (Avery *et al.*, 1944), whose experiments showed that growing cells of *D. pneumoniae* were able to take up DNA and acquire new genetic capacity. Ultimately, this work led to the recognition that DNA is the carrier of genetic information. Only a few strains are able to take up DNA in this way; more often the cells have to be treated to make them 'permeable' or competent for transformation (e.g. the calcium chloride treatment of *E. coli*).

Konvalinkova *et al.* (1977) found that early log phase cultures of *S. virginae* were competent for DNA uptake and were able to transfect the mycelium with purified actinophage DNA. However, this was not developed into a generalised system for transformation. The use of protoplasts in genetic studies has received much attention and protoplast fusion has now been demonstrated in a wide variety of cells : fungal (Ferenczy *et al.*, 1974), bacterial (Fodor & Alfoldi, 1976; Schaeffer *et al.*, 1976) and eukaryotic cells (Pontecorvo *et al.*, 1977). In 1974, Okanishi showed that some *Streptomyces* form stable protoplasts when treated with lysozyme, and on suitable agars these could regenerate into mycelial forms (Okanishi *et al.*, 1974). This was developed into a generalised system for genetic recombination in *Streptomyces*, when Hopwood found that polyethylene glycol (PEG) could stimulate fusion of *Streptomyces* protoplasts (Hopwood *et al.*, 1977). The versatility of protoplast systems was further demonstrated by Moritia, who found that PEG increased the efficiency of transfection of *E. coli* spheroplasts by bacteriophage DNA (Morita *et al.*, 1977). Again, it was possible to apply this information to the *Streptomyces* system and develop methods for PEG-assisted transformation (Bibb *et al.*, 1978) and transfection (Suarez & Chater, 1980(a)). The precise mechanism of these processes is not understood but it is possible that the PEG induces the protoplasts to aggregate and fuse. The DNA may become trapped in the aggregates and be incorporated into the cells when the fusion occurs.

4.3 Development of cloning vectors for *Streptomyces*

The first vectors used for *in vitro* cloning in *Streptomyces* were natural plasmids isolated from *S. coelicolor* (SCP2) and *S. lividans* (SLP1.2) (Bibb *et al.*, 1980; Thompson *et al.*, 1980). One of the advantages of transformation in *Streptomyces* is that under appropriate conditions transformants can be detected by visual inspection. When a derivative of a strain harbouring a conjugative plasmid is grown in contact with a strain lacking the plasmid on a suitable agar, a visible reaction is produced at the interface. This appears as a region of

retarded growth (Fig. 4.1). The inhibition reaction has been called 'lethal zygosis' (Bibb *et al.*, 1977) but this may be misleading as there is no evidence that cell death is involved. When protoplasts are regenerated to produce confluent lawns after a transformation experiment, the cells containing the plasmid (i.e. the transformants) are growing in a background of plasmid-free cells, and so are surrounded by a ring of retarded growth and appear as 'pocks' (Fig. 4.2) (Chater *et al.*, 1982).

Fig. 4.1. M180 growing in a lawn of *S. lividans* 66 showing a 'lethal zygosis' reaction at the interface.

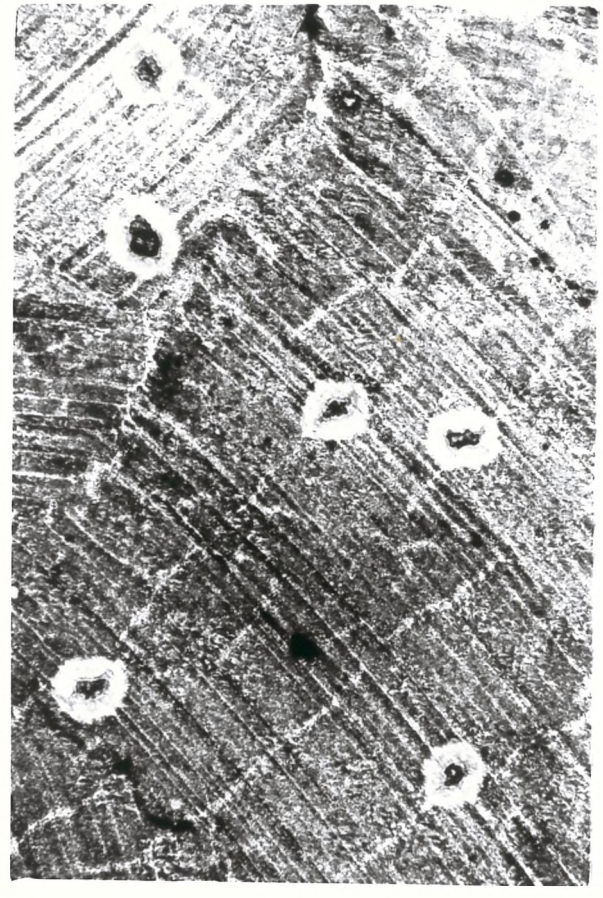


Fig. 4.2. SLP1.2 transformed into *S. lividans* 66 protoplasts which were allowed to regenerate on R2YE agar as lawns. Transformants appear as 'pocks'.

Although these natural plasmids were successfully used in cloning experiments, their use was restricted by a narrow host-range. More recently, a broad host-range plasmid pIJ101 has been isolated (Kieser *et al.*, 1982). This is an important advance as it is often not convenient to perform cloning experiments in *S. lividans* or its close relatives. For instance, when attempting to clone genes involved in antibiotic biosynthesis, it may be necessary to do the experiments in derivatives of the producing organism, so that the clones can be selected by complementation. The availability of cloned antibiotic resistance determinants (Thompson *et al.*, 1980) has allowed the development of the natural plasmids into specialised cloning vectors, with markers for positive selection and insertional inactivation (Chater *et al.*, 1982; Kieser *et al.*, 1982).

All of the fundamental requirements for *in vitro* gene cloning in *Streptomyces* have now been fulfilled and many types of cloning experiment are now possible. Also, specialised cloning vectors, such as promotor cloning plasmids have been developed (Bibb & Cohen, 1982). One appealing class of specialised vectors are shuttle vectors. These are vectors that are able to replicate in more than one host and allow exploitation of the different advantages of other host/vector systems. For example, when a gene is cloned in yeast, it is often more convenient to isolate large amounts of the cloned DNA for physical analysis by growing it in *E. coli* (Faye & Simon, 1983). Several vectors that can replicate in both *E. coli* and *Streptomyces* have been described (Suarez & Chater, 1980(b); Schottel *et al.*, 1981; Kieser *et al.*, 1982). One of the requirements for a successful bifunctional vector is a method of selecting transformants in either host. The *APH* gene of *S. fradiae* has been engineered and expressed in *E. coli* (Rodgers *et al.*, 1982). The *APH* gene of *B. circulans* is expressed in *E. coli*; if it is also expressed in *Streptomyces*, it could be of considerable use in the construction of bifunctional vectors.

4.4 Transfer of the *B. circulans APH* gene to *S. lividans* 66

The vector used for cloning in *S. lividans* 66 was SLP1.2, which was purified from M180 (Table 2.7). The intention was to

ligate the 2.7 kb Sal I insert from pCH5 into SLP1.2, transform the mixture into *S. lividans* 66 and see if ribostamycin resistant transformants could be detected. One difficulty with this approach was that SLP1.2 contains 3 Sal I sites (Fig. 4.3), only 1 of which is known to be in a non-essential region. This was overcome by using vector that had been partially digested with Sal I. For the cloning experiment, 0.5 µg of SLP1.2 was partially digested with Sal I (to give 1-2 cuts per molecule) and mixed with 1 µg of Sal I digested pCH5. The DNA was ethanol precipitated and the pellet resuspended in 30 µl of T_4 DNA ligase buffer. T_4 DNA ligase (0.1 units) was added and the mixture was incubated overnight at 4°C. The ligation mixture was used to transform protoplasts of *S. lividans* 66. These were spread on 4 R2YE plates and allowed to regenerate for 10 days at 28°C. This gave a total of approximately 1,500 'pocks'. The regeneration plates were replica-plated onto minimal agar supplemented with 5 µg/ml ribostamycin and incubated at 30°C. After 3 days' incubation, 51 colonies had grown; 2 of these (CL1023 and CL1024) were picked off and purified for further study.

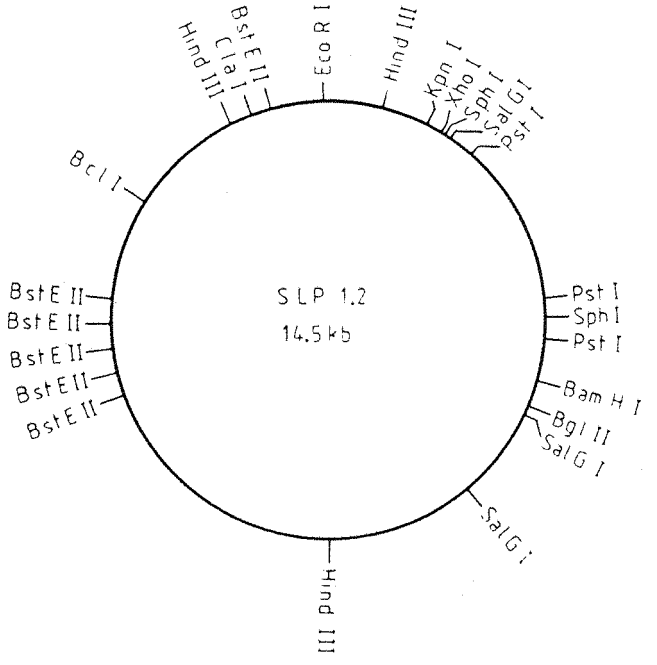


Fig. 4.3. Restriction map of SLP1.2 (Chater *et al.*, 1982).

CL1023 and CL1024 were patched out onto a R2YE agar plate with M180 and *S. lividans* 66 as controls. After 5 days' incubation at 30°C, this was replica-plated onto minimal agar supplemented with 5 μg/ml ribostamycin and R2YE agar seeded with enough spores of *S. lividans* 66 to produce a confluent lawn. These plates were incubated at 30°C for 7 days and then inspected. M180, CL1023 and CL1024 all produced a 'lethal zygosis' reaction in the lawn of *S. lividans* 66 but only CL1023 and CL1024 were ribostamycin resistant (Fig. 4.4). This indicated that CL1023 and CL1024 had acquired a plasmid and become ribostamycin resistant. The presence of plasmids in these strains was confirmed by agarose gel electrophoresis of small scale plasmid preparations. Both CL1023 and CL1024 harboured a plasmid which was larger than SLP1.2.

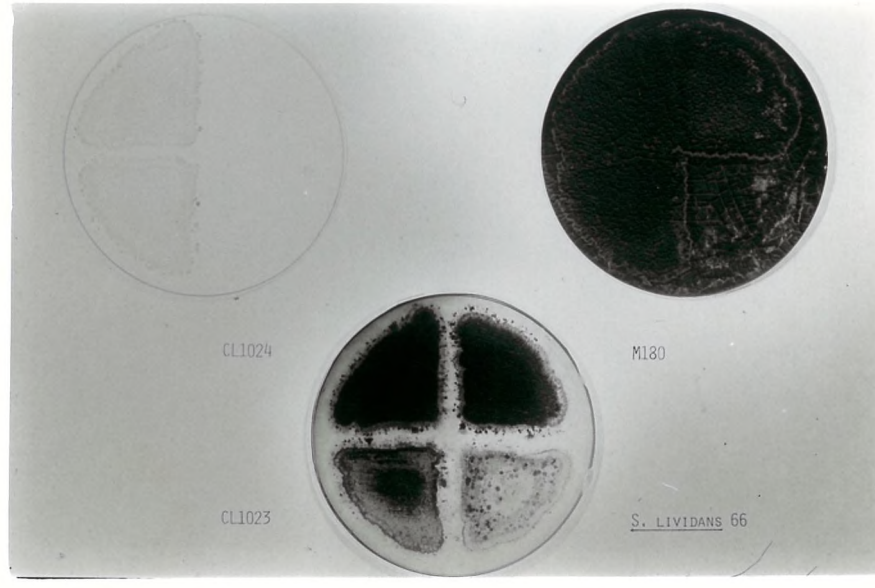


Fig. 4.4. *S. lividans* 66, M180, CL1023 and CL1024 replica plated onto lawns of *S. lividans* 66 and agar supplemented with ribostamycin. M180, CL1023 and CL1024 harbour a plasmid but only CL1023 and CL1024 are ribostamycin resistant.

To demonstrate that the ribostamycin resistance is encoded on the plasmids, one of them (pCL1) was purified from CL1023 (Table 2.7). The purified pCL1 was used to transform *S. lividans* 66 protoplasts. When these were allowed to regenerate, 20-30 'pocks' were visible on each plate. Five of these plates were replica plated onto minimal agar supplemented with 5 µg/ml ribostamycin. After the replicas had grown, they were compared with the master plates; each pock on the master plate had given rise to a ribostamycin resistant colony (Fig. 4.5). In control experiments, purified pCL1 was spread onto R2YE agar plates but no colonies grew, indicating that the DNA was not contaminated with cells. This experiment showed that the ribostamycin resistance was encoded on pCL1.

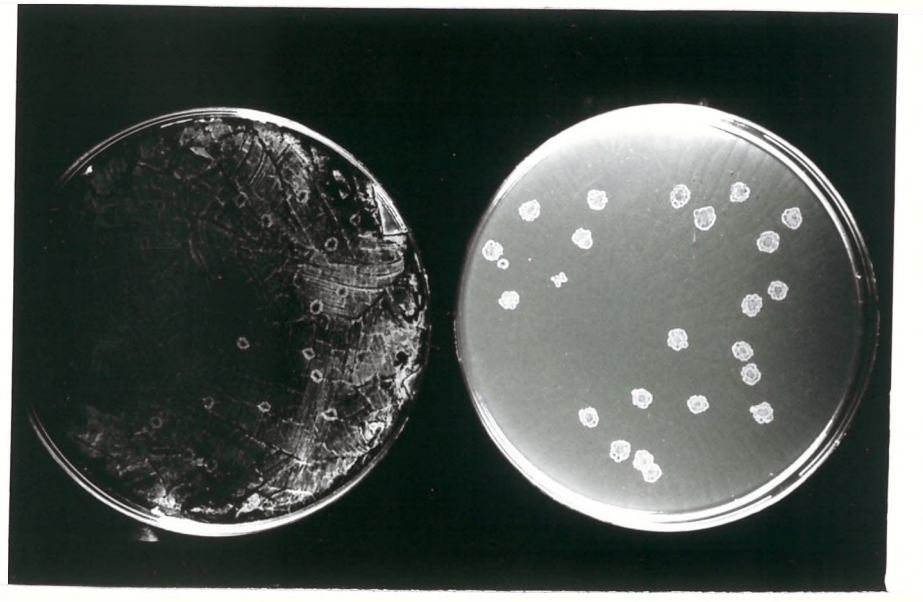


Fig. 4.5. pCL1 transformed into *S. lividans* 66 protoplasts and replica plated onto agar supplemented with ribostamycin after regeneration. All transformants were ribostamycin resistant.

The purified pCL1 was digested with Sal I, Pst I and Bgl II. Agarose gel electrophoresis of the products showed that all of SLP1.2 was included in pCL1 and allowed a restriction map of pCL1 to be deduced (Fig. 4.6).

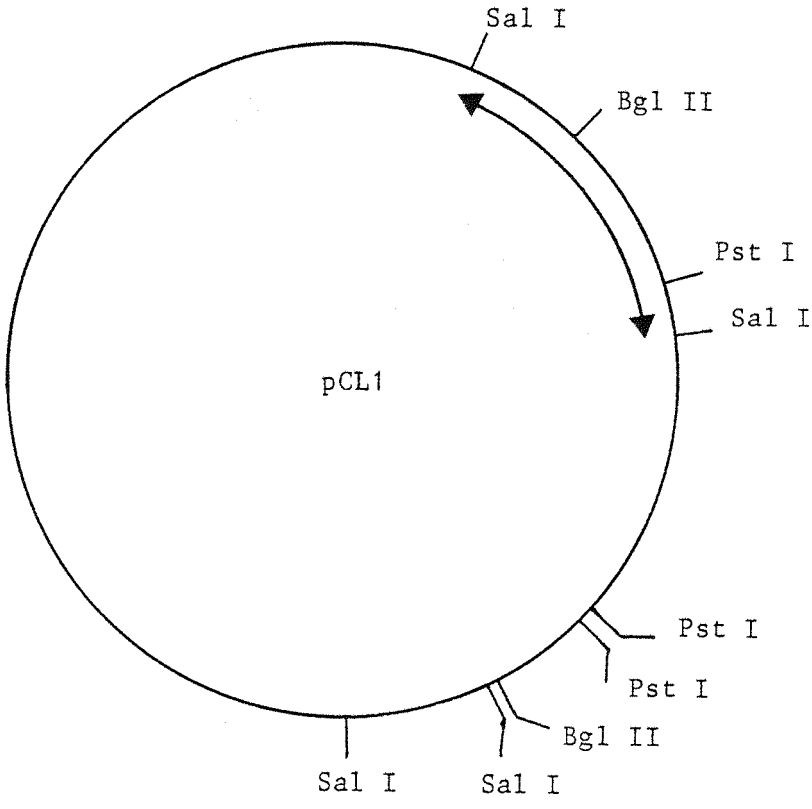


Fig. 4.6. Restriction map of pCL1. The 2.7 kb insert from pCH5 is marked by a double-headed arrow.

4.5 Assay of the phosphotransferase from *S. lividans* derivatives

To confirm that the ribostamycin resistance in CL1023 and CL1024 was due to a phosphotransferase, the levels of the enzyme were measured in crude extracts, using butirosin as the phosphate acceptor (Section 2.12). M180 and *B. circulans* were used as controls (Table 4.1).

Table 4.1. Phosphotransferase activity in *S. lividans* derivatives

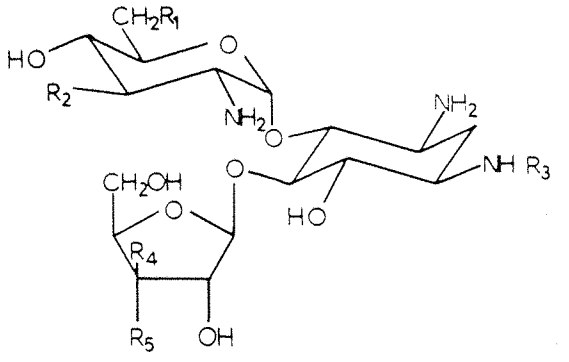
See Methods section 2.12.

Strain	Protein concentration (mg/ml)	Activity (dpm transferred/ 10 µl extract)	Relative activity (dpm transferred/ µg protein)
B. circulans	3.2	11,440	358
M180	2.8	890	32
CL1023	3.4	10,840	319
CL1024	3.2	11,230	351

4.6 Hybridisation of the *APH* gene to DNA from other aminoglycoside producing organisms

The suggestion that antibiotic producing organisms are the original source of clinically significant antibiotic resistance (Benveniste & Davies, 1973) has been tested in two ways : by immunological methods (Davies & Smith, 1978) and by nucleic acid hybridisation (Courvalin *et al.*, 1978). Neither of these methods has detected homology between different classes of aminoglycoside phosphotransferases. However, the hybridisation techniques used relied on the detection of heteroduplex molecules under the electron microscope. A more sensitive method would be to use the 'Southern Blot' technique (Southern, 1975). This involves the denaturing of DNA in an agarose gel and its transfer to a nitrocellulose filter. Denatured probe DNA, that has been labelled with [^{32}P] by nick-translation, is added and the conditions are changed to favour hybridisation. The level of homology required for hybridisation can be reduced by increasing the salt concentration in the wash. Finally, the filters are dried and used to expose a sheet of X-ray film. If sequence specific hybridisation has occurred, it will be detected as a dark band on the autoradiogram.

It was decided to use the Southern Blot technique to see if homology could be detected between the cloned *APH* gene of *B. circulans* and other strains that produce members of the neomycin family of aminoglycoside antibiotics (Pearce & Rinehart, 1981) (Fig. 4.7).



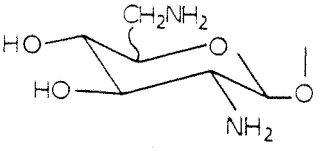
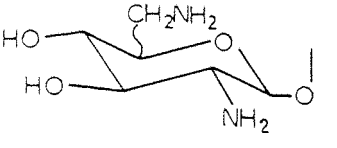
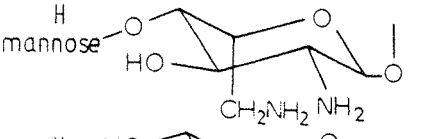
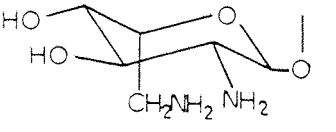
	R_1	R_2	R_3	R_4	R_5
NEOMYCIN	NH_2	OH	H	H	
PAROMOMYCIN	OH	OH	H	H	
RIBOSTOMYCIN	NH_2	OH	H	H	OH
BUTIROSIN A	NH_2	OH	$CO.CHOH.CH_2CH_2NH_2$	OH	H
B	NH_2	OH	$CO.CHOH.CH_2CH_2NH_2$	H	OH
LIVIDOMYCIN A	OH	H	H	H	
B	OH	H	H	H	

Fig. 4.7. Structures of aminoglycosides produced by the strains used in the blotting experiment (Pearce & Rinehart, 1981).

Total DNA was isolated from *B. circulans*, *S. fradiae* ATCC10745, *S. ribosdificus* ATCC21294, *S. rimosus* forma *paromomycinus* ATCC14827 and *S. lividus* ATCC21178. The genomic DNA was digested with EcoRI and two samples of each were run on an agarose gel; EcoRI digested pCH3 and Sal I digested pCL1 were also included on the gel. The DNA was transferred to a nitrocellulose filter, which was subsequently cut into strips. The 2.7 kb Sal I insert from pCH5 was labelled by nick-translation with [P^{32}]dATP and used as a probe for hybridisation. For the wash, when hybridisation was complete, the duplicate digests were treated differently: one set was given a stringent wash (0.1x standard buffer), so that only hybridisation between closely homologous sequences would survive; the other set was given a non-stringent wash (2x standard buffer) so that hybridisation between less homologous sequences would survive. After this, the nitrocellulose filter was re-assembled and used to expose a sheet of X-ray film. A comparison of the original gel and the autoradiograph (Fig. 4.8) showed that the 2.7 kb SalI fragment hybridises to the 7.5 kb EcoRI fragment of pCH3, a 7.5 kb EcoRI fragment of *B. circulans* DNA and the 2.7 kb Sal I fragment of pCL1. This shows that all the cloned fragments are derived from a section of *B. circulans* DNA. However, no hybridisation to DNA from any of the other aminoglycoside producing strains was detected; even when the X-ray film was pre-fogged and exposed for 7 days at -70°C,using an intensifying screen, no extra bands appeared. This result was surprising, as considerable homology might be expected between these strains. One of the difficulties of the Southern Blot technique is that it is not possible to use it to quantify homology. A short sequence of complete homology might be detected, whereas a much longer sequence with 50% homology might not be detected. The only completely reliable way in which to detect homology is to compare sequences.

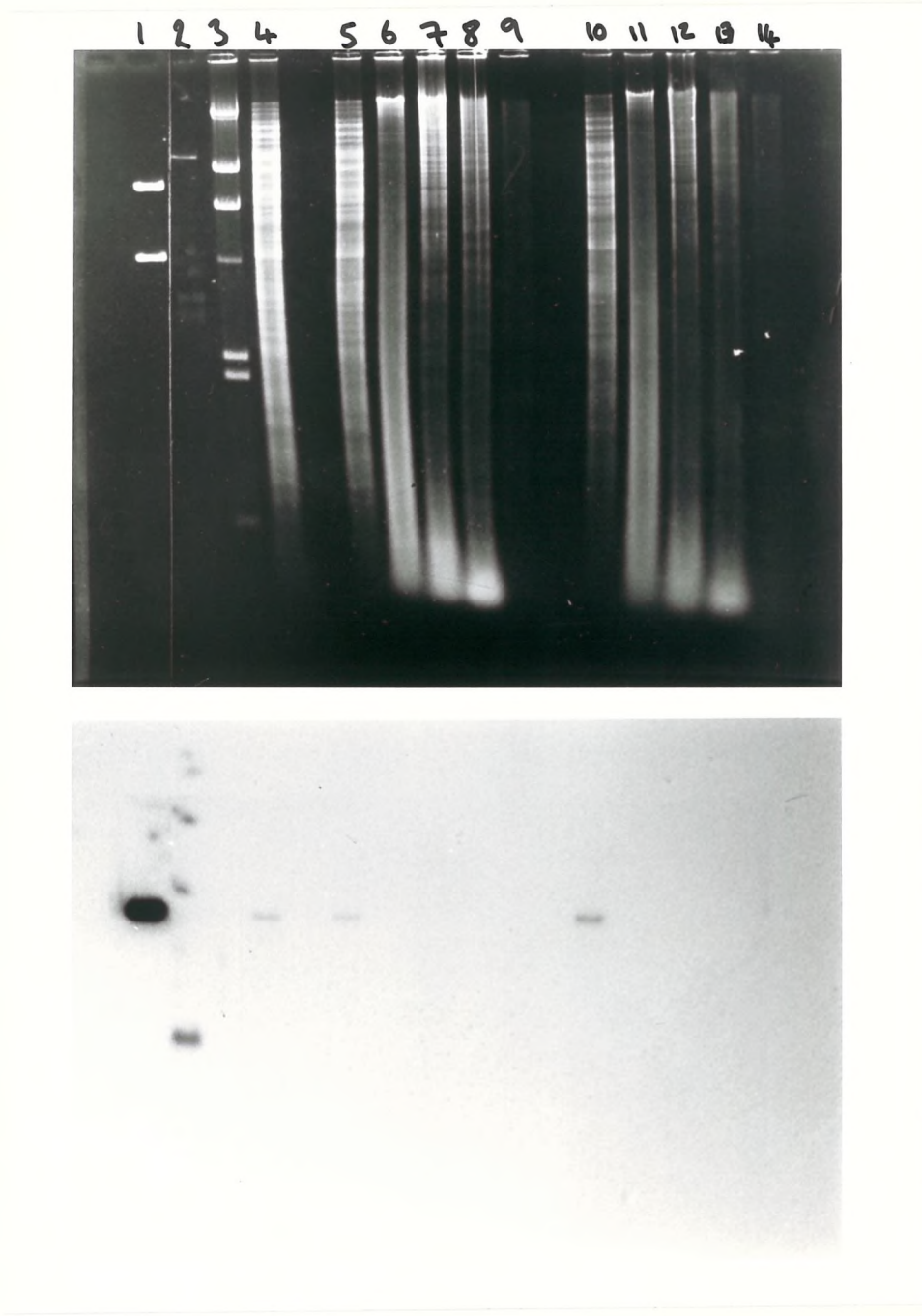


Fig. 4.8. Results of the Southern Blot experiment.

Track: 1) Eco RI digested pCH3; 2) Sal I digested pCL1; 3) Hind III digested λ DNA; 4), 5) and 10) Eco RI digested *B. circulans* DNA; 6) and 11) Eco RI digested *S. fradiae* DNA; 7) and 12) Eco RI digested *S. ribosdificus* DNA; 8) and 13) Eco RI digested *S. rimosus* forma *paromomycinus* DNA; 9) and 14) Eco RI digested *S. lividus* DNA. Hybridisation was detected only in tracks 1, 2, 4, 5 and 10. Tracks 10-14 had a non-stringent wash.

4.7 Summary

The *APH* gene of *B. circulans* has been cloned in *S. lividans* 66, using the low copy number plasmid SLP1.2. The presence of the gene can be detected by the antibiotic resistance phenotype and the phosphotransferase activity in crude cell extracts is approximately the same as in *B. circulans*. When the cloned *APH* gene is used as a probe in Southern Blot experiments, no hybridisation to DNA from other aminoglycoside producing organisms could be detected. The Southern Blot experiments were performed with the assistance of S. Talwar of the Department of Biology, University of Southampton.

CHAPTER 5

LOCALISATION OF THE *B. circulans APH* GENE AND CHARACTERISATION OF THE PRODUCT

5.1 Introduction

Mechanism of action of aminoglycoside antibiotics

The aminoglycoside antibiotics all have effects on protein biosynthesis. When mutagenised cultures are used to select antibiotic resistant mutants, the most common class of mutants found have altered ribosomal proteins. This presumably affects the binding of the antibiotic to the ribosome (Franklin & Snow, 1981). The exact molecular action of aminoglycosides on the ribosome is not known. However, members of the 2-deoxystreptamine series of aminoglycosides have been shown to have three separate concentration dependent effects on isolated ribosomes : (i) at low concentrations < 3 μg/ml there is a strong inhibition of total protein synthesis; (ii) at 5-50 μg/ml, misreading of the mRNA template becomes evident, especially reading through termination signals, so that abnormally long polypeptides accumulate; and (iii) at higher concentrations the inhibition of protein synthesis is re-established (Zierhut *et al.*, 1979). This is taken as an indication that there are several antibiotic binding sites; also, competition experiments show that different aminoglycosides bind at different sites on the ribosome.

Aminoglycoside phosphotransferases

It is well documented that the presence of an aminoglycoside phosphotransferase is sufficient to confer resistance on an host cell but the mechanism which mediates this resistance is unclear. The phosphotransferase from *B. circulans* is classed as an APH(3') enzyme and it modifies the 3'-hydroxyl of the amino-hexose I of the 2-deoxystreptamine antibiotics (Davies & Smith, 1978) (Fig. 5.1). One interesting observation is that when resistant cells are growing in the presence of antibiotic, modified antibiotic does not accumulate in the medium, and the concentration of antibiotic in the medium remains unchanged when measured using conventional assays. Attempts

to localise the enzyme within the cell have indicated that it is periplasmic or associated with the cell membrane (Goldman & Northrop, 1976). These results have led to the hypothesis that a small amount of antibiotic is modified and this somehow blocks the transport of more antibiotic into the cell. There is also some evidence that the modified antibiotic binds less well to ribosomes (Bryan & Vanden Elzen, 1977) but the relative contribution of these factors to the overall resistance phenotype is not known.

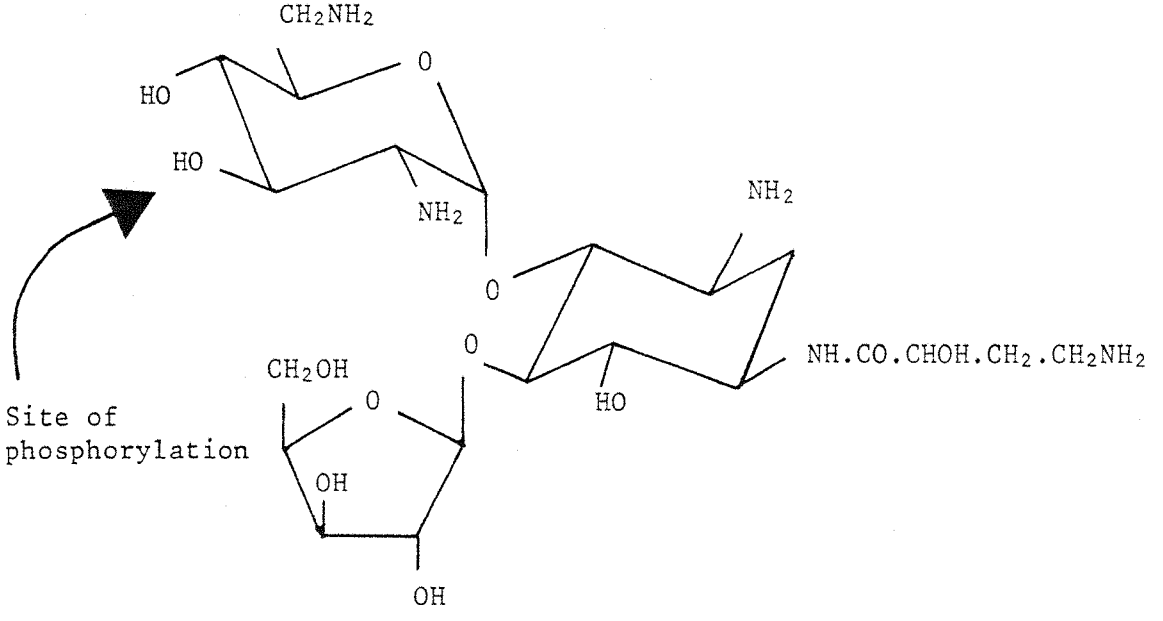


Fig. 5.1. Structure of Butirosin A showing the site of phosphorylation by APH(3') enzymes.

5.2 Localisation of the *APH* gene

When trying to locate a gene within a stretch of DNA, there are two simple methods which may be tried : (i) smaller fragments can be subcloned and tested to see if they contain the gene, and (ii) pieces of foreign DNA can be inserted into different restriction sites within the fragment to see if insertional inactivation occurs. The predicted size of 1,200 bp for the *APH* gene (Chapter 3) and the position of the Pst I and Bgl II sites within the 2.7 kb Sal I fragment (Fig. 3.9) made it convenient to use these methods in an attempt to further localise the gene. As a series of subclones of the 2.7 kb Sal I fragment had been constructed in M13 phage vectors, for the sequencing experiments (Chapter 6), these were used in the first localisation experiments.

The recombinant M13 phages carrying sections of the 2.7 kb Sal I fragment (Fig. 5.2) were inoculated into 1.5 ml aliquots of 2TY broth seeded with *E. coli* JM103 and incubated at 37°C for 5 h. These cultures were then used to inoculate LB agar plates and LB agar plates supplemented with 5 µg/ml ribostamycin. *E. coli* JM103 infected with M13 mp 9 was used as a control. After 48 h incubation at 37°C, slight growth could be detected on all the antibiotic-free plates but on the plates supplemented with ribostamycin growth could be detected only when the infecting phage carried the 2.7 kb Sal I fragment or the large Pst I - Sal I fragment. This indicated that the *APH* gene was located within the large Pst I - Sal I fragment and that the Bgl II site was within the *APH* gene.

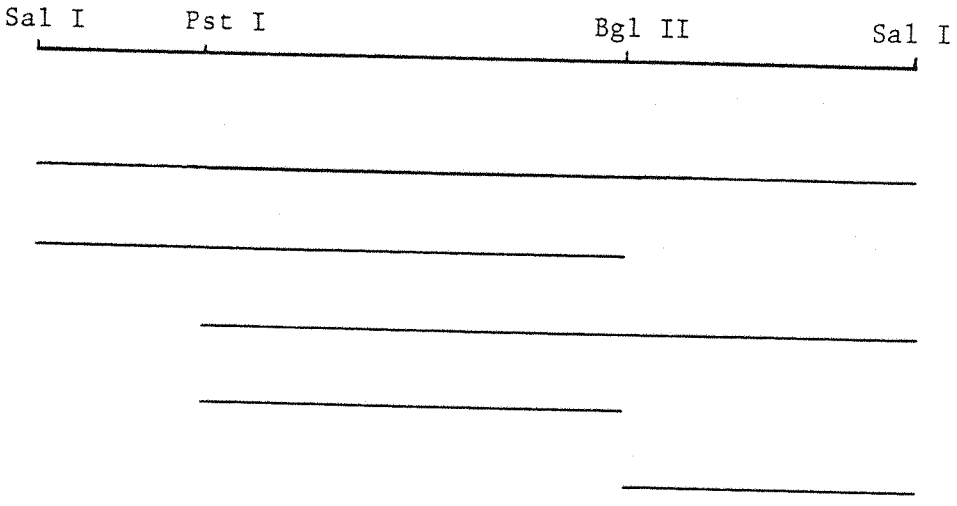


Fig. 5.2. Map of the 2.7 kb Sal I insert with the sections cloned in the recombinant M13 phages shown underneath.

5.3 Insertional inactivation of the *APH* gene

If the Bgl II site does lie within the *APH* gene, then the insertion of foreign DNA into this site should inactivate the gene. The foreign DNA used in this experiment was phage λ; ~ 0.5 μg of pCH5 and 1 μg of phage λ DNA were digested with Bgl II, mixed and the reaction was stopped by ethanol precipitating the DNA. The pellet was resuspended in 100 μl of T_4 DNA ligase buffer and ~ 0.1 units of T_4 DNA ligase was added. After incubation at 4°C overnight, the ligation mixture was used to transform 100 μl of transformation competent *E. coli* HB101 and transformants were selected for ampicillin resistance. Using velvet pads, the masterplates were replica plated onto LB agar supplemented with 50 μg/ml ampicillin and LB agar plates supplemented with 25 μg/ml ribostamycin. After incubation at 37°C for 18 h, the replicas were inspected. Several Amp^r Rib^s colonies were found; two of these were picked off and purified (CC1106 and CC1107). Agarose gel electrophoresis of rapid plasmid preparations from these strains showed that they both harboured a plasmid (pCH6

and pCH7 respectively) larger than pCH5. On digestion with Bgl II and agarose gel electrophoresis of the products, both pCH6 and pCH7 gave two bands, one equivalent to pCH5 and the other equivalent to a Bgl II fragment of phage λ (Fig. 5.3).

If it is the insertion of the Bgl II fragments of phage λ which has caused the inactivation of the *APH* gene, it should be possible to excise the fragment of phage λ and regain phosphotransferase activity. Aliquots of pCH6 and pCH7 (~ 0.2 μg) produced in a rapid plasmid preparation (Table 2.5) were digested with Bgl II, ethanol precipitated and resuspended in 100 μl of T_4 ligase buffer. T_4 DNA ligase (~ 0.1 units) was added and the mixture was incubated at 4°C overnight. This was used to transform 100 μl of transformation competent *E. coli* HB101. Transformants were selected for ribostamycin resistance. Two ribostamycin resistant transformants were purified and maintained from each experiment (CC1108 and CC1109 from the experiment with pCH6, and CC1110 and CC1111 from the experiment with pCH7). Agarose gel electrophoresis of rapid plasmid preparations and Bgl II digests of rapid plasmid preparations from these strains show that all harboured a plasmid (pCH8, pCH9, pCH10 and pCH11 respectively) indistinguishable from pCH5 (Fig. 5.3).

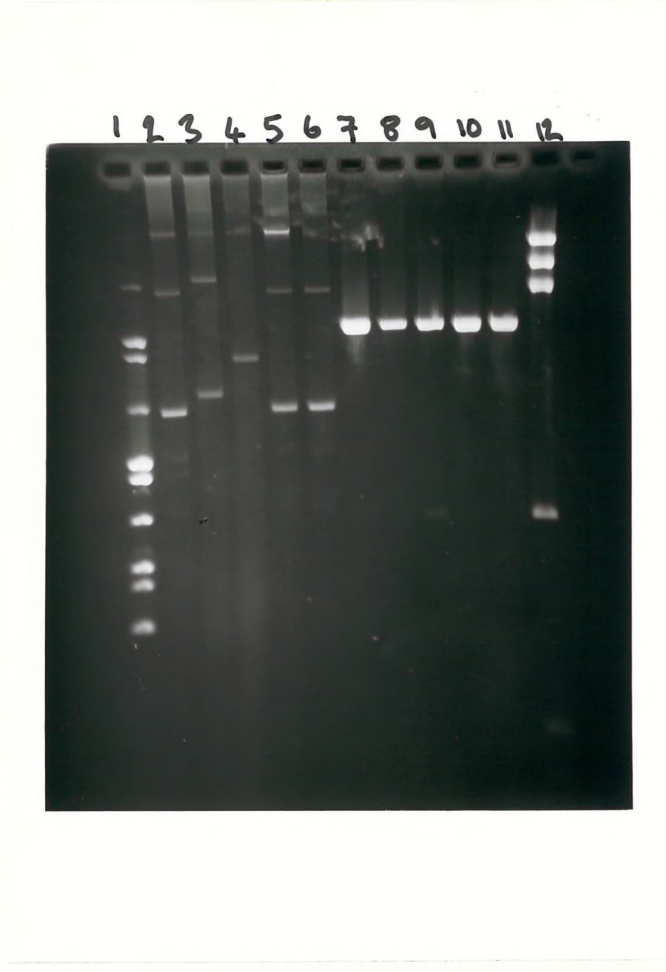


Fig. 5.3. Insertional inactivation of *APH* gene.

Track: 1) Plasmid M.W. standards purified from *E. coli* V517; 2) pCH5; 3) pCH6; 4) pCH7; 5) pCH8; 6) pCH10; 7) Bgl II digest of pCH5; 8) Bgl II digest of pCH6; 9) Bgl II digest of pCH7; 10) Bgl II digest of pCH8; 11) Bgl II digest of pCH10; 12) Bgl II digest of λ DNA.

The 0.3 kb insert in pCH6 (track 8) may not be visible.

To confirm the results obtained in the growth tests, the phosphotransferase activity was measured in crude cell extracts using butirosin as the phosphate acceptor (Section 2.12) (Table 5.1).

Table 5.1. Phosphotransferase activity in crude extracts of *E. coli* strains

See Methods section 2.12.

Strain	Protein concentration (mg/ml)	Activity (dpm transferred/ 10 µl of extract)	Relative activity (dpm transferred/ µg protein)
CC1105	10.6	30,570	288
CC1106	11.1	860	8
CC1107	10.3	580	6
CC1108	9.8	31,120	318
CC1109	12.1	32,300	270
CC1110	10.8	29,880	277
CC1111	11.3	30,400	269

These results confirm that the insertion of fragments of phage λ into the Bgl II site causes the inactivation of the *APH* gene. The simplest explanation of this is that the Bgl II site is within part of the *APH* gene. However, it is not possible to predict if it is a coding or regulatory region.

5.4 Characterisation of the *APH* gene product

One of the standard ways of characterising an aminoglycoside inactivating enzyme is to examine its activity *in vitro* against a range of different aminoglycosides (Dowding, 1979). The activity of the phosphotransferase produced from the *B. circulans APH* gene in three different hosts (*B. circulans*, *E. coli* and *S. lividans* 66) was measured in crude extracts using a variety of aminoglycosides as the phosphate acceptor (Section 2.12) (Fig. 5.4). As expected, the spectrum of activity is essentially the same in each case with butirosin, neomycin and ribostamycin being the best substrates and gentamicin, tobramycin and amikacin being the worst.

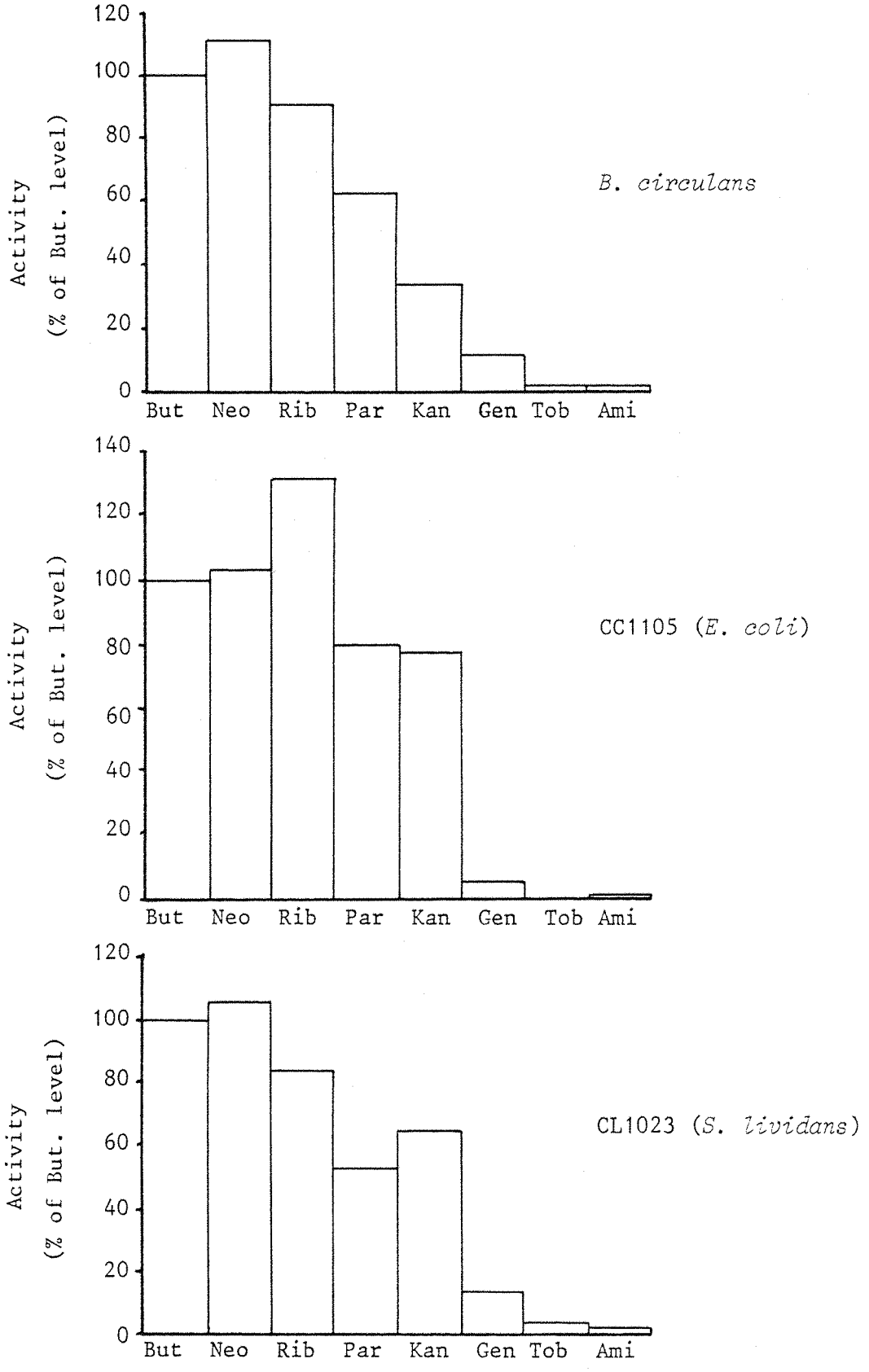


Fig.5.4. Phosphotransferase activity in crude extracts with different aminoglycosides used as phosphate acceptor.

The activity is expressed relative to the activity with butirosin.

But: Butirosin; Neo: Neomycin; Rib: Ribostamycin; Par: Paromomycin; Kan: Kanamycin A; Gen: Gentamicin; Tob: Tobramycin; Ami: Amikacin.

5.5 Summary

In summary, the 2.7 kb Sal I fragment has been further subcloned and insertional inactivation has shown that the Bgl II site lies within the *APH* gene. The activity of the phosphotransferase produced in different hosts has been measured against a range of aminoglycosides and the spectrum of activity is essentially the same in each.

CHAPTER 6

SEQUENCING OF THE *B. circulans APH* GENE

6.1 Introduction

In recent years, several methods have become available for rapidly accumulating DNA sequence information. These methods can be divided into two classes : a chemical method (Maxam & Gilbert, 1977), which relies on the selective chemical cleavage of bonds in single stranded DNA, and the enzymic methods of Sanger. The first of the enzymic methods was the 'plus and minus' method (Sanger & Coulson, 1975). This was used to deduce the sequence of the bacteriophage ϕX174 (Sanger *et al.*, 1977(b)). However, this method on its own was not considered completely reliable and was soon replaced by the chain termination method (Sanger *et al.*, 1977(a)).

6.2 The chain termination method of sequencing

It was already known that dideoxythymidine triphosphate (ddTTP) was a substrate for DNA polymerase and that when incorporated into growing chains it caused the termination of chain growth (Atkinson *et al.*, 1969). This was extended by the synthesis of the other three deoxynucleoside triphosphate analogues : ddATP, ddGTP and ddCTP. These were also found to cause chain termination. Conceptually, it is not difficult to see how these chain terminators could be used to sequence stretches of DNA. If a suitably primed piece of single stranded DNA is used as a template for DNA synthesis by DNA polymerase I (Klenow fragment) in the presence of the four deoxy-nucleoside triphosphates, and one dideoxy analogue is added (e.g. ddTTP), then at some T residues the analogue will be incorporated and cause chain termination. If the ratio of deoxy/dideoxy is just right, then for a certain distance after the primer (usually 500-600 bases) termination will occur in some chains at each T residue. This produces a nested set of fragments, whose lengths represent the distance from the primer to the T residue. These fragments can be separated by

gel electrophoresis under denaturing conditions and, when the DNA is radiolabelled (usually by including $\alpha[^{32}P]$dATP in the incubation), the fragments can be detected by autoradiography of the dried gel. If a similar experiment is performed for each nucleotide, and the resulting fragments are run next to each other on gels, the sequence can be read from the position of the bands on the autoradiograph (Fig. 6.1).

The difficulty with this method was in obtaining a suitably primed single stranded DNA template. Initially, this was overcome by using the single stranded DNA bacteriophage ϕX174. The latter has a double stranded DNA replicative form (RF DNA) that can be isolated from infected cells, and a single stranded DNA infective form which is excreted. Small restriction fragments from the RF DNA could be isolated, heat denatured and allowed to anneal to the single stranded infective form, thus producing a primed single stranded DNA template.

A more general approach has been to isolate large restriction fragments, separate the strands on denaturing gels, and use small restriction fragments to prime synthesis on the larger fragments. The disadvantage of this system is that a very detailed restriction map must be available before sequencing can be attempted.

6.3 Development of the M13 phage vectors

Messing realised that filamentous single stranded DNA phages offered a convenient vector system for producing single stranded DNA templates and chose the bacteriophage M13 for development as a cloning vector. (The development of the M13 vectors is summarised in Messing, 1982.) M13 is a male-specific bacteriophage; after the single stranded genome (+ strand) has infected a cell, the complementary-strand is synthesised to form the double stranded replicative form. The replicative form is used as a template to produce more + strand phage genomes; these are packaged and extruded without lysis of the host. On agar plates, a turbid plaque is seen in the bacterial lawn as the infected cells are slower growing. This life cycle has several features which are desirable for a cloning vector : (i) the double stranded replicative form DNA can be isolated like a plasmid and used to clone restriction fragments in the normal way; (ii) the replicative form DNA can be efficiently introduced into transformation competent

E. coli and subsequently give rise to infective phage particles; and (iii) M13 is a filamentous phage and it has been found that there is little or no size constraint on the fragments that can be cloned. Thus, a restriction fragment can be cloned in the replicative form DNA, the cloning mixture introduced into *E. coli* and the plaques produced used to infect small cultures. The single stranded DNA can be isolated from the excreted phage particles.

The problem remaining was how to detect phages carrying inserts. Screening by size would be laborious, insertional inactivation of antibiotic resistance would involve replica plating, and so insertional inactivation of the α complementation of β-galactosidease was used. The first 145 amino acids of β-galactosidease (the α peptide) can complement β-galactosidease molecules, which have part of this portion of the protein missing, to give an active enzyme. The promotor region and enough of the *lac* Z gene to code for the α peptide were inserted into M13. The *E. coli* host used was a *pro-lac* deletion and carried an F' factor which complemented the chromosomal *pro* deletion, and also carried a partially deleted *lac* Z gene, which could be complemented by the α peptide. Thus, the F' factor needed for M13 infection can be maintained by growing the cell without added proline, and when M13 infects the cell the M13 encoded α peptide will complement the partially deleted *lac* Z gene to give an active β-galactosidase. In the presence of the *lac* operon inducer IPTG and the lactose analogue BCIG, an active β-galactosidase results in a blue plaque (when cleaved by β-galactosidease BCIG releases the blue dye bromo-chloroindole). A series of vectors were produced with different cloning sites introduced into the promotor and proximal region of the α peptide gene in M13 (Fig. 6.2). When fragments are cloned in these sites, the production of the α peptide is stopped and a white plaque is produced. This enables phages with inserts to be detected by visual inspection of the plates. Also, because the cloning sites are all within a short segment of the vector, a synthetic primer complementary to a region just prior to these sites can be used to prime the sequencing of all cloned fragments.

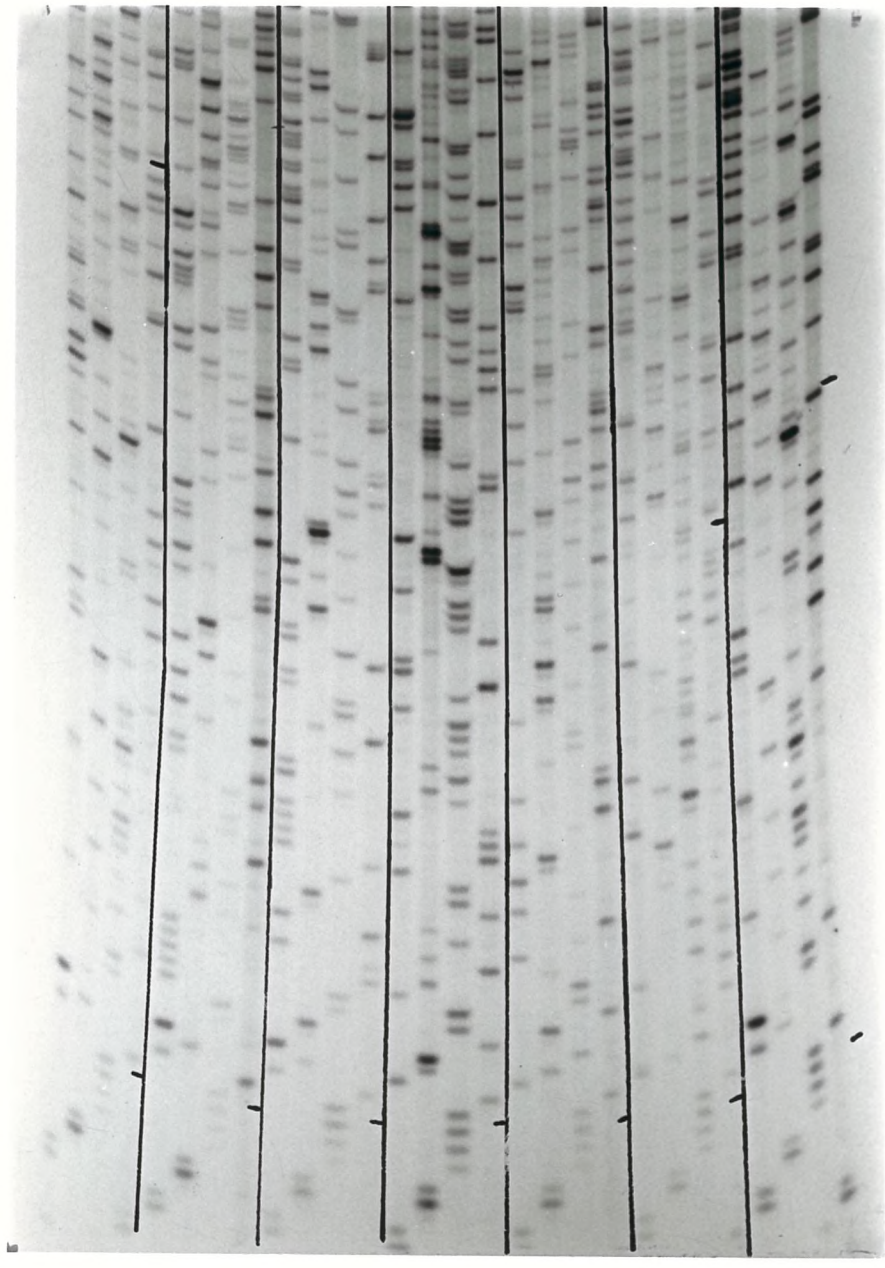


Fig. 6.1. Autoradiograph of a 'short' sequencing gel.

Seven clones are shown; all are Sau 3A fragments cloned into the Bam HI site. The four tracks are T, C, G and A (left to right) and the start of the insert is marked in the A track.

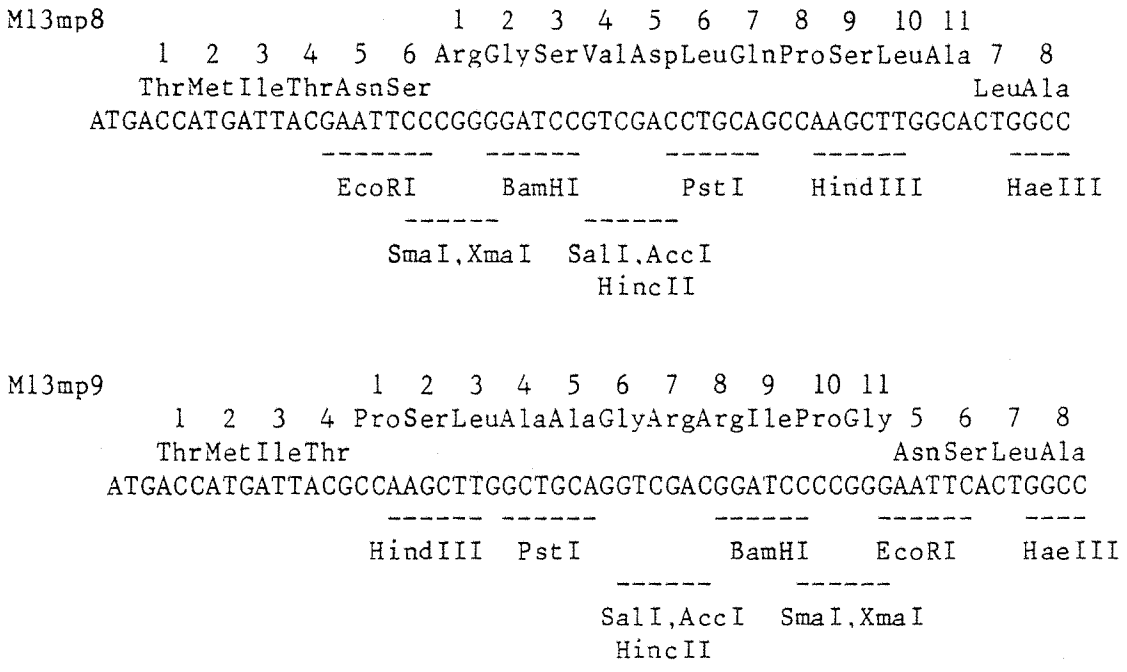


Fig. 6.2. Restriction enzyme sites available for cloning in M13mp8 and M13mp9. These are the two vectors used for sequencing in this work.

The various methods for producing libraries of fragments are discussed in Chapter 3. The final assembly of the sequence data is computer aided. The programmes used in this work were those of Roger Staden (listed in Staden, 1982) and run locally on an ICL 2970 machine by Dr. I.G. Giles.

6.4 Sequencing of the *B. circulans APH* gene

In general, there are two strategies that can be used when sequencing a section of DNA using the chain termination method. One is random: small fragments are generated and cloned non-selectively; the other is non-random , specific restriction fragments are isolated and cloned. In order to sequence the 2.7 kb Sal I insert in pCH5, a combination of these approaches was used. Initially, specific fragments were cloned in the phages M13 mp 8 and M13 mp 9. This allowed assymetric fragments (e.g. Bgl II-Sal I) to be cloned in both orientations. When the various fragments had been cloned and sequenced, a considerable amount of primary data had been accumulated. To complete the sequence, these data had to be confirmed and the intervening sequences determined.

Libraries for further sequencing experiments were generated by purifying the 2.7 kb Sal I fragment, digesting it with frequent cutting enzymes (Sau 3A, Taq I and Hpa II) and cloning the fragments in appropriately cut M13 mp 9. When sequenced, these libraries provided new information but, for several reasons, they proved to be an inefficient use of resources. Firstly, many of the fragments produced were small (50-100 bp) and small fragments seem to be preferentially cloned, so some of these would be sequenced many times, while the larger fragments might not be sequenced at all. Secondly, in each sequencing experiment, it is possible to sequence 250-350 bp of each clone, so if the cloned fragments are short (> 100 bp), then over half of the sequencing capacity of the experiment is not used. Finally, if only short fragments are sequenced, then the number of fragments required to produce all the overlaps and complete the sequence increases enormously.

To try and overcome these problems a new method of generating the libraries was developed. The 2.7 kb Sal I fragment was purified and polymerised, using T_4 DNA ligase. The polymerised fragment was

Fig. 6.3. Sequence of the *B. circulans APH* gene

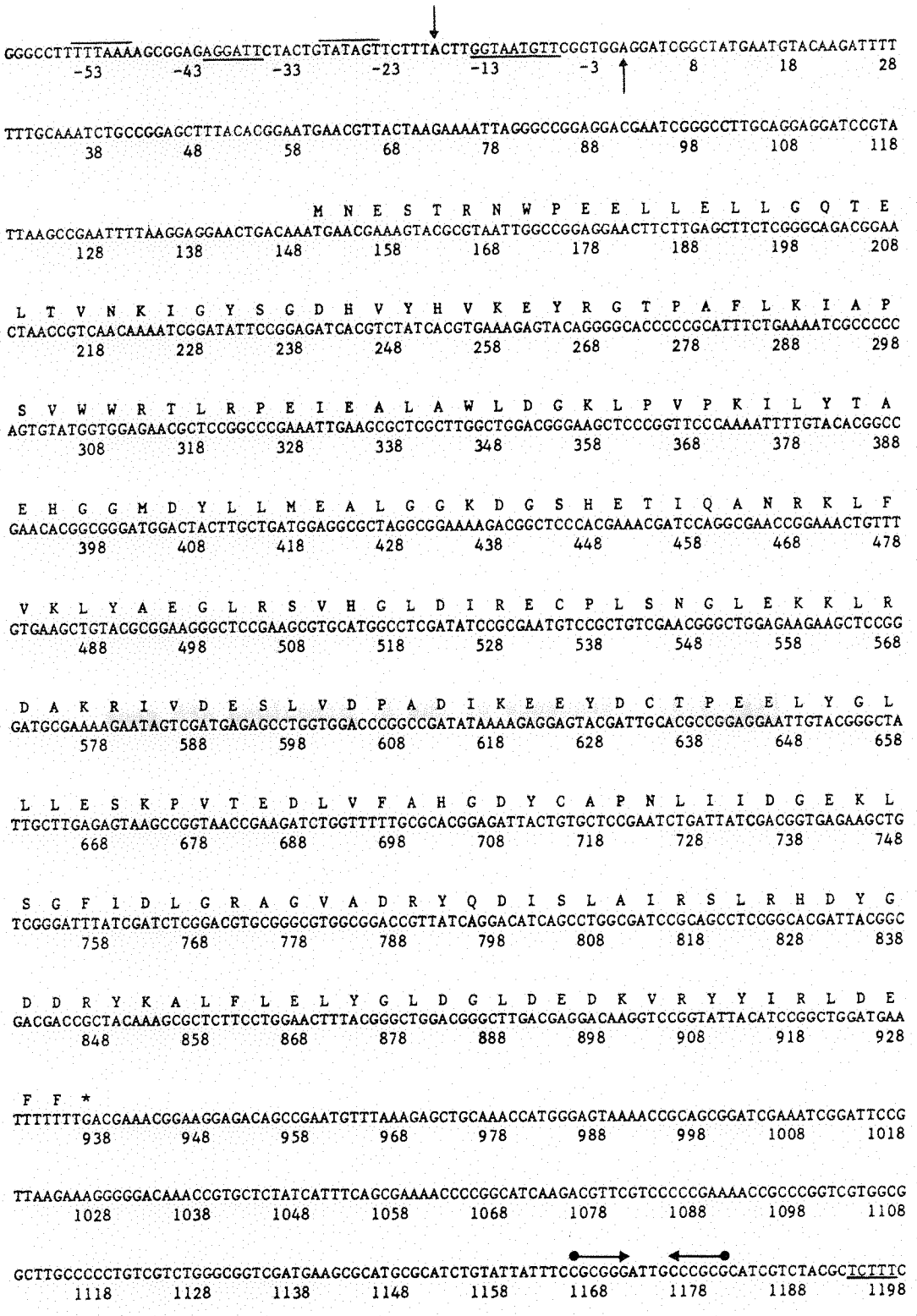


Fig. 6.3. Sequence of the *B. circulans APH* gene.

The possible promotors, transcriptional start and terminator are marked (see Figs. 6.8 and 6.9 for details). The translated protein sequence is shown above the coding region.

partially digested with a frequent cutting enzyme (Sau 3A, Taq I or Hpa II) and the partial digest was run on an agarose gel. The section of the gel containing fragments of 400-1,000 bp was cut out and the DNA purified. These fragments were then cloned into appropriately cut M13 mp 9 and sequenced. This method ensured that the full sequencing capacity of each experiment was used and allowed the sequence of the entire 2.7 kb Sal I fragment to be determined.

It had already been established that the Bgl II site was within the *APH* gene (Chapter 5), so the sequence in the region of this site was translated in all three reading frames and in both directions. This revealed the possibility of one long open reading frame for 786 bases (Fig. 6.3) which could code for a protein of 29,800 daltons. This is in broad agreement with the observed molecular weight of 28,500 daltons (Davies & Smith, 1978). The +1 and -1 reading frames contain a total of 20 stop codons in this region. The consensus sequence shown in Fig. 6.3 is built up of many individual gel readings. In this area the sequence has been determined at least once from each strand. A schematic representation of the gels that make up the consensus in this area is shown in Fig. 6.4. A full listing of the data for the whole sequence is presented in Appendix I.

The sequence can now be used in an attempt to answer some of the questions posed in Chapter 1 : are there any homologies with the *APH* genes from other organisms, and do the regulatory sequences of the *APH* gene from *B. circulans* indicate how the antibiotic metabolising genes are linked to differentiation?

6.5 Homologies with phosphotransferases from other organisms

Three other *APH* genes had been sequenced. These were derived from TN903 (Oka *et al.*, 1981), TN5 (Beck *et al.*, 1982) and *S. fradiae* (Thompson & Gray, 1983), so their deduced protein sequences were available for comparison with the one deduced from the *APH* gene of *B. circulans*. It was decided to search for homology using the DIAGON program of Staden (1982). This is a program which looks for homology between pairs of protein or nucleic acid sequences. The program is designed such that areas of homology will be highlighted,

even if insertions or deletions occur. In the simplest mode of usage, the program displays a graph with one sequence aligned along the abscissa and the second along the ordinate. For each cell a point is plotted if the two sequences contain the same amino acid; for a perfect match a diagonal line will be produced; hence, the name of the program. If an insertion (or deletion) occurs in proteins of otherwise identical structures, two diagonal lines will be observed, one leading up to the site of insertion (or deletion) and the other after it. The size of the discontinuity is a measure of the size of the insertion. One of the problems with this approach is that there is a 1 in 20 chance of the same amino acid occurring in the two sequences by chance. Consequently, the areas of homology will be present in a background of random matches. This approach requires absolute identity between the amino acids, whilst it may be relevant to take into account the degree of similarity between them. In the program already described, an absolute identity is given a score of 1 and non-identity a score of 0.

The existence of 7 matches in a 10 amino acid stretch (window) is taken to be highly significant, so a score of 7 can be set. The display can be calculated by comparing all possible stretches of 10 residues for one sequence against all those for the second. For each pair of sequences, an identity score is calculated and, if it is 7 or greater, the sequences are plotted on the diagonal plot. The user can specify the window length and the minimum level of significance to be displayed. The present version of the program uses fractional scores to allow for similarities between amino acids; thus, isoleucine → leucine would only be considered a minor change.

The output of the DIAGON program comparing the *B. circulans* phosphotransferase with the other phosphotransferases in turn is shown in Fig. 6.5. Regions of homology between the *B. circulans* phosphotransferase and the other phosphotransferases were easily detected. Interestingly, the homology was confined to the C terminal portion of the molecule. The DIAGON program was also used to study the homology between the *B. circulans* phosphotransferase and the *S. fradiae* phosphotransferase in more detail. The two sequences were compared and the output was recorded at various levels of significance (0.1 - 0.00001% significance) (Fig. 6.6). An examination of the output from

this experiment clearly shows that the most highly conserved sequences are in the C-terminal portion of the molecule and only under the least stringent conditions used (0.1% significance) could homology be detected in the N-terminal portion of the molecule.

The output of the DIAGON program could also be used to construct a direct comparison of the phosphotransferase sequences. The regions of strong homology detected by the DIAGON program were lined up above each other and the intervening stretches were fitted together using as few padding characters as possible (Fig. 6.7). The conservation of sequence in the second half of the molecule is easily seen and the most highly conserved areas presumably identify regions of the protein that are involved in the formation of the active site of the enzyme. The level of homology between the phosphotransferases is good evidence that they may have a common origin.

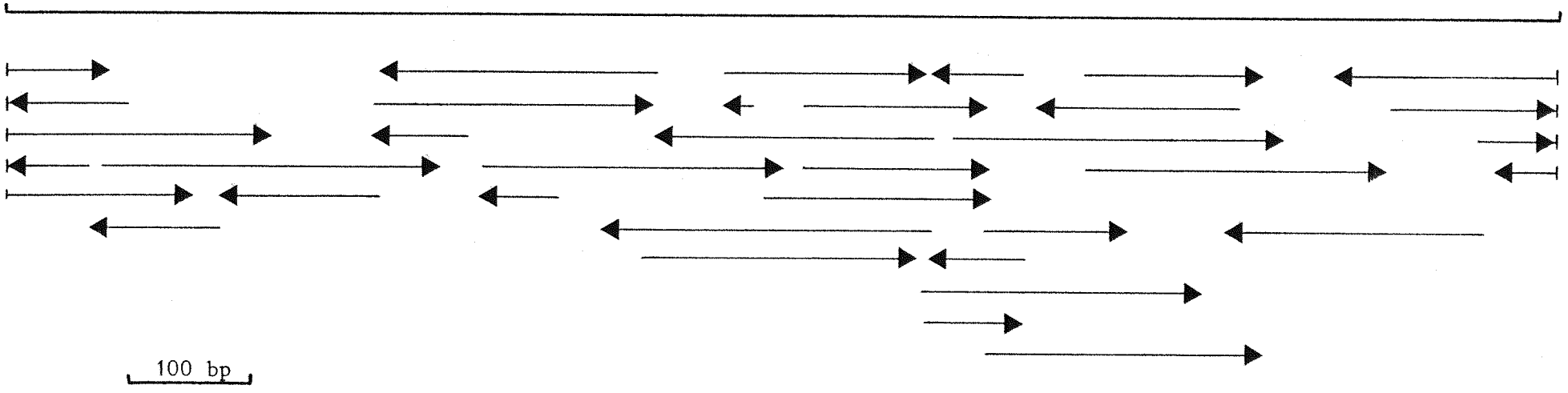


Fig. 6.4. Diagramatic representation of the gel readings used to construct the consensus sequence shown in Fig. 6.3. The arrows indicate the direction in which the sequence data were obtained.

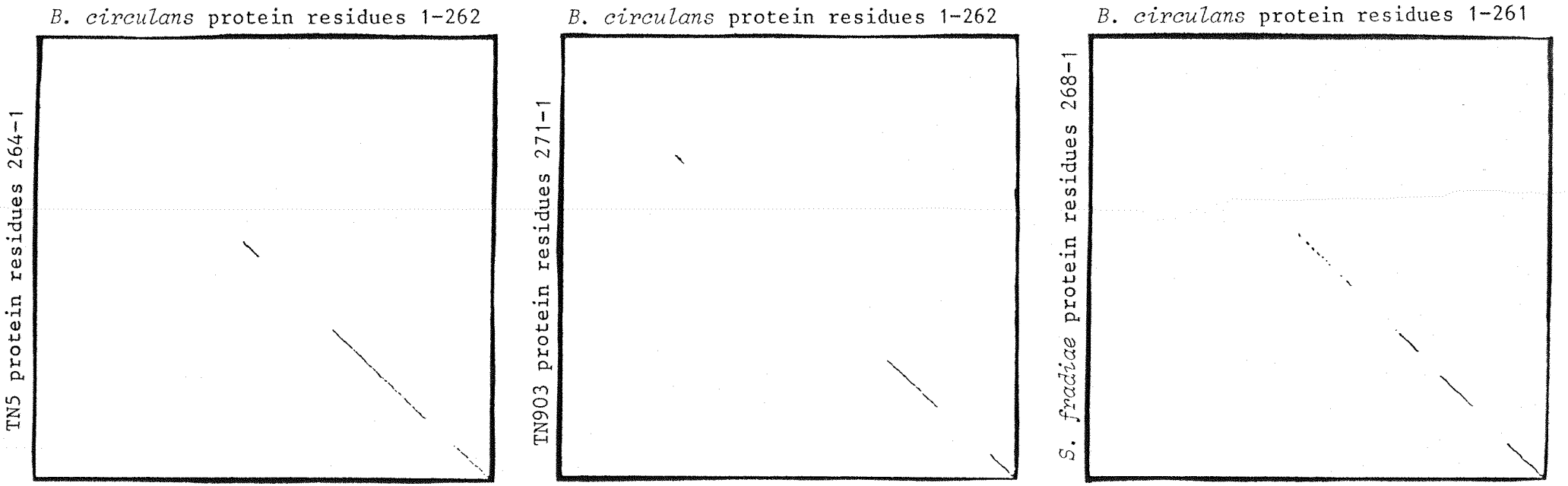
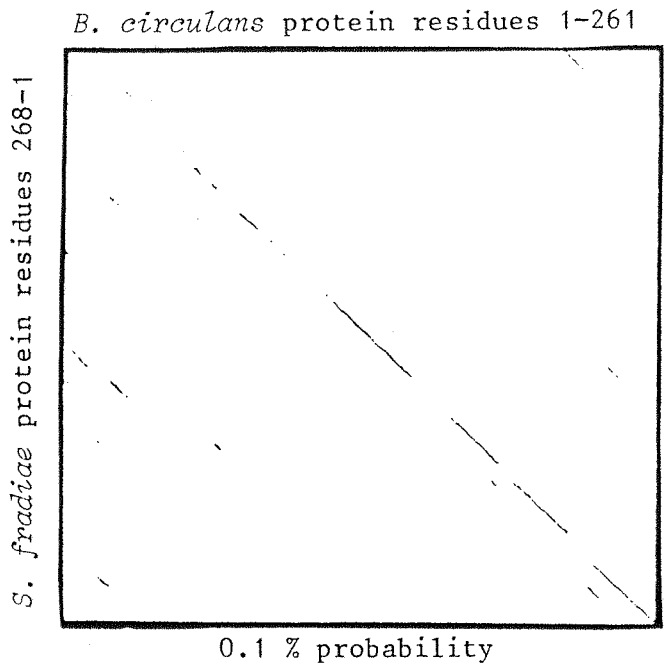
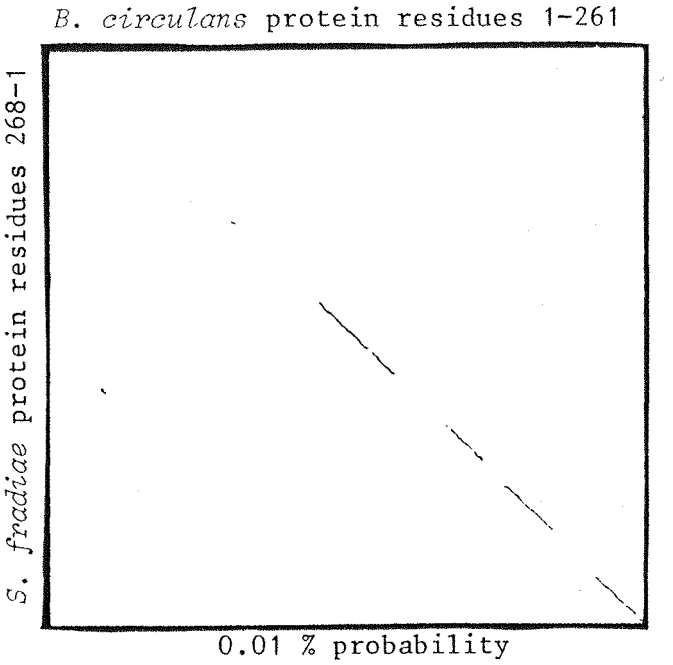
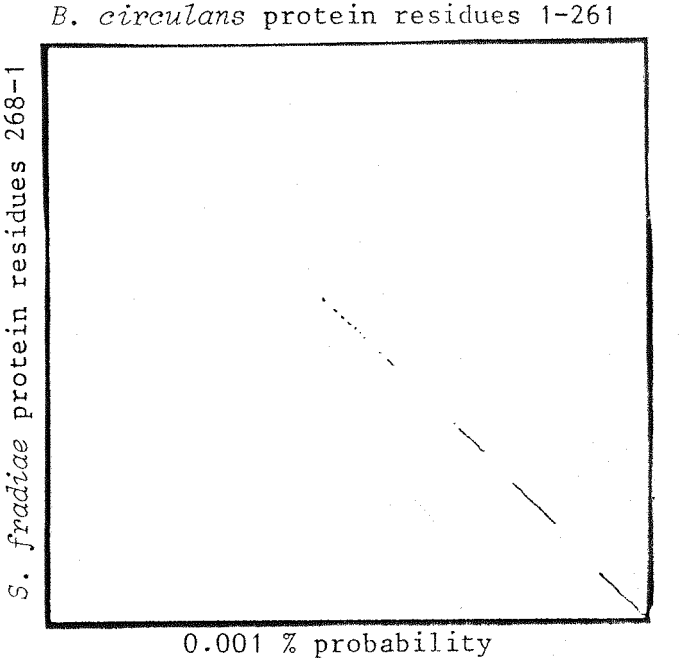
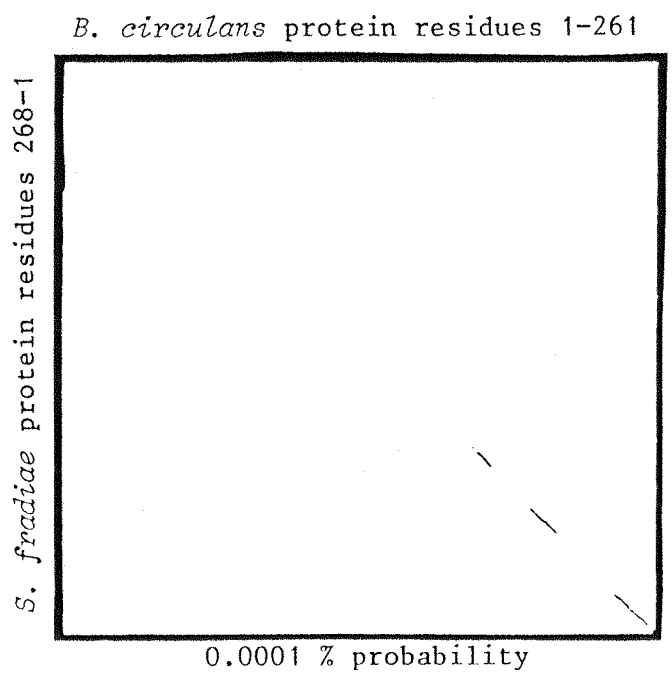


Fig. 6.5. DIAGON plots of homology between pairs of phosphotransferase sequences. The figures were obtained using a double-matching expectation score equivalent to a 0.001% probability level and a window length of 21 residues.









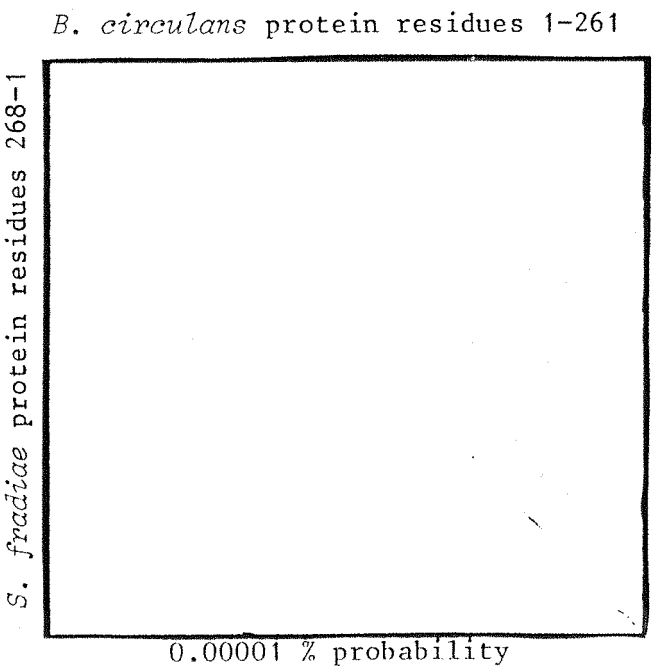


Fig. 6.6. DIAGON plots of homology between the *B. circulans* and *S. fradiae* phospho-transferases. The figures were obtained using a window length of 21 residues and a double-matching expectation score equivalent to the probabilities indicated on the graphs.

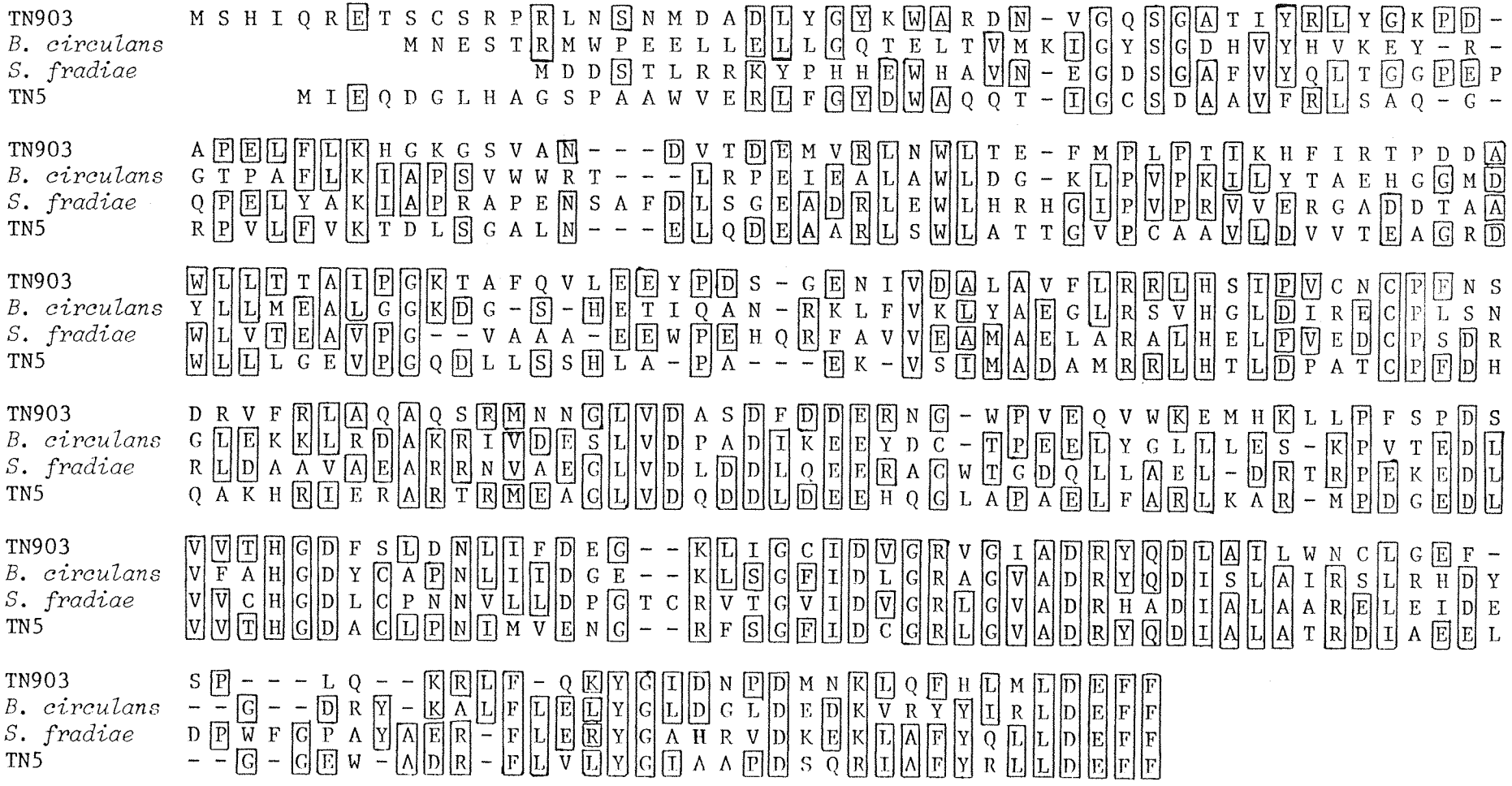


Fig. 6.7. Comparison of the four phosphotransferase protein sequences showing the regions of homology. Initially the sequences were lined up using the regions of homology revealed by the DIAGON plots; the intervening regions were then fitted using as few padding characters as possible. Pairs of amino acids (I/L, D/E and K/R) are considered to be equivalent.

6.6 Regulatory sequences of the *B. circulans APH* gene

At least two processes are involved in the expression of a protein-encoding gene, transcription and translation. The gene contains nucleotide sequences which act as signals to initiate and regulate these processes.

In transcription, the two most important types of sequences are promotors and terminators, although other sequences may be involved, such as binding sites for regulatory proteins (e.g. the binding site for the *lac* reperssor in the *lac* operon of *E. coli*). Promotors are RNA polymerase binding sites and specify the start sites for transcription. They comprise two short sequences (5-7 bases) centred approximately 10 and 35 bases before the first transcribed base. RNA polymerase is a complex enzyme and the apo-enzyme is characterised as an $\alpha_2\beta\beta'$ structure, which will initiate transcription randomly. However, when a σ subunit binds to the complex, the polymerase will only initiate transcription at specific sites (promotors) (Losick & Pero, 1981). In *E. coli* considerable homology exists between the known promotors (reviewed in Rosenberg & Court, 1979) and this allows a composite or generalised promotor sequence to be deduced (Table 6.1). Although fewer *B. subtilis* promotors have been sequenced, the normal vegetative promotors fit into the same pattern as the *E. coli* promotors (Moran *et al.*, 1982(a), Table 6.1). Once the RNA polymerase has bound to the promotor, transcription usually starts with a purine nucleotide, normally an A, 6-9 bases beyond the -10 region.

When the sequence leading up to the translational start of the *APH* gene was examined, no good matches to the consensus promotor sequence could be found. However, in *B. subtilis* several other promotor sequences have been identified (listed in Johnson *et al.*, 1983). These promotors are involved in the expression of developmentally controlled genes. A working hypothesis is that during endospore formation, new σ factors for RNA polymerase are produced which bind to the RNA polymerase apo-enzyme and change the promotor recognition specificity. Thus, genes which have these different promotors are only expressed when the different σ factors are produced. As it is thought that the genes for antibiotic metabolism are also under developmental control (Demain, 1974), a search was made for promotors similar to the other known *B. subtilis* promotors. One such promotor sequence was

found; this was similar to those of the *spoVC* and *spoVG* genes (Moran *et al.*, 1981; Moran *et al.*, 1982(b)) and was located 170 bases before the start of translation (Fig. 6.3, Table 6.1). The *spoVC* and *spoVG* genes are both transcribed by an RNA polymerase which has a σ factor of 37,000 daltons and hence are known as σ^{37} promotors.

Table 6.1. Sequences of σ^{55} and σ^{37} promotors.

Promotor	-35 region	-10 region
E. coli (consensus)	TTGACA	TATAAT
B. subtilis σ^{55} (consensus)	TTGACA	TATAAT
APH σ^{55}	TTTAAA	TATAGT
spo VC σ^{37}	AGGTTT	GGTATTGTT
spo VG σ^{37}	AGGATT	GGAATTGAT
APH σ^{37}	AGGATT	GGTAATGTT

It is known that σ^{37} promotors are not transcribed by *E. coli* RNA polymerase (Johnson *et al.*, 1983) but the *B. circulans APH* gene was expressed when cloned in *E. coli* on pBR322. If we assume that the σ^{37} promotor is the gene's normal promotor, there are two possible explanations of this: either the transcription is initiated in the vector and reads through into the insert, or there is a second promotor within the cloned fragment that is recognised by *E. coli* RNA polymerase. The first explanation seems unlikely as the *APH* gene was cloned into three different sites in the vector and a similar level of activity was produced in each case (see Chapter 3). This would indicate that we are not dealing with read through transcription. If we assume that a second promotor is present, it would not need to be a strong promotor, as the specific activity of the phosphotransferase in crude extracts of *E. coli* is only twice that in *B. circulans*, even though it is present on a plasmid with a high copy number. An examination of the sequence reveals several possible -10 regions but only one of these is preceded by a possible -35 region (Fig. 6.3 and Table 6.1). Interestingly, the possible σ^{55} promotor is interlaced with the σ^{37} promotor (Fig. 6.8). This has important implications for the regulation of the gene (Chapter 7).

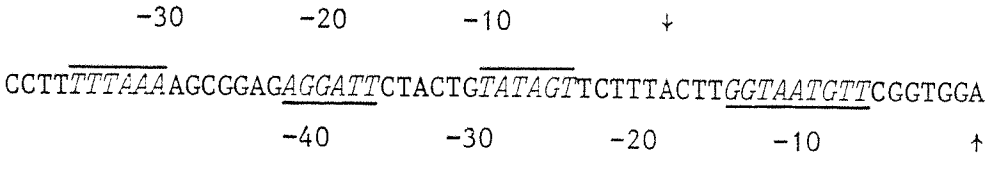


Fig. 6.8. Possible interlaced promotors. The σ^{55} promotor is shown above the sequence and the σ^{37} below. Consensus promotor sequences are marked and proposed transcriptional starts are marked with an arrow.

Terminators are sequences that are involved in the termination of transcription. A comparison of the known terminator sequences has revealed three common features : (i) an inverted repeat precedes the termination site; (ii) U residues are found in the terminus of the RNA transcript; and (iii) G/C rich sequences, variable in length from 3 to 11 continuous G/C base pairs, are found preceding the stop site (Rosenberg & Court, 1979). An examination of the sequence following the translated region revealed only one region that fitted this pattern (Figs. 6.3 and 6.9).

```
      TT
     A  G
      GC
      GC
      GC
      CG
      GC
TTATTTCCGCATCGTCTACGCTCTTTCGG
```

Fig. 6.9. Possible terminator structure. The bases at which termination might occur are underlined.

The centre of the inverted repeat is 240 bases after the end of translation and, as expected, the U residues at the end of the RNA transcript are located approximately 20 bases after this (Holmes *et al.*, 1983).

The sequences involved in the initiation of translation are the ATG initiation codon and the 'Shine-Delgarno' sequences. The Shine-Delgarno sequence is a short sequence just prior to the ATG initiation codon (usually 7-14 bases before in gram positive organisms, Moran *et al.*, 1982(a)), which is complementary to the 3' of the 16s ribosomal RNA. It is thought that the 16s rRNA base pairs with the Shine-Delgarno sequence and helps to stabilise the initiation complex. The sequence of the 3' end of the *B. subtilis* 16s rRNA is known (McLaughlin *et al.*, 1981). When the region immediately before the initiation codon of the *APH* gene was examined, a 7 base region capable of base-pairing with the 3' end of the *B. subtilis* 16s rRNA was found 10 bases before the initiation codon (Figs. 6.3 and 6.10).

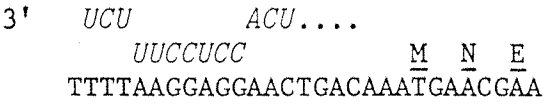


Fig. 6.10. Possible Shine-Delgarno sequence showing base-pairing with the 3' end of the 16s rRNA.

In summary, the sequence of the *B. circulans APH* gene has been examined and possible regulatory signals and a putative coding region have been identified. The significance of these will be discussed in Chapter 7.

CHAPTER 7

GENERAL DISCUSSION

The hypothesis that antibiotic-producing organisms are the original source of clinically significant antibiotic resistance (Benveniste & Davies, 1973) has been tested using both immunological and nucleic acid hybridisation techniques, but it has not been possible to demonstrate any homology between different classes of phosphotransferases (Davies & Smith, 1978; Courvalin *et al.*, 1978). The nucleic acid hybridisation experiments relied on the detection of heteroduplex DNA under the electron microscope. A more sensitive hybridisation experiment is to use the Southern Blot technique (Southern, 1975), so we used the *B. circulans APH* gene as a probe against total DNA from four aminoglycoside producing *Streptomyces* (Chapter 4). Even under conditions of low stringency, no hybridisation could be detected.

Significant homology between the TN5 and TN903 phosphotransferases had already been reported (Beck *et al.*, 1982). When the DIAGON program was used to compare the deduced protein sequences for the TN5, TN903, *S. fradiae* and *B. circulans* phosphotransferases, significant homology could be detected between all pairwise combinations. This allowed a direct comparison of the sequences to be made (Fig. 6.7) and also offers possible explanations for the failure of other experiments to detect homology. In the hybridisation experiments, short stretches of direct homology will be detected more easily than larger stretches of 50-70% homology. The codon usage in the genes varies from one organism to another (in particular, the third base of each triplet varies considerably), so although the protein product may be identical the coding sequences can vary significantly. This would make the formation of stable heteroduplexes unlikely and is also the reason for using deduced protein sequences in the DIAGON comparisons rather than nucleic acid sequences.

When the protein sequences are compared, we can see that certain regions are highly conserved but that the intervening regions are less conserved. This may indicate that the conserved sequences are

involved in the formation of the active site of the enzyme. If this is the case, we might expect these sequences to be buried in the tertiary structure and the variable regions to be presented at the surface. This would explain the lack of immunological cross-reactivity (Davies & Smith, 1978) and the inability to detect homology by immunological methods. One interesting feature is the high level of conservation of proline residues, indicating that they may be important in maintaining the tertiary structure.

The evidence presented here is consistent with the hypothesis that the resistance determinants come from a common source, with enzymic function being maintained and codon usage being altered to fit the normal pattern of the host organism. However, it cannot be used to argue that the antibiotic-producing organisms are the original source of the resistance determinants; indeed, it is not clear what evidence could be used in this way. A much larger collection of sequences from different organisms would be needed before a serious attempt to construct a phylogeny for the *APH* genes was possible.

As we have seen, the several thousand different antibiotics that have been characterised possess diverse chemical structures and biological actions. However, one striking similarity is that virtually all antibiotics are produced by differentiating micro-organisms, both prokaryotic and eukaryotic. Also, many mutations which affect antibiotic metabolism have pleiotrophic effects on differentiation and *vice versa*. These and other observations have led to the hypothesis that antibiotic metabolism and differentiation are linked, or that antibiotic metabolism is under developmental control (Demain, 1974). Before this can be discussed in more detail, a brief description of endospore formation in the genus *Bacillus* is necessary.

Endospores are produced at the end of normal, vegetative growth; they are able to withstand destructive agents (e.g. heat) and remain viable for long periods in the absence of nutrients. The exact trigger for sporulation is not known but, in general, it seems to be a response to nutrient depletion. The probability that cells will become committed to sporulate during vegetative growth is dependent on the carbon and nitrogen sources. When easily metabolised carbon and nitrogen sources are present, the probability is low, but if either

is replaced by a less readily metabolised compound, the probability increases. Under some conditions, phosphate limitation can also induce sporulation (Chambliss, 1979).

Endospore formation can be divided into eight different stages (Hanson, 1979) (Fig. 7.1). Stage 0 is defined as the end of logarithmic growth; Stage I is marked by the formation of an axial filament of condensed chromatin and Stage II by the formation of an acentric forespore septum. During Stage III, the cytoplasmic membranes of the mother cell (sporangium) grow and engulf the forespore. At Stage IV, the peptidoglycan of the germ cell wall is deposited. Subsequently, the cortical peptidoglycan is deposited between the inner and outer forespore membranes; this continues until Stage VI. During Stage V, the spore coat proteins are deposited outside the outer forespore membrane. Maturation of the spore occurs during Stage VI: the spore coat proteins are modified and there is hydrolysis of some bonds in the cortical peptiglycan to produce a loosely cross-linked expandable polymer. The sporulation process is completed in Stage VII by liberation of the mature endospore via autolysis of the sporangium.

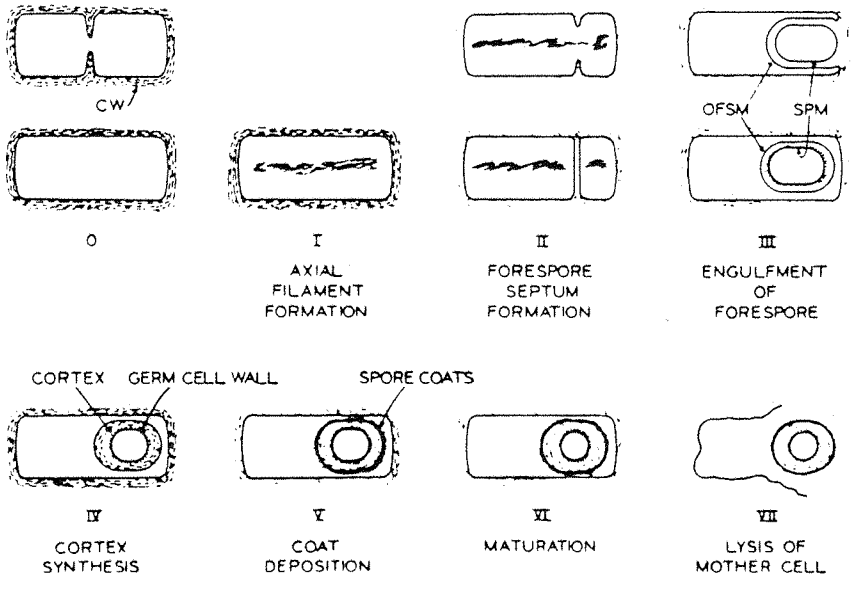


Fig. 7.1. Cytological stages in sporogenesis: CW, cell wall; OFSM, outer forespore membrane; SPM spore plasma membrane.

Although neither process is understood in great detail, there are interesting and possibly significant similarities between the factors that control the onset of sporulation and those that control antibiotic biosynthesis. In batch cultures of antibiotic-producing organisms growing in rich complex media, high levels of antibiotic are normally only produced after most of the cellular growth has stopped (the control of antibiotic biosynthesis has recently been reviewed - Martin & Demain, 1980). Also, many fermentation studies have shown that, while glucose is an excellent carbon source for growth, its antibiotic yield is poor. The antibiotic yield is often increased when more slowly utilised carbon sources, such as glycerol, are used (Howells *et al.*, 1972). The effect of different nitrogen sources on antibiotic production is not as well documented as that of carbon sources but a similar picture emerges. Readily metabolisable nitrogen sources, such as ammonia, seem to delay the onset of antibiotic production. Studies to identify suitable nitrogen sources for antibiotic production in defined media often result in the selection of slowly-metabolised amino acids, such as proline, which may cause nitrogen limitation (Dulaney, 1948). Finally, phosphate inhibits the synthesis of many antibiotics and industrial fermentations are frequently phosphate limited. In batch culture, antibiotic production often occurs only after the phosphate is depleted, Thus, *S. gresius* will grow in the presence of 300 mM phosphate but candicidin production is abolished by the presence of 10 mM phosphate (Liras *et al.*, 1977). As the antibiotics, whose production has been shown to be regulated by phosphate, belong to several different chemical groups, it seems unlikely that specific biosynthetic enzymes are being inhibited. It is tempting to speculate that phosphate regulation reflects a general control mechanism; however, the nature of this is unknown.

At the physiological level, we can see many parallels between the onset of sporulation and antibiotic production and resistance; we can now ask if there are similar parallels at the molecular level. One important method of control in the sporulation process is the production of different σ factors that bind to the RNA polymerase and alter the promotor specificity (Losick & Pero, 1981). So far, seven different σ factors have been described in *B. subtilis*, and each is associated with a different promotor sequence (listed in Johnson *et al.*, 1983).

Five of these are derived from *B. subtilis* and two are phage encoded being involved in the development of phage SPO1. An examination of the sequence for the *APH* gene of *B. circulans* revealed a possible σ^{37} promotor. Two other genes are known to have σ^{37} promotors: these are the *spoVC* and *spoVG* genes (Moran *et al.*, 1981(a); Moran *et al.*, 1982(b)). Both of these genes are under complex control and the *spoVG* gene is regulated by at least five out of the eight known *spo0* loci (mutations in any *spo0* locus blocks the onset of sporulation). Phenotypically mutants in the *spoVC* and *spoVG* genes block sporulation at a late stage (Stage IV or V) but the exact function of these genes is not known. When the RNA transcripts of these genes were measured, they were found to be actively transcribed at the earliest stages of sporulation; also, a temperature sensitive mutant of *spoVC* has been isolated and it has been demonstrated that the temperature sensitive step occurs several hours before the phenotype is expressed. When the levels of the σ^{37} subunit were measured, they were present in vegetative cells (at 10% of the normal σ^{55} level), early sporulating cells and stationary phase cells (Johnson *et al.*, 1983). This corresponds well with the level of transcription of the *spoVG* gene, which is low in vegetative cells but high in early sporulating cells and stationary phase cells. It is not known if the increase in transcription is due to an increase in the level of σ^{37} subunit or an increased association of the σ^{37} subunit with the RNA polymerase apo-enzyme.

In the sequence as numbered (Fig. 6.3), the σ^{37} promotor may seem to be too far back from the first transcribed base. However, this was chosen as the first A residue after the promotor. There are several G residues which could be utilised as the first transcribed base and give a more normal spacing of the -35 and -10 regions. An A+T rich sequence (9 consecutive bases centred on -55) is found in the *B. circulans APH* gene and a similar but more extensive sequence is found in the *spoVG* gene. The construction of deletions has shown that this region is important for efficient promotor usage (Moran *et al.*, 1981(b)). At present, the evidence indicates that at the earliest stages of sporulation there is active transcription from σ^{37} promotors. The transcription of the *B. circulans APH* gene from

a σ^{37} promotor is consistent with the timing of the onset of high antibiotic resistance (Demain, 1974) and also indicates how the resistance is linked to the developmental cycle. It is tempting to speculate that genes involved in antibiotic biosynthesis may also have developmentally expressed promotors and so extend the parallels between the onset of sporulation and antibiotic biosynthesis from the physiological to the molecular level.

The possibility exists that the *B. circulans APH* gene σ^{37} promotor is interlaced with a normal vegetative σ^{55} promotor (Fig. 6.8). There is a precedent for interlaced promotors in this type of system as both the *spoVG* and *spoVC* σ^{37} promotors are interlaced with σ^{32} promotors (Johnson *et al.*, 1983). The exact function of this arrangement is not known. The production of σ^{32} seems to parallel the production of σ^{37} but at ~ 10% of the σ^{37} level. If the possible σ^{55} promotor can be utilised at low level by the normal σ^{55} RNA polymerase, it is not difficult to construct a model to explain this. When growing in a normal habitat, a fraction of the population will be undergoing sporulation at any time, thus low levels of antibiotic may be present all the time. Low level transcription of the *APH* gene by the vegetative σ^{55} RNA polymerase may produce enough phosphotransferase to protect growing cells from the low level of antibiotic. Active transcription by the σ^{37} RNA polymerase may only be needed when the cell starts to sporulate and becomes committed to antibiotic production. At this point, it should be emphasised that the control sequences have only been identified on the basis of homology with other known sequences. It is assumed that the regulation of development in *B. circulans* is substantially the same as in *B. subtilis*. The nature of the promotors and the transcriptional start points still need to be confirmed by *in vitro* RNA 'run off' experiments with the different purified RNA polymerases.

The observation that the *B. circulans APH* gene is expressed in *S. lividans* 66 can be explained in four ways: (i) transcription is initiated in the vector and continues through into the cloned fragment; (ii) the σ^{55} promotor is utilised; (iii) the σ^{37} promotor is utilised; or (iv) another as yet unrecognised promotor sequence is used. At the moment, there is no evidence to favour any of these

possibilities; however, as *Streptomyces* are also Gram-positive differentiating organisms, it is tempting to speculate that a similar control mechanism to the one seen in *B. circulans* may operate.

McLaughlin *et al.* (1981) have proposed that the ribosome binding sites (Shine-Delgarno sequences) of Gram-positive mRNAs are able to form highly stable complexes with the 3' terminal region of the 16s rRNA (see also Moran *et al.*, 1982), whereas in *E. coli* the potential for base pairing varies considerably. The proposed ribosome binding site for the *B. circulans APH* gene (Fig. 6.10) fits this model well, with a seven base direct match. The putative transcriptional start and stop signals give long leader sequences of 150 bp and 260 bp respectively. Using the computer, these were searched for secondary structures,such as hairpins, but no plausible structures could be found.

In summary, the *APH* gene of *B. circulans* has been cloned, sequenced and found to be expressed in *E. coli* and *S. lividans* 66. A comparison of the deduced protein sequence with those from TN5, TN903 and *S. fradiae* indicates that the phosphotransferases may all have a common origin. An examination of the regulatory sequence of the gene indicates that it is at least partly under the control of a developmentally expressed promotor, which would cause it to be actively transcribed at the onset of sporulation. This casts new light on the way that genes involved in antibiotic metabolism are linked to the developmental cycle. Also, as the gene is expressed in a variety of hosts, it could be useful as a marker for positive selection in the construction of shuttle vectors.

APPENDIX I

Full listing of sequence data for the
B. circulans APH gene (2.7 kb Sal I fragment)

Contig description lines

Column 1 - gel archive file

Column 2 - gel number in contig

Column 3 - position of first base of the gel in the contig

Column 4 - length of gel. A negative number indicates that the gel is used in the opposite orientation

Column 5 - gel number of gels nearest left neighbour

Column 6 - gel number of gels nearest right neighbour

	99	2706	0	66	21
CJH.S1	66	1	271	0	67
CJH.SS6	67	95	205	66	68
CJH.P8	68	161	-301	67	65
CJH.SS1	65	400	-265	68	63
CJH.A14	63	439	-289	65	62
CJH.P16	62	456	298	63	64
CJH.H10	64	527	-145	62	61
CJH.HP13	61	624	-209	64	69
CJH.SS2	69	667	-299	61	58
CJH.HP7	58	833	273	69	57
CJH.HP9	57	934	-193	58	70
CJH.T40	70	935	304	57	59
CJH.SS12	59	962	-129	70	56
CJH.H9	56	1017	149	59	55
CJH.S12B	55	1091	-108	56	71
CJH.T41	71	1112	270	55	46
CJH.S21A	46	1199	-128	71	41
CJH.HP2	41	1320	-79	46	42
CJH.SS3	42	1321	222	41	1
CJH.S7	1	1327	-215	42	2
CJH.H101	2	1407	-64	1	33
CJH.HP11	33	1409	245	2	27
CJH.SS4	27	1504	-270	33	22
CJH.S10A	22	1539	230	27	39
CJH.BS6	39	1549	-225	22	15
CJH.T3A	15	1603	213	39	20
CJH.T11A	20	1607	-23	15	30
CJH.T30	30	1636	184	20	17
CJH.T12	17	1669	146	30	16
CJH.T4	16	1669	147	17	31
CJH.HP3	31	1764	227	16	3
CJH.S3B	3	1770	75	31	5
CJH.S15	5	1774	-71	3	6
CJH.S27B	6	1777	-68	5	40
CJH.BS9	40	1787	272	6	44
CJH.T19	44	1812	118	40	32
CJH.T2	32	1817	274	44	38
CJH.A2	38	1861	-232	32	43
CJH.S13	43	1895	194	38	37
CJH.SS11	37	1897	241	43	45
CJH.T9	45	2013	-208	37	36
CJH.SS17	36	2102	-245	45	34
CJH.HP10	34	2152	-173	36	35
CJH.T36	35	2219	139	34	14
CJH.A52	14	2230	-262	35	10
CJH.S2	10	2365	-314	14	12
CJH.S12A	12	2384	111	10	18
CJH.T11B	18	2431	-54	12	19
CJH.T3B	19	2440	-45	18	13
CJH.S21B	13	2495	41	19	11
CJH.S3A	11	2499	-41	13	21
CJH.BS5	21	2652	-52	11	0

```
          10         20         30         40         50         60
 66 GTCGACGGCATGGGGCGCCGGCGGCTTCTCGCCGATCGAAGCTGCTGACGTGGTGTCGCA
    GTCGACGGCATGGGGCGCCGGCGGCTTCTCGCCGATCGAAGCTGCTGACGTGGTGTCGCA

          70         80         90        100        110        120
 66 GTTCGCGGCCCTGCTGACGCGGGCCCTCGCGCTCGATCCGGCGACGGCCGCTGGCGGACT
 67                                    GATCCGGCGACGGCCGCTGGCGGACT
    GTTCGCGGCCCTGCTGACGCGGGCCCTCGCGCTCGATCCGGCGACGGCCGCTGGCGGACT

         130        140        150        160        170        180
 66 TCACGACATCCCCGGCGACGCGTGGTATGCCGGCGCGTCGGCGGCCGTTAGGCCCGGCTG
 67 TCACGACATCCCCGGCGACGCGTGGTATGCCGGCGCGTC8G8GGCCGTTAGGCCCGGCTG
-68                                        GCGGCCGTTAGGCCCGGCTG
    TCACGACATCCCCGGCGACGCGTGGTATGCCGGCGCGTCGGCGGCCGTTAGGCCCGGCTG

         190        200        210        220        230        240
 66 CGGACGGTTTGAAACCGGGGAGTTCCGTCCCAGAAGGGCCTGACAGGAACAGATGAGCGT
 67 CGGACGGTTTGAAACDGGGGAGTTCCGTCCCAGAAGGGCCTGACAGGAACAGATGAGCGT
-68 CGGACGGTTTGAAACCGGGGAGTTCCGTCCCAGAAGGGCCTGACAGGAACAGATGAGCGT
    CGGACGGTTTGAAACCGGGGAGTTCCGTCCCAGAAGGGCCTGACAGGAACAGATGAGCGT

         250        260        270        280        290        300
 66 GATGCTATGCGCGCCGTGAAGCTTGCCGTAT
 67 GATGCTATGCGCGCCGTGAAGCTTGCCGTATCCGGACGGAGGAGAACTCGCTGCGCGTG
-68 GATGCTATGCGCGCC-T-AAGCTTGCCGTATCCGGACGGAGGA-AACTCGCTGCGCGTGC
    GATGCTATGCGCGCCGTGAAGCTTGCCGTATCCGGACGGAGGAGAACTCGCTGCGCGTGC

         310        320        330        340        350        360
-68 CTGGCGCTCTTCTCGGACCGGGGCAAATCTCCGGCTGGACGGCCGCTGCCGTTTCGCAGG
    CTGGCGCTCTTCTCGGACCGGGGCAAATCTCCGGCTGGACGGCCGCTGCCGTTTCGCAGG

         370        380        390        400        410        420
-68 CGGTGGAAGCCGGCCTGCTGGACGGACGCGGGGACGGACTTCGCGCCCGGCG8ACCGGCT
-65                                      TTCGCGCCCGGCGGACCGGCT
    CGGTGGAAGCCGGCCTGCTGGACGGACGCGGGGACGGACTTCGCGCCCGGCGGACCGGCT

         430        440        450        460        470        480
-68 ACCCTGCGGAAGGCGCGGCCGTGCTGAAACGGCTGCTGCAG
-65 ACCCTGCHGAAHGCGCGGCCGTGCTGAAACGGCTGCTGCAGACAGCCGGTTTCGTGAACT
-63                   CCGTGCTGAAACGGCTGCTGCAGACAGCCGGTTTCGTGAACT
 62                                    CTGCAGACAGCCGGTTTCGTGAACT
    ACCCTGCGGAAGGCGCGGCCGTGCTGAAACGGCTGCTGCAGACAGCCGGTTTCGTGAACT
```

```
          490        500        510        520        530        540
-65  AG8GCATCTGCGCATTCACGGCTCCGCGGCACGCGCGCCGGTTTAAACCCGGAAAGGCGC
-63  AGCGCATCTGCGCATTCACGGCTCCGCGGCACGCGCGCCGGTTTAAACCCGGAAAGGCGC
 62  AGCGCATCTGCGCATTCACGGCTCCGCGGCACGCGCGCCGGTTTAAACCCGGAAAGGCGC
-64                                                  ACCCGGAAAGGCGC
     AGCGCATCTGCGCATTCACGGCTCCGCGGCACGCGCGCCGGTTTAAACCCGGAAAGGCGC

          550        560        570        580        590        600
-65  ATACCACTCAGGTATGCGCCTTTTACGATATTGAGCCGGTCGCGAATTTCGTCCGGTTTG
-63  ATACCACTCAGGTATGCGCCTTTTACGATATTGAGCCGGTCGCGAATTTCGTCCGGTTTG
 62  ATACCACTCAGGTATGCGCCTTTTACGATATTGAGCCGGTCGCGAATTTCGTCCGGTTTG
-64  ATACCACTCAGGTAT---CCTTTTACGATATTGAGCCGGTCGCGAATTTCGTCCGGTTT8
     ATACCACTCAGGTATGCGCCTTTTACGATATTGAGCCGGTCGCGAATTTCGTCCGGTTTG

          610        620        630        640        650        660
-65  CGCCGCTTTTTTACGAAGATTTGCTTAGAACGCGCCGGAGACCCGATTTCCTTCTCCAAT
-63  CGCCGCTTTTTTACGAAGATTTGCTTAGAACGCGCCGGAGACCCGATTTCCTTCTCCAAT
 62  CGCCGCTTTTTTACGAAGATTTGCTTAGAACGCGCCGGAGACCCGATTTCCTTCTCCAAT
-64  CGC8GCTTTTTTACGAAGATTTGCTTAGAAC4CGCCGGAGACCCGATTTCCTTCTCCAAT
-61                          CTTAGAAC-CGCCGGAGACCCGATTTCCTTCTCCAAT
     CGCCGCTTTTTTACGAAGATTTGCTTAGAACGCGCCGGAGACCCGATTTCCTTCTCCAAT

          670        680        690        700        710        720
-65  GATC
-63  GATCCAAAAGGCCGTCTCGGCGGCTATTCGGGTTCCATCATTTGCTTCCGGAATTAGAAT
 62  GATCDAAAAGGCCGTCTCGGCGGCTATTCGGGTTCCATCATTTGCTTCCGGAATTAGAAT
-64  GATCCAAAAGG
-61  GATCCAAAAGGCCGTCTCGGCGGCTATTCGGGTTCCATCATTTGCTTCCGGAATTA-AAT
-69        AAAGGCCGTCTCGGCGGCTATTCGGGTTCCATCATTTGCTTCCGGAATTAGAAT
     GATCCAAAAGGCCGTCTCGGCGGCTATTCGGGTTCCATCATTTGCTTCCGGAATTAGAAT

          730        740        750        760        770        780
-63  TTGATTA
 62  TTGATTATAACTTTGCGATAGGCCGATGGCGGT
-61  TTGATTAT-ACYTTG88ATAGGCCGA-GGCGGTGAGCAGTACACAAAGAAGCAGAATGAC
-69  TTGATTATAACTTTGCGATAGGCCGATGGCGGTGAGCAGTACACAAAGAAGCAGAATGAC
     TTGATTATAACTTTGCGATAGGCCGATGGCGGTGAGCAGTACACAAAGAAGCAGAATGAC

          790        800        810        820        830        840
-61  TACCGTATTCGGGTAACGGCTACCTCTTTCGCTTTCTTCATGACAACTGTCC
-69  TACCGTATTCGGGTAACGGCTACCTCTTTCGCTTTCTTCATGACAACTGTCCGGAACGTG
 58                                                  GGAAC6TG
     TACCGTATTCGGGTAACGGCTACCTCTTTCGCTTTCTTCATGACAACTGTCCGGAACGTG

          850        860        870        880        890        900
-69  GCAACGATGAGGCACAGCGGAGTACAGCAGCAGGGAGACTACTTTCGAAAGGATGGGTTT
 58  GCAACGATGAGGCACAGCGGAGTACAGCAGCAGGGAGACTACTTTCGAAAGGATGGGTTT
     GCAACGATGAGGCACAGCGGAGTACAGCAGCAGGGAGACTACTTTCGAAAGGATGGGTTT
```

```
          910        920        930        940        950        960
-69  CTTAGTATAGAGACAGGAGCAGCCGTCCTAATATAAAGATTTCTAAATCTGTGCTCTGGA
 58  CTTAGTATAGAGACAGGAGCAGCCGTCCTAATATAAAGATTTCTAAATCTGTGCTCTGGA
-57                                    TAAAGATTTCTAAATCTGTGCTCTGGA
 70                                     AAAGATTTCTAAATCTGTGCTCTGGA
     CTTAGTATAGAGACAGGAGCAGCCGTCCTAATATAAAGATTTCTAAATCTGTGCTCTGGA


          970        980        990       1000       1010       1020
-69  AGATC
 58  AGATCGGCCTTAAAAAGGAATTTTGTTCTAACGGCGAATAACAGTAATTTCCGTTTAAAG
-57  AGATCGGCCTTAAAAAGGAATTTTGTTCTAACGGCGAATAACAGTAATTTCCGTTTAAAG
 70  AGATCGGCCTTAAAAAGGAATTTTGTTCTAACGGCGAATAACAGTAATTTCCGTTTAAAG
-59   GATCGGCCTTAAAAAGGAATTTTGTTCTAACGGCGAATAACAGTAATTTCCGTTTAAAG
 56                                                          AAAG
     AGATCGGCCTTAAAAAGGAATTTTGTTCTAACGGCGAATAACAGTAATTTCCGTTTAAAG


        1030       1040       1050       1060       1070       1080
 58  GCGGGCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG
-57  GCGGGCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG
 70  GCGGGCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG
-59  GCGGGCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG
 56  GCG6RCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG
     GCGGGCCTTTTTAAAAGCGGAGAGGATTCTACTGTATAGTTCTTTACTTGGTAATGTTCG


        1090       1100       1110       1120       1130       1140
 58  GTHGAGGATCGGCTATGAATGTACA
-57  GTGGAGGATCGGCTATGAATGTACAA8ATTTTTTTGCAAATCTGCC
 70  GTGGAGGATCGGCTATGAATGTACAAGATTTTTTTGCAAATCTGCCGGAGCTTTACACGG
-59  GTGGAGGATC
 56  GTGGAGGATCGGCTATGAATGTACAAGATTTTTTTGCAAATCTGCCGGAGCTTTACACGG
-55            GGCTATGAATGTACAAGATTTTTTTGCAAATCTGCCGGAGCTTTACACGG
 71                                    TTTTGCAAATCTGCCGGAGCTTTACACGG
     GTGGAGGATCGGCTATGAATGTACAAGATTTTTTTGCAAATCTGCCGGAGCTTTACACGG


        1150       1160       1170       1180       1190       1200
 70  AATGAACGTTACTAAGAAAATTAGGGCCGGAGGACGAATCGGGCCTTGCAGGAGGATCCG
 56  5ATHAACGTTACTAAGAAAATTAGG
-55  AATGAACGTTACTAAGAAAATTAGGGCCGGAGGACGAATCGGGCCTTGCAGGAGGATC
 71  AATGAACGTTACTAAGAAAATTAGGGCCGGAGGACGAATCGGGCCTTGCAGGAGGATCCG
-46                                                          CG
     AATGAACGTTACTAAGAAAATTAGGGCCGGAGGACGAATCGGGCCTTGCAGGAGGATCCG


        1210       1220       1230       1240       1250       1260
 70  TATTAAGCCGAATTTTAAGGAGGAACTGACAAATGAAC
 71  TATTAAGCCGAATTTTAAGGAGGAACTGACAAATGAACGAAAGTACGCGTAATTGGCCGG
-46  TATTAAGCCGAATTTTAAGGAGGAACTGACAAATGAACGAAAGTACGCGTAATTGGCCGG
     TATTAAGCCGAATTTTAAGGAGGAACTGACAAATGAACGAAAGTACGCGTAATTGGCCGG
```

```
          1270       1280       1290       1300       1310       1320
 71  AGGAACTTCTTGAGCTTCTCGGGCAGACGGAACTAACCGTCAACAAAATCGGATATTCCG
-46  AGGAACTTCTTGAGCTTCTCGGGCAGACGGAACTAACCGTCAACAAAATCGGATATTCCG
-41                                                             G
     AGGAACTTCTTGAGCTTCTCGGGCAGACGGAACTAACCGTCAACAAAATCGGATATTCCG


          1330       1340       1350       1360       1370       1380
 71  GAGATCACGTCTATCACGTGAAAGAGTACAGGGGCACCCCCGCATTTCTGAAAATCGCCC
-46  GAGATC
-41  GAGATCACGTCTATCACGTGAAAGAGTACAGGGGCACCCCCGCATTTCTGAAAATCGCCC
 42  GAGATCACGTCTATCACGTGAAAGAGTACAGGGGCACCCCCGCATTTCTGAAAATCGCCC
 -1        ACGTCTATCACGTGAAAGAGTACAGGGGCACCCCCGCATTTCTGAAAATCGCCC
     GAGATCACGTCTATCACGTGAAAGAGTACAGGGGCACCCCCGCATTTCTGAAAATCGCCC


          1390       1400       1410       1420       1430       1440
 71  C
-41  CCAGTGTATGGTGGAGAA
 42  CCAGTGTATGGTGGAGAACGCTCCGGCCCGAAATTGAAGCGCTCGCTTGGCTGGACGGGA
 -1  CCAGTGTATGGTGGAGAACGCTCCGGCCCGAAATTGAAGCGCTCGCTTGGCTGGACGGGA
 -2                              CCCGAAATTGAAGCGCTCGCTTGGCTGGACGGGA
 33                                CGAAATTGAAGCGCTCGCTTGGCTGGACGGGA
     CCAGTGTATGGTGGAGAACGCTCCGGCCCGAAATTGAAGCGCTCGCTTGGCTGGACGGGA


          1450       1460       1470       1480       1490       1500
 42  AGCTCCCGGTTCCCAAAATTTTGTACACGGCCGAACACGGCGGGATGGACTACTTGCTGA
 -1  AG5TCCCGGTTCCCAAAATTTTGTACACGGCCGAACACGGCGGGATGGACTACTTGCTGA
 -2  AGCTCCCGGTTCCCAAAATTTTGTACACGG
 33  AGCTCCCGGTTCCC5AAATTTTGTACACGGCCGAACACGGCGG8ATGGACTACTTGCT-A
     AGCTCCCGGTTCCCAAAATTTTGTACACGGCCGAACACGGCGGGATGGACTACTTGCTGA


          1510       1520       1530       1540       1550       1560
 42  TG8AGGCGCTAGGCGGAAAAGACGGCTCCCACGAAACGATCC
 -1  TGGAGGCGCTAGGCGGAAAAGACGGCTCCCACGAAACGATC
 33  TGGAGGCGCTAGGCGGAAAAGACGGCTCCCACGAAACGATCCA88CGAACCGG55ACTGT
-27     AGGCGCTAGGCGGAAAAGACGGCTCCCACGAAACGATCCAGGCGAACCGGAAACTGT
 22                                              ATCCAGGCGAACCGGAAACTGT
-39                                                    ACCGGAAACTGT
     TGGAGGCGCTAGGCGGAAAAGACGGCTCCCACGAAACGATCCAGGCGAACCGGAAACTGT


          1570       1580       1590       1600       1610       1620
 33  TTGTGAAGCTGTACGCGGAAGGGCTCCGAAGCGTGCATGGCCTCGATATCCGCGAATGTC
-27  TTGTGAAGCTGTACGCGGAAGGGCTCCGAAG8GTGC77GGCCTCGATATCCGCGAATGTC
 22  TTGTGAAGCTGTACGCGGAAGGGCTCCGAAGCGTGCATGGCCTCGATATCCGCGAATGTC
-39  TTG78AAGCT8TACGCGGAAGGGCTCCG5AGC-TGCATGGCCTCGATATCCGCGAATGTC
 15                                              TCGATBTCCGCGAATGTC
-20                                                 TATCCGCGAATGTC
     TTGTGAAGCTGTACGCGGAAGGGCTCCGAAGCGTGCATGGCCTCGATATCCGCGAATGTC
```

```
        1630       1640       1650       1660       1670       1680
 33 CGCTGTCGAACGGGCTGGAGAAGAAGCTCCGGG
-27 CGCTGTCGAACGGGCTGGAGAAGAAGCTCCGGGATGCGAAAAGAATAGTCGATGAGAGCC
 22 CGCTGTCGAACGGGCTGGAGAAGAAGCTCCGGGATGCGAAAAGAATAGTCGATGAGAGCC
-39 CGCTGTCGAA-GGGCTG8AGAAGAAGCTCCRG-ATGCG7AAAGAATAGTCGATGAGAGCC
 15 CGCTGTCGAACGGGCTGGAGAAGAAGDTCCGGGATGCGAAAAGAATAGTCGATGAGAGCC
-20 CGCTGTCGA
 30                TGGAGAAGAAGCTCCGGGATGCGAAAAGAATAGTCGATGAGAGCC
 17                                                    TCGATGAGAGCC
 16                                                    TCGATGAGAGCC
    CGCTGTCGAACGGGCTGGAGAAGAAGCTCCGGGATGCGAAAAGAATAGTCGATGAGAGCC


        1690       1700       1710       1720       1730       1740
-27 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC
 22 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGG-
-39 TGGTGGACCCGGCCGA-ATAAAARAGGAGTAC-ATT5CAC8YCGGAGGAATT8TACGGGC
 15 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC
 30 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC
 17 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC
 16 TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC
    TGGTGGACCCGGCCGATATAAAAGAGGAGTACGATTGCACGCCGGAGGAATTGTACGGGC


        1750       1760       1770       1780       1790       1800
-27 TATTGCTTGAGAGTAAGCCGGTAACCGAAGATC
 22 TATTGCTTGAGAGTAAGCCGGTAACCGA
-39 TATT-C77777A-TAAGCCGGTAACCGAAGATC
 15 TATTGCTTGAGAGTAAGCCGGTAACCGAAGATCTGGTTTTTGCGCACGGAGATTACTGTG
 30 TATTGCTTGAGAGTAAGCCGGTAACCGAAGATCTGGTTTTTGCGCA1GGAGATTACTGTG
 17 TATTGCTTGAGAGTAAGCCGGTAACCGAAGATCTGGTTTTTGCGCACGGAGATTACTGTG
 16 TATTGCTTGAGAGTAAGCCGGTAACCGAAGATCTGGTTTTTGCGCAYGGAGATTACTGTG
 31                          CCGAAGATCTGGTTTTTGCGCACGGAGATTACTGTG
  3                                GATCTGGTTTTTGCGCACGGAGATTACTGTG
 -5                                    TGGTTTTTGCGCACGGAGATTACTGTG
 -6                                       TTTTTGCGCACGGAGATTACTGTG
 40                                                    CGGAGATTACTGTG
    TATTGCTTGAGAGTAAGCCGGTAACCGAAGATCTGGTTTTTGCGCACGGAGATTACTGTG


        1810       1820       1830       1840       1850       1860
 15 CTCCGAATCTGATTA
 30 CTCCGAATCTGATTATCGA
 17 CTCCGAATCTGATT
 16 CTCCGAATCTGATTA
 31 CTCCGAATCTGATTATCGACGGTGAGAAGCTGTCGGG3TTTATCGATCTCGGACGTGCGG
  3 CTCCGAATCTGATTATCGACGGTGAGAAGCTGTCGGGATTTATC
 -5 CTCCGAATCTGATTATCGACGGTGAGAAGCTGTCGGGATTTATC
 -6 CTCCGAATCTHATTATCGACGGTGAGAAGCTGTCGGGATTTATC
 40 CTCCGAATCTGATTATCGACGGTGAGAAGCTGTCGGGA7TTATCGATCTCGGACGTGCGG
 44             ATTATCGACGGTGAGAAGCTGTCGGGATTTATCGATCTCGGACGTGCGG
 32                  CGACGRTGAGAAGCTGTCGGGATTTATCGATCTCGGACGTGCGG
    CTCCGAATCTGATTATCGACGGTGAGAAGCTGTCGGGATTTATCGATCTCGGACGTGCGG
```

```
          1870       1880       1890       1900       1910       1920
 31  GCGTGGCGGACCGTTATCAGGACATCAGCCTGGCGATCCGCAGCCTCCGGCACGATTACG
 40  GCGTGGCGGACCGTTATCAGGACATCAGCCTGGCG77CCGCAGCCTCC8-8ACGATTACG
 44  GCGTGGCGGACCGTTATCAGGACATCAGCCTGGCGATCCGCAGCCTCCGGDACGATTACG
 32  GCGTGGCGGACCGTTATCAGGACATCAGCCTGGCGATCCGCAGCCTCCGGCACGATTACG
-38  GC-TGGCGGACCGTTATCAGGACATCAGCCT-HCGATCCGCAGCCTCCGGCACGATTACG
 43                                      GATCCGCAGCCT5CG-8ACGATTA8G
 37                                        TCCGCAGCCTCCGGCACGATTACG
     GCGTGGCGGACCGTTATCAGGACATCAGCCTGGCGATCCGCAGCCTCCGGCACGATTACG


          1930       1940       1950       1960       1970       1980
 31  GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGACGGGCTTGACGAGG
 40  GCGACGACCGCTA8AAAGCGCTCTTCCTGGAACTTTACGGGC2GGACGGGCTTGACGAGG
 44  GCGACGACC
 32  GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGACGGGCTTGACGAGG
-38  GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGACGGGCTTGACGAGG
 43  GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGACGGGCTTGACGAGG
 37  GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGYYGGGCTTGBCGAGG
     GCGACGACCGCTACAAAGCGCTCTTCCTGGAACTTTACGGGCTGGACGGGCTTGACGAGG


          1990       2000       2010       2020       2030       2040
 31  ACAAGGTCCG
 40  ACAAGGTCCGGTATTACATCCGGCTGGATGAATTTTTTTGACGAAACGGAAGGAGACAGC
 32  ACAAGGTCCGGTATTACATCCGGCTGGATGAATTTTTTTGACGAAADGGAAGGAGACAGC
-38  ACAAGGTCCGGTATTACATCCGGCTGGATGAATTTTTTTGACGAAACGGAAGGAGACAG8
 43  ACAAGGTCCGGTATTACA6CCGG8TGGATGAATTTTTTTGACGAAACGGAAGGAGACAGC
 37  ACAAGGTCC-HTATTACATCCGGCTGGATGAATTTTTTTGACGAAADGGAAGGAGACAGC
-45                                      TTTTTTTGACGAAACGGAAGGAGACAHC
     ACAAGGTCCGGTATTACATCCGGCTGGATGAATTTTTTTGACGAAACGGAAGGAGACAGC


          2050       2060       2070       2080       2090       2100
 40  CGAAT4TTTAAAGAGCTG
 32  CGAATGTTTAAAGAGCTGCAAACCATGGGAGTAAAACCGCAGCGGATCGA
-38  CGAATGTTTAAAGAGCTGCAAACCATGGGAGTAAAACCGCAGCGGATCGAAA
 43  CGARTGTTTAAAGAGCTGCRARCCATGGGAGTAAAACCGCAGCGGATC
 37  CGAATGTTTAAAGAGCTGCAAA-DATGGGAGTAAAA-CGCAGCGGATCGAAATCGGATTC
-45  CGAATGTTTAAAGAGCTGCAAACCATGGGAGTAAAACCGCAGCGGATCGAAATCGGATTC
     CGAATGTTTAAAGAGCTGCAAACCATGGGAGTAAAACCGCAGCGGATCGAAATCGGATTC


          2110       2120       2130       2140       2150       2160
 37  CGTTAAGAAAGGGGGACAAA-DGTGCTCTATCATTTC
-45  CGTTAAGAAAGGGGGACAAACCGTGCTCTATCATTTCAGCGAAAACCCCGGCATCAAGAC
-36   GTTAAGAAA-HGGGACAAACCGTGCTCTATCAT7TCAGCGAAAACCCCGGCATCAAG55
-34                                                    CATCAAGAC
     CGTTAAGAAAGGGGGACAAACCGTGCTCTATCATTTCAGCGAAAACCCCGGCATCAAGAC
```

```
          2170       2180       2190       2200       2210       2220
-45  GTTCGTCCCCCGAAAA-DGCCCGGTCGTGGCGGCTTGCCCCCTGTCGTCTGGGCGGTCGA
-36  GTTCGTCCCCCGAAAACCGCCCGGTCGTGGCGGCTTGCCCCCT-TCGTCTGGGCGGTCGA
-34  GTTCGTCCCCCGAAAACCGCCCGGTCGTGGCGGCTTGCCCCCTGTCGTCTGHGCGGTCGA
 35                                                              GA
     GTTCGTCCCCCGAAAACCGCCCGGTCGTGGCGGCTTGCCCCCTGTCGTCTGGGCGGTCGA


          2230       2240       2250       2260       2270       2280
-36  TGAA8888ATGCGCATCTGTATTATTTCC8CGGGATTGCCCGCGCATCGTCTACGCTCTT
-34  TGAAGCGCATGCGCATCTGTATTATTTCC8CGGGATTGCCCGCGCATCGTCTACGCTCTT
 35  TGAAGCGCATGCG-ATCTGTATTATTTCCGCGGGATTGCCCGCGCATCGTCTACGCTCTT
-14            TGCGCATCTGTATTATTTCCGCGGGATTGCCCGCGCATCGTCTACGCTCTT
     TGAAGCGCATGCGCATCTGTATTATTTCCGCGGGATTGCCCGCGCATCGTCTACGCTCTT


          2290       2300       2310       2320       2330       2340
-36  TCGGACGGAGTCTCCHACGCGGATAAGCGGCTGCTGTTCGGACCGGGCGACTCGGGTACC
-34  TCGGACGGAGTCTCCGACGCGGATAAGCGGCTGCTGTTCGGACC
 35  TCGGACGGAGTCTCCGACGCGGATAAGCGGCTGCTGTTCGGACCGGGCGACTCGGGTACC
-14  TCGGACGGAGTCTCCGACGCGGATAAGCGGCTGCTGTTCGGACCGGGCGACTCGGGTACC
     TCGGACGGAGTCTCCGACGCGGATAAGCGGCTGCTGTTCGGACCGGGCGACTCGGGTACC


          2350       2360       2370       2380       2390       2400
-36  GTGATC
 35  GTGATCGCGATAGAGCG
-14  GTGATCGCGATAGAGCGCCTGGCTGGAGCGGCTGAGAGAAACGGATCTGTACCGGTATAC
-10                           GGAGCGGCTGAGAGAAACGGATCTGTACCGGTATAC
 12                                                   GATCTGTACCGGTATAC
     GTGATCGCGATAGAGCGCCTGGCTGGAGCGGCTGAGAGAAACGGATCTGTACCGGTATAC


          2410       2420       2430       2440       2450       2460
-14  CTTCGATGAGCGGGACTTCCGGCTGCTCGATGCGACAGCGGGCTACTATACGGCCGAGCG
-10  CTTCGATGAGCGGGACTTCCGGCTGCTCGATGCGACAGCGGGCTACTATACGGCCGAGCG
 12  CTTCGATGAGCGGGACTTCCGGCTGCTCGATGCGACA88GGGCTACTATACGGCCGAGCG
-18                                  TGCGACAGCGGGCTACTATACGGCCGAGCG
-19                                              GGGCTACTATACGGCCGAGCG
     CTTCGATGAGCGGGACTTCCGGCTGCTCGATGCGACAGCGGGCTACTATACGGCCGAGCG


          2470       2480       2490       2500       2510       2520
-14  GCAGGTAACGCCCTTATCGGTCGAACGGGAG
-10  GCAGGTAACGCCCTTATCGGTCGAACGGGAGGGGGATCTGCTCGCCAAGCTGGCAGCCAC
 12  GCAGGTAACGCCCTTATCGGTCGAACGGGAGGGG
-18  GCAGGTAACGCCCTTATCGGTCGA
-19  GCAGGTAACGCCCTTATCGGTCGA
 13                                        GATCTGCTCGCCAAGCTGGCAGCCAC
-11                                            TGCTCGCCAAGCTGGCAGCCAC
     GCAGGTAACGCCCTTATCGGTCGAACGGGAGGGGGATCTGCTCGCCAAGCTGGCAGCCAC
```

```
          2530       2540       2550       2560       2570       2580
-10  GGGAGCCGAGCTTCGGATCGTGCCGAATCTTCACCCTTTGCGTGAGGCGCTTCTGGCTTC
 13  GGGAGCCGAGCTTCG
-11  GGGAGCCGAGCTTCGGATC
     GGGAGCCGAGCTTCGGATCGTGCCGAATCTTCACCCTTTGCGTGAGGCGCTTCTGGCTTC


          2590       2600       2610       2620       2630       2640
-10  GTCCGTTTCGCTTTTTCCCGTCATTCGTTTCCGCAATGCACAGCCTCTTAGGATGAACCT
     GTCCGTTTCGCTTTTTCCCGTCATTCGTTTCCGCAATGCACAGCCTCTTAGGATGAACCT


          2650       2660       2670       2680       2690       2700
-10  GCCAGCAGGCTATTTTTTTTCTTTGTATTTGCCTGCAT
-21             ATTTTTTTTCTTTGTATTTGCCTGCATGAGTGATGTCCBTCTTCACGTC
     GCCAGCAGGCTATTTTTTTTCTTTGTATTTGCCTGCATGAGTGATGTCCATCTTCACGTC


          2710       2720       2730       2740       2750       2760
-21  GAC
     GAC
```

REFERENCES

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

Akagawa, H., Okanishi, M. & Umezawa, H. (1975).
J. Gen. Microbiol. 90, 336-346.

Akagawa, H., Okanishi, M. & Umezawa, H. (1979).
J. Antibiotics 32, 610-620.

Atkinson, M.R., Duetscher, M.P., Kornberg, A., Russell, A.F. & Moffatt, J.G. (1969).
Biochemistry 8, 4897-4904.

Avery, O.T., Macheal, C.M. & McCarty, M. (1944).
J. Exp. Med. 79, 137-158.

Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. & Shaller, H. (1982).
Gene 19, 327-336.

Benveniste, R. & Davies, J. (1973).
Proc. Natl. Acad. Sci. USA 80, 2276-2280.

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

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

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

Bibb, M.J., Schottel, J.L. & Cohen, S.N. (1980).
Nature 284, 526-531.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. & Falkow, S. (1977).
Gene 2, 95-113.

Bradford, M.M. (1976).
Anal. Biochem. 72, 248-254.

Britten, R.J., Graham, D.E. & Neufeld, B.R. (1974).
In Methods in Enzymology, Vol. XXIX,
Eds. Grossman, L. & Moldare, K.
Academic Press, New York p.363-418.

Bryan, L.E. & Vanden Elzen, H.M. (1977).
Antimicrob. Ag. Chemother. 12, 163-177.

Chambliss, G.H. (1979).
In Developmental Biology of Prokaryotes
Ed. Parish, J.H.
Blackwell Scientific Publications, Oxford. p.57-72.

Chang, A.C.Y., Nunberg, J.H., Kaufman, R.J., Erlich, H.A., Schunke, R.T. & Cohen, S.N. (1978)
Nature 275, 617-624.

Chater, K.F., Hopwood, D.A., Kieser, T. & Thompson, C.J. (1982).
Current Topics in Microbiol. and Immunol. 96, 69-95.

Chung, S-T. (1982)
Gene 17, 239-246.

Clarke, L. & Carbon, J. (1975)
Proc. Natl. Acad. Sci. USA 72, 4361-4365.

Clewell, D.B. (1972)
J. Bacteriol. 110, 667-676.

Courvalin, P., Weisblum, B. & Davies, J. (1977).
Proc. Natl. Acad. Sci. USA 74, 999-1003.

Courvalin, P., Fiandt, M. & Davies, J. (1978).
In Microbiology 78,
Am. Soc. Microbiol., Washington, D.C. 262-266.

Daniels, D.L. & Blattner, F.R. (1982).
Virology 117, 81-92.

Davies, J. & Smith, D.I. (1978).
Ann. Rev. Microbiol. 32, 469-518.

Davis, R.W., Botstein, D. & Roth, J.R. (1980).
Advanced Bacterial Genetics, p.140-142,
Cold Spring Harbour, New York.

Demain, A.L. (1974).
Ann. N.Y. Acad. Sci. 235, 601-612.

Dowding, J.E. (1979).
FEMS Microbiol. Lett. 6, 95-98.

Dulaney, E.L. (1948).
J. Bacteriol. 56, 305-313.

Faye, G. & Simon, M. (1983).
Cell 32, 77-87.

Ferenczy, L., Kevei, F. & Zsolt, J. (1974).
Nature 248, 793-794.

Fodor, K. & Alföldi, L. (1976).
Proc. Natl. Acad. Sci. USA 73, 2147-2150.

Franklin, T.J. & Snow, G.A. (1981).
In Biochemistry of Antimicrobial Action, Third Edition.
Chapman and Hall, London. p.118-122.

Goldman, P. & Northrop, D. (1976).
Biochem. Biophys. Res. Commun. 69, 230-236.

Greene, P.J., Poonian, M.S., Nussbaum, A.L., Tobias, L., Garfin, D.E.,
Boyer, H.W. & Goodman, H.M. (1975).
J. Mol. Biol. 99, 237-261.

Grunstein, M. & Hogness, D.S. (1975).
Proc. Natl. Acad. Sci. USA 72, 3961-3965.

Haas, M.J. & Dowding, J.E. (1975).
In Methods in Enzymology, Vol. XLIII,
Ed. Hash, J.H.
Academic Press, New York, p.611-628.

Hanson, R.S. (1979).
In Developmental Biology of Prokaryotes,
Ed. Parish, J.H.
Blackwell Scientific Publications, Oxford. p.37-56.

Hohn, B. & Murray, K. (1977).
Proc. Natl. Acad. Sci. USA, 74, 3259-3263.

Holmes, W.M., Platt, T. & Rosenberg, M. (1983).
Cell 32, 1029-1032.

Hopwood, D.A. (1978).
Ann. Rev. of Microbiol. 32, 373-392.

Hopwood, D.A. (1979).
J. Nat. Prod. 42, 596-602.

Hopwood, D.A. (1981).
In Genetics as a tool in microbiology,
(Society for General Microbiology Symposium 31,
Eds. Glover, S.W. & Hopwood, D.A.)
Cambridge University Press, Cambridge. p.187-218.

Hopwood, D.A., Wright, H.M., Bibb, M.J. & Cohen, S.N. (1977).
Nature 268, 171-174.

Hornemann, U. & Hopwood, D.A. (1980).
In Antibiotics, Biosynthesis, vol. IV.
Ed. Corcoran, J.W.
Springer-Verlag, New York. p.123-131.

Howells, J.D., Anderson, L.E., Coffey, G.L., Senos, G.D.,
Underhill, M.A., Vogler, D.L. & Erlich, J. (1972).
Antimicrob. Agents. Chemother. 2, 79-83.

Ish-Horowicz, D. (1982).
In Molecular Cloning,
Maniatis, T., Fritsch, E.F. & Sambrook, J.
Cold Spring Harbour, New York. p.368-369.

Johnson, W.C., Moran, C.P. & Losick, R. (1983).
Nature 302, 800-804.

Kado, C.I. & Liu, S-T. (1981).
J. Bacteriol. 145, 1365-1373.

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

Konvalinkova, V., Roelants, P. & Mergeary, M. (1977).
Biochem. Soc. Trans. 5, 941-943.

Liras, P., Villanuera, J.R. & Martin, J.F. (1977).
J. Gen. Microbiol. 102, 269-277.

Lobban, P.E. & Kaiser, A.D. (1973).
J. Mol. Biol. 78, 453-471.

Losick, R. & Pero, J. (1981).
Cell 25, 582-584.

Malik, V.S. & Reusser, F. (1979).
Plasmid 2, 627-631.

Mandel, M. & Higa, A. (1970).
J. Mol. Biol. 53, 159-162.

Marcina, F.L., Kopecko, D.J., Jones, K.R., Ayers, D.J. &
McCowen, S.M. (1978).
Plasmid 1, 417-420.

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

Maxam, A.M. & Gilbert, W. (1977).
Proc. Natl. Acad. Sci. USA 74, 560-564.

Messing, J. (1982).
In Genetic Engineering, vol.4.
Eds. Setlow, J.K. & Hollaender, A.
Plenum Press, New York. p.19-35.

Miller, J.H. (1972).
Experiments in Molecular Genetics.
Cold Spring Harbour, New York.

Moran, C.P., Lang, N. & Losick, R. (1981(a)).
Nucleic Acids Res. 9, 5979-5990.

Moran, C.P., Lang, N., Banner, C.D.B., Haldenwang, W.G. & Losick, R. (1981(b)). Cell 25, 783-791.

Moran, C.P., Lang, N., Le Grice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J. & Losick, R. (1982(a)). Mol. Gen. Genet. 186, 339-346.

Moran, C.P., Johnson W.C. & Losick, R. (1982(b)). J. Mol. Biol. 162, 709-713.

Morita, J., Nishida, T. & Komano, T. (1977). Agric. Biol. Chem. 41, 567-571.

Murray, N.E. & Murray, K. (1974). Nature 251, 476-481.

McLaughlin, J.R., Murray, C.L. & Rabinowitz, J.C. (1981). J. Biol. Chem. 256, 11283-11291.

Oka, A., Sugisaki, H. & Tanakami, M. (1981). J. Mol. Biol. 147, 217-226.

Okanishi, M. & Umezawa, H. (1978). In Genetics of the Actinomycetales, Eds. Freerksen, E., Tarnok, I. & Thumin, J.H. Gustave Fischer Verlag, New York. p.19-38.

Okanishi, M., Suzuki, K. & Umezawa, H. (1974). J. Gen. Microbiol. 80, 389-400.

Old, R.W. & Primrose, S.B. (1980). Principles of Gene Manipulation. Blackwell Scientific Publications, London.

Pearce, C.J. & Rinehart, K.L. (1981). In Antibiotics IV Biosynthesis. Springer-Verlag, Berlin. p.74-100.

Pontecorro, G., Riddle, P.N. & Hales, A. (1977). Nature 265, 257-258.

Rhodes, P.M., Winskill, N., Friend, E.J. & Warren, M. (1981). J. Gen. Microbiol. 124, 329-338.

Rodgers, W.H., Springer, W. & Young, F.E. (1982). Gene 18, 133-141.

Roozen, K.J., Fenwick,Jr., R.G. & Curtiss, R. (1971). J. Bacteriol. 107, 21-33.

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

Rudd, B.A.M. & Hopwood, D.A. (1979).
J. Gen. Microbiol. 114, 35-43.

Rudd, B.A.M. & Hopwood, D.A. (1980).
J. Gen. Microbiol. 119, 333-340.

Sancar, A., Hack, A.M. & Rupp, W.D. (1979).
J. Bacteriol. 137, 692-693.

Sanger, F. & Coulson, A.R. (1975).
J. Mol. Biol. 94, 441-448.

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

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchinson III, C.A., Slocombe, P.M. & Smith, M. (1977b).
Nature 265, 687-695.

Sanzey, B., Mercereau, O., Ternynck, T. & Kourilsky, P. (1976).
Proc. Natl. Acad. Sci. USA 73, 3394-3397.

Schaeffer, P., Cami, B. & Hotchkiss, R.D. (1976).
Proc. Natl. Acad. Sci. USA 73, 2151-2155.

Schottel, J.A., Bibb, M.J. & Cohen, S.N. (1981).
J. Bacteriol. 146, 360-368.

Schrempf, H. (1982).
J. Bacteriol. 151, 701-707.

Schrempf, H. & Gobel, W. (1979).
In Plasmids of Medical,Environmental and Commercial Importance.
Eds. Timmis, K.N. & Pühler, A.
Elsevier, North-Holland. p.259-268.

Sherratt, D. (1981).
In Genetics as a tool in microbiology.
(Society for General Microbiology, Symposia 31.
Eds. Glover, S.W. & Hopwood, D.A.)
Cambridge University Press, Cambridge. p.35-47.

Smith, M. & Gillam, S. (1981).
In Genetic engineering, vol.3.
Ed. Setlow, J.K. & Hollaender, A.
Plenum Press, New York. p.1-32.

Southern, E.M. (1975).
J. Mol. Biol. 98, 503-517.

Staden, R. (1982)
Nucl. Acids Res. 10, 2951-2961.

Suarez, J.E. & Chater, K.F. (1980(a)).
J. Bacteriol. 142, 8-14.

Suarez, J.E. & Chater, K.F. (1980(b)).
Nature 286, 527-529.

Sutcliffe, J.G. (1979).
Cold Spring Harbour Symposium 43, 77-90.

Taniguchi, T. & Weissmann, C. (1978).
J. Mol. Biol. 118, 533-565.

Thompson, C.J. & Gray, G.S. (1983).
Proc. Natl. Acad. Sci. USA. (In press).

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

Timmis, K.N. (1981).
In Genetics as a tool in mircobiology.
(Society for General Microbiology, Symposium 31,
Eds. Glover, S.W. & Hopwood, D.A.).
Cambridge University Press, Cambridge. p.49-109.

Vogel, H.J. & Bonner, D.M. (1956).
Microbial. Genet. Bull. 13, 43-44.

Yagisawa, M., Huang, T-S.R. & Davies, J. (1978).
J. Antibiotics 31, 809-813.

Zierhut, G., Piepersberg, W. & Böck, A. (1979).
Eur. J. Biochem. 98, 577-583.

Zubay, G., Chambers, D.A. & Cheong, L.C. (1970).
In The Lactose Operon.
Eds. Beckwith, J.R. & Zipser, G.
Cold Spring Harbour, New York. p.375-391.