β-LACTAMASE GENES OF GRAM-NEGATIVE BACTERIA

JOAN IYABO AMIEMENOGHENA CAMPBELL

A Thesis presented for the degree of
Doctor of Philosophy

Department of Molecular Biology
University of Edinburgh
June 1986

2

Declaration

I hereby declare that this thesis has been composed by myself and that except where stated, the work described within it is my own.

Abstract

Two β-lactamase gene sequences encoded by *Ps. aeruginosa* RMS 149 plasmid and *Rps. capsulata* sp 108 were investigated. The genes were located using DNA recombinant techniques and their nucleic acid sequences were determined using the Sanger dideoxy-sequencing technique. The amino acid sequences were identified and compared with other characterised β-lactamases. They are both class A enzymes (Ambler classification).

The pseudomonad plasmid encoded enzyme is expressed constitutively, but its gene sequence has an attenuator sequence - reminiscent of inducible bacterial synthetic operons. It also has three putative loop-forming sequences in the middle of the gene. RNA mapping studies indicate that, the attenuator is read through from an upstream promoter. There is low level initiation from its own promoter and the transcripts sometimes terminate around the second internal stem-loop. The full message is also made. Thus, it is likely, that the pseudomonad gene is normally highly regulated. Its constitutive expression may be as a result of some control mutation.

The rhodopseudomonad enzyme is unlike other characterised Gram-negative class β-lactamases because it is inducible. Gene hybridization experiments suggest that it may be chromosomally encoded in strain sp 108 as well as in the Pen^S strain sp 109. β-lactamase active bands were also observed in Pen^S *Rps. capsulata* St. Louis and *Rps. sphaeroides*. If this is the usual state of affairs in photosynthetic bacteria which are not normally subject to the selective pressures of the presence of β-lactam antibiotics by virtue of their aquatic habitat, the sp 108 strain may also be producing the enzyme in large quantities due to some control mutation.

It is postulated then, that, β-lactamase genes in Gram-negative bacteria may be of two kinds - one that is chromosomal and is highly regulated, and another

which has lost the regulation and over-expresses the enzyme. The latter may be representative of the common plasmid-borne β-lactamase genes.

Dedication

To the memory of my dear father & friend,
Gabriel Opumelu Aidomokhai Campbell

Acknowledgements

I would like to thank my supervisor, Dr. R.P. Ambler for his guidance and support throughout the duration of this project. I am very much indebted to Drs. D. Finnegan, N. Murray, A. Coulson, R. Hayward J. Maule, R. Moore, G. Coupland, A. Robinson, W. Loenen, J. Flemming and M. Daniels, for their friendly help and advice on many of the aspects of this work. Special thanks go to Dr. T. Gibson for patiently taking me through the steps of DNA sequencing.

There is hardly any person in this department whom I have not asked for one help or the other. So, I thank everyone. I would like to give particular mention to I. Scragg, D. Anderluzzi, G. Cowan and R. Saunders who have been such great pals these years.

Immense thanks also go to Yemi Aderinoye and family and to all the members of Crew House (both past and present) who in their own unique way, provided for me, a home from home.

I thank God for the gift of my parents and friends who have loved and encouraged me in all my undertakings.

Finally, I am grateful to the Association of Commonwealth Universities, the African Trust and to the generosity of my brother, Joseph, for financing me for this course.

Abbreviations

Amp	Ampicillin
Bp	Base-pair
BSA	Bovine serum albumin
Cb	Carbenicillin
Cm	Chloramphenicol
COS	Cold osmotic shock
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra-acetic acid
EtBr	Ethidium bromide
Kb	Kilo-base
Kn	Kanamycin
MIC	Minimal inhibitory concentration
ORF	Open reading frame
PBP	Penicillin binding protein
pI	Isoelectric point
PMSF	Phenyl methyl-sulfonyl fluoride
PSE	Penicillin sensitive enzyme
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
Tc	Tetracycline
Tris	Tris(hydroxymethl) aminomethane

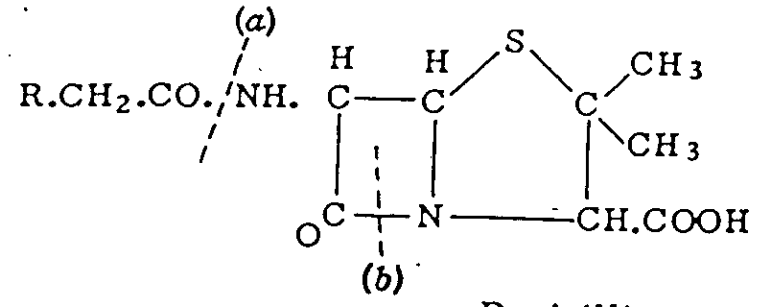
All other abbreviations used in this thesis are as defined in the Biochemical Journal - Instructions to authors.

Abbreviations for bacterial genera

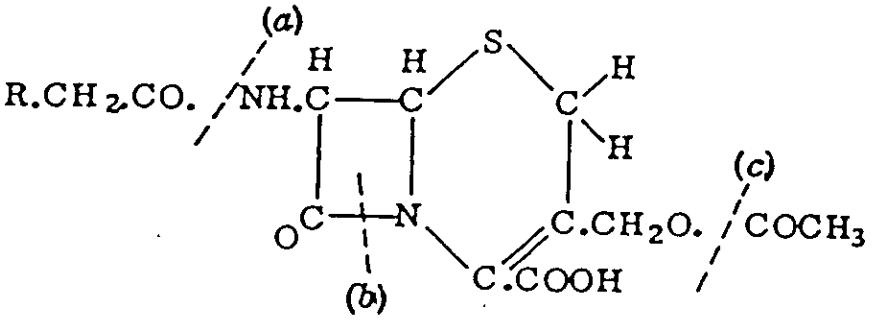
B.	*Bacillus*
E.	*Escherichia*
Ps.	*Pseudomonas*
Rps.	*Rhodopseudomonas*
Staph.	*Staphylococcus*
Strep.	*Streptococcus*

TABLE OF CONTENTS

Fig. 1.1: General formulae of β-lactams



Penicillin



Cephalosporin

Fig. 1.1: Action of enzymes attacking penicillins and cephalosporins. (a) Amidases; (b) β-lactamases; (c) Acyl esterases.

CHAPTER 1

INTRODUCTION

1.1. Introduction

The production of a penicillin destroying enzyme by bacteria was first reported by Abraham and Chain (1940), soon after the introduction of penicillin into clinical medicine. Penicillin inactivating enzymes have been found to be of three kinds: the amidase (E.C.3.5.1.11), which splits off the side chain (R) (Fig.1.1a) (Batchelor *et al*, 1961); penicillinase and cephalosporinase, which break the β-lactam bond (Fig.1.1b); and thirdly, the penicillin esterase, which liberates a penicillanate from Penicillin-3-amide (Fig.1.1c) (Huang *et al*, 1963). The penicillinases and cephalosporinases constitute the β-lactamases (Penicillin amido β-lactam hydrolase, E.C.3.5.2.6; Pollock, 1960).

Though penicillin amidases have been used extensively in the industrial production of semi-synthetic penicillins by attachment of special side chains to aminopenicillanic acid, the clinically important antipenicillin enzymes are β-lactamases. Other intrinsic cellular properties contribute to β-lactam resistance. These include a diffusion barrier (as provided by the Gram-negative bacterial outer membranes), as well as a decrease in affinity of target sites (Spratt, 1978). β-lactamases are however of prime importance because they are the most common cause of high level β-lactam resistance (Coulson, 1985).

The enzymes can probably be found in any bacterial group which is carefully examined, as illustrated in an extensive recent review (Hamilton-Miller and Smith, 1979). β-lactamase activity has also been reported in yeast (Mehta and Nash, 1978) and mammalian tissue (Hamilton-Miller, 1982), but these non-bacterial enzymes have not been characterised in great detail. All the work reported here is therefore concerned with bacterial enzymes.

1.2. Historical development

β-lactamase genes provide a readily selectable marker and the protein is often easily expressed and purified. They have thus served as convenient systems for much molecular biological research. Some recent examples of these are the testing of DNA sequencing methods with pBR322 (Sutcliffe, 1978); the studies of prokaryotic promoters and the control of gene expression (Stuber and Bunjard, 1981; Kreft *et al.* 1983; etc.).

Research directed at understanding the function of β-lactamases themselves started with the work of Pollock on the induction of *Bacillus* enzymes (Pollock, 1950). The only group of organisms whose enzymes posed a clinical problem at this time were the staphylococci (Kirby, 1944). Following the introduction of phenethicillin and methicillin to combat these, it was soon observed that several Gram-negative bacteria could readily hydrolyse the new compounds. Indeed, the gene coding for a β-lactamase in *E. coli* (designated TEM) was found to be on a plasmid (Datta and Kontomichalou, 1965). It was thus predicted that β-lactam resistance could spread to other bacterial species. This prediction was further strengthened by the discovery that the TEM gene was carried on a transposon (Hedges and Jacob, 1974). Using a sensitive assay with a chromogenic substrate and high resolution isoelectric focusing, Matthew and Harris (1976) found β-lactamases in all the bacteria they examined including highly sensitive species such as streptococci. This observation supported the idea earlier put forward by Abraham and Chain (1940) that β-lactamases may have a physiological role other than that of conferring β-lactam resistance.

Work had already begun into the molecular basis of β-lactamase activity. By 1979, four β-lactamase sequence analyses had been achieved by Ambler and co-workers (Ambler, 1975; Ambler and Meadway, 1969; Thatcher, 1975; Ambler & Scott, 1978). The enzymes are produced at high concentrations in certain Gram-positive bacteria (Kogut *et al.* 1956; Citri and Pollock, 1966) where much of

the β-lactamase is excreted into the culture. In Gram-negative bacteria, synthesis is at a lower level and the mature enzymes are located in the periplasmic space. DNA sequence analysis of β-lactamases with known protein sequences, (Sutcliffe, 1978; Lai *et al*,1981; Nielsen *et al*, 1981) led to the elucidation of the signal peptides and hence the mode of secretion.

Subsequent active site modification studies on β-lactamases from Gram-positive (Cohen and Pratt, 1980; Cartwright and Coulson, 1980) and Gram-negative (Fisher *et al*, 1980; Knott-Hunziker *et al*, 1982) bacteria has provided some insight into the mechanism of action of certain β-lactamases. The present active state of β-lactamase-associated research is indicated by the number of recent reviews (Waxman and Strominger, 1983; Waley & Cartwright, 1983; Coulson, 1985; Frere and Joris, 1985).

1.3. Classification of β-lactamases

1.3.1. Variety of β-lactamases

Studies of the physiology, chemistry and genetics of β-lactamases show that there are a large variety of types. The earliest method used to classify β-lactamases involved the use of "substrate profiles", whereby rates of destruction of different β-lactam antibiotics are related to benzylpenicillin or cephaloridine. When grouped like this, the enzymes fall into three categories:- penicillinases (which are more active against compounds having the basic penicillin structure), cephalosporinases (which preferentially hydrolyse those with the β-lactam ring fused to a dihydrothiazine nucleus; O'Callaghan *et al*, 1968) and broad spectrum enzymes. However the several penicillinases and cephalosporinases react with penicillins and cephalosporins at varying rates. Other differences in molecular weights, isoelectric points and susceptibility to inhibitors necessitated more discriminatory methods of classification.

1.3.2. The Richmond and Sykes Classification

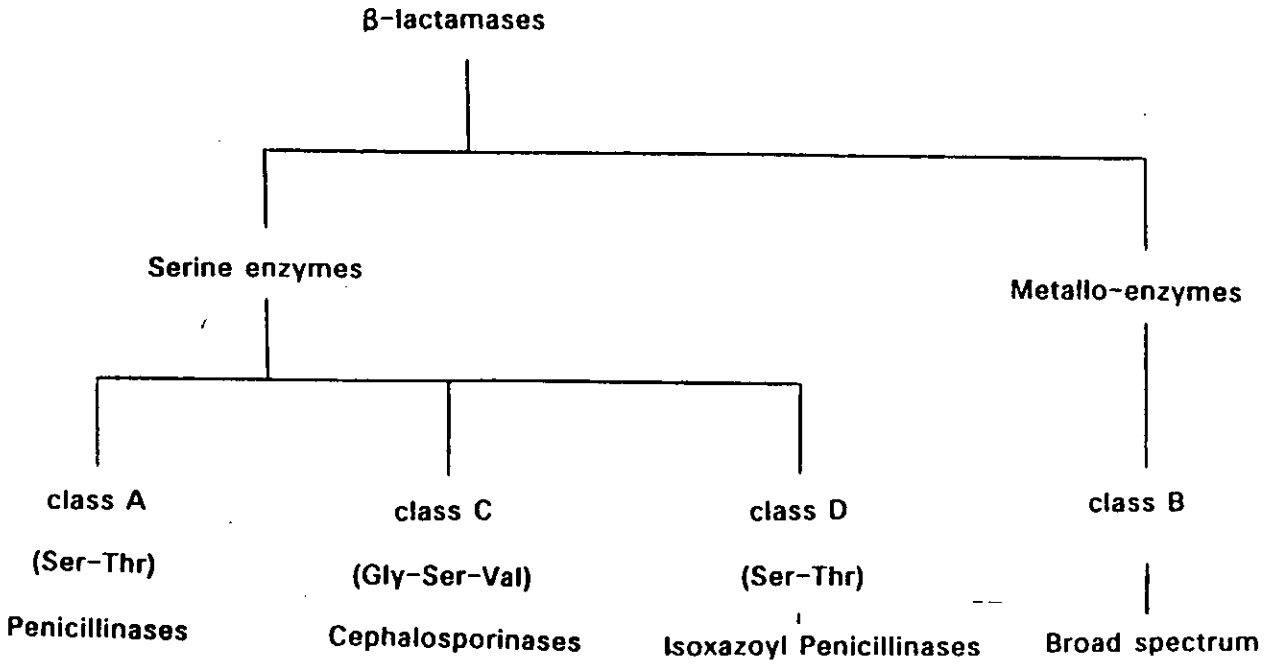
Richmond and Sykes (1973) proposed the first extensive and systematic classification of β-lactamases. The criteria used included substrate profile, molecular weight, susceptibility to various classes of inhibitors, as well as the location of the gene.

The Gram-negative β-lactamases were thus divided into five groups. Class I comprised the cephalosporinases. In class II were placed the chromosomally mediated penicillinases which are resistant to cloxacillin inhibition. The broad spectrum β-lactamases make up classes III and IV, but are distinguished by the cloxacillin sensitivity of the latter class. Class V enzymes are basically penicillinases which differ from class II in being sensitive to sulphydryl agents.

The scheme had a serious disadvantage in that it did not cover all bacterial species. Hence, the staphylococcal and bacillary enzymes, which are similar to the Gram-negative TEM in many respects, could not be easily fitted in. This is because β-lactamase genes are chromosomal in *Bacillus* sp (Nielsen, 1983) but can be on the chromosome as well as on plasmids in *Staph. aureus* (Dyke, 1979). In addition, some of the measurements on which the scheme depends, such as molecular weight, are unreliable, as different values can be reported for the same protein in different laboratories, (e.g. molecular weights of 16,000 (Datta and Richmond, 1966) from ultracentrifugation studies, 21,000 (Dale & Smith, 1971) using gel filtration measurements and 27,000 (Scott, 1972) from SDS gel electrophoresis, have all been reported for the *E. coli* TEM enzyme. Finally, there is evidence that the substrate specificity of β-lactamases can be altered by a small number of amino acid substitutions. Hall and Knowles (1976) described the isolation of point mutants of a penicillinase which had acquired cephalosporinase activity.

However, the main purpose in classifying enzymes is to gain some insight

Fig. 1.2: β-lactamase classes



into the underlying principles of their function and mechanism of action (Coulson, 1985). The best mode of classification is therefore a natural one obtained by comparing their primary structures (Ambler, 1975). However, sequence information takes a long time to gather, so the Richmond and Sykes classification might still serve for preliminary classification.

1.3.3. Ambler Classification

Ambler (1979) started a sequence-based classification for β-lactamases. Not only were the enzymes from Gram-positive and negative bacteria reconciled, but a clearer pattern for the mechanisms of action and evolution of β-lactamases emerged. On the basis of the sequence information available at the time, Ambler (1980) suggested a polyphyletic origin for β-lactamases.

At present, four classes are recognised. The relationship between them is set out in Fig. 1.2. Essentially, they are mainly serine enzymes and a group of metallo-enzymes.

Class A

Class A consists of four main β-lactamases: *Staph. aureus* PC1; *B. licheniformis* 749/C; *B. cereus* 569/H β-lactamase I and *E. coli* TEM. They have molecular weights approximately 29,000, and pI varying from 5.9 to 9.0. They also differ widely in their substrate profiles as well as in other kinetic parameters. However, about 20% of the amino acid sequences are identical in all the proteins and very few insertions or deletions need to be postulated in aligning the sequences. The extent and distribution of the sequence similarities suggest that they have all diverged from a common ancestral gene (Ambler, 1980).

They are all penicillinases. Chemical modification of active site residues by inhibitors has led to the implication of an invariant serine residue. Incubation of class A enzymes with clavulanic acid (a metabolite of *Streptomyces clavuligerus*)

gives rise to an absorption at 280 nm. Judging from the plausible arrangements of bonds and the multiple bond capacity of clavulanate, it was suggested that inhibition was due to the formation of an α,β, unsaturated acyl enzyme (Cartwright & Coulson, 1980; Fisher *et al.*, 1980). Using other derivatives of the inhibitor, an equivalent serine residue was observed to be labelled in both the TEM (Fisher *et al.*, 1980) and *Staph. aureus* enzymes (Cartwright & Coulson, 1980), as well as in the *B. cereus* enzyme (Knott-Hunziker *et al.*, 1979).

Class B

The fully characterised representative of this class is the broad spectrum *B. cereus* β-lactamase II (Hussein *et al.*, 1985). It has a molecular weight of about 25,000 and is unique in its requirement for a metal cofactor Zn^{2+}. The metal binding involves three histidine residues and a solitary cysteine (Baldwin *et al.*, 1979). Sequence information clearly shows that it is different structurally and perhaps in its mechanism of action from the class A group.

Another Zn^{2+} requiring β-lactamase has been reported in *Ps. maltophilia* (Saino *et al.*, 1982). This enzyme differs from the *B. cereus* β-lactamase II in being a penicillinase and a tetramer of subunit molecular weight 31,000. Bicknell *et al.* (1985) determined the sequence of the N-terminal 32 residues and found no recognizable similarity to other β-lactamases. Hence, it is not known if the *Ps. maltophilia* L1 enzyme is homologous to the *B. cereus* Zn^{2+} requiring enzyme.

Class C

The inducible cephalosporinase from *Ps. aeruginosa* (Sabath *et al.*, 1965; McPhail & Furth, 1973) had been purified and shown to have a molecular weight of 40,000, suggesting that it was unlikely to be closely related to either the class A or B enzymes. Jaurin and Grundstrom (1981) sequenced the amp C gene of *E. coli* which encodes a closely related enzyme and found this to be

representative of a new class of β-lactamases. Other class C enzymes have now been characterised from *Enterobacter cloacae* P99 (Charlier *et al*, 1983), *Citrobacter freundii* and *Shigella sonei* (Bergstrom *et al*, 1983). In fact, the use of cloned amp C gene as a probe has revealed that a wide range of Gram-negative bacteria contain homologous genes (Bergstrom *et al*, 1983).

Class C enzymes are more active against cephalosporins and apart from a requirement for an active site serine, they are structurally unrelated to class A enzymes. Protein/β-lactam interaction here also involves the formation of an acyl enzyme. The acyl enzyme situation is better established for class C β-lactamases because, when they interact with benzylpenicillin in aqueous methanol, an intermediate (acyl enzyme) accumulates (Knott-Hunziker *et al*, 1982).

Class D

The newly sequenced oxa-2 enzyme specified by the resistance plasmid R46 (a *Salmonella typhimurium* isolate; Anderson & Datta, 1965), has been assigned to a new class based on its sequence relationship (Dale *et al*, 1985). It further differs from other β-lactamases in its affinity for anthraquinone dyes such as Cibacron blue (Dale & Smith, 1976). It has been suggested that it might have a nucleotide binding domain. Though there is no evidence for nucleotide binding in β-lactamases, it is interesting that the complete sequence shows homology in some regions to rho (a DNA binding protein involved in control of bacterial gene expression) (D. Mossakowska, Pers. Comm.).

Oxa-2 shows a preference for hydrolysing isoxazoyl penicillins. Mechanistic studies have not been done for the class D enzyme. However, it has the ST-K----L sequence, characteristic of the active site residues of classes A and C enzymes. It is thus possible that it may still act via acyl enzyme formation.

The Ambler classification is by no means closed. There are other enzymes

being studied in different laboratories and judging from preliminary reports e.g. 12,000 molecular weight for the *Ps. aeruginosa* R151 specified enzyme (Matthew, 1978); the Zn^{2+} requiring penicillinase from *Ps. maltophilia* (Saino *et al*, 1982); 14,900 molecular weight for a *Streptomyces* sp UCSM-104 β-lactamase (Campos *et al*, 1985); etc., new classes might need to be created.

1.4. Genetics of β-lactamases

Reference has already been made to the variation in the location of β-lactamase genes in section 1.2.1. Class A β-lactamases are chromosomally mediated in *Bacillus* sp, while in *Staph. aureus*, they may be on the chromosome or on plasmids. The closely related TEM enzyme is often plasmid borne in Gram-negative bacteria. For example, of 363 plasmid encoded penicillin resistant isolates characterised by isoelectricfocusing, 62% were the TEM type (Matthew *et al*, 1979). There is no clear evidence yet of a chromosomal TEM type gene, although the known properties of R factors suggest that integration should readily occur (Meynell *et al*, 1968; Novick, 1969).

Consequently, in contrast to the evolutionary relatedness, there seems to be no set pattern for the genetic organisation of class A β-lactamase genes. The flexible arrangement has been attributed to interspecies transfer of the plasmid borne genes. Evidence in support of this hypothesis comes from the location of the TEM gene to a transposable element (Richmond *et al*, 1980). This situation is not peculiar to class A enzymes. Most class C genes so far studied are chromosomal. None has been found on a transposon, but a plasmid encoded related gene has been reported (Jack & Richmond, 1970).

An interesting revelation from isoelectricfocusing of β-lactamases, was that natural isolates possess more than one kind of enzyme (Matthew & Harris, 1976; Matthew & Hedges, 1976). The presence of multiple β-lactamases had also been observed in *B. cereus* 569 (Pollock, 1956; Nielsen & Lampen, 1983). Here, two

class A genes are specified by the same organism in addition to a class B type. The two class A enzymes are immunologically distinct and, while the protein of type I can be loosely cell-bound or free, the type III is a lipoprotein which is more tightly bound to the membrane. A cell-free form of the type III enzyme is also produced in very small quantities. Nielsen & Lampen (1983) studied the type III form and reported that it differed from the type I enzyme in having a cysteine residue in its leader sequence (to which is attached the lipid moeity).

Most β-lactamases from Gram-negative bacteria are constitutive in that their rate of expression is not affected by the presence of β-lactam antibiotics. Exceptions have been certain class C enzymes from *Enterobacter cloacae* (Hennessey, 1967), *Citrobacter freundii* (Lindberg *et al*, 1986), and *Ps. aeruginosa* "Sabath enzyme" (Sabath *et al*, 1965). Induction of β-lactamase is more common with Gram-positive bacteria (Citri & Pollock, 1966). Inducibility is characterised by a long lag before maximum enzyme synthesis is achieved. In the pseudomonad, induction was only possible with exceptionally high concentrations of penicillin or cephalosporins (Sabath *et al*, 1965).

In summary, some β-lactamase genes are chromosomal in some bacteria but are on plasmids in others. Bacteria have and express more than one type of the genes. Though most enzyme synthesis is constitutive, induction in inducible β-lactamases generally follows a peculiar pattern (characterised by a long lag phase), when compared with normal inducible enzyme systems (see section 1.5.3.).

1.5. Evolution of β-lactamases

1.5.1. Penicillin-bacterial cell interactions

β-lactam antibiotics kill bacteria by interfering with the biosynthesis of the cell wall, an action that usually leads to lysis of the cell. Bacterial cell walls consist of crosslinked polysaccharide chains made up of alternating units of two

Fig. 1.3: Schematic representation of the structure of bacterial cell wall

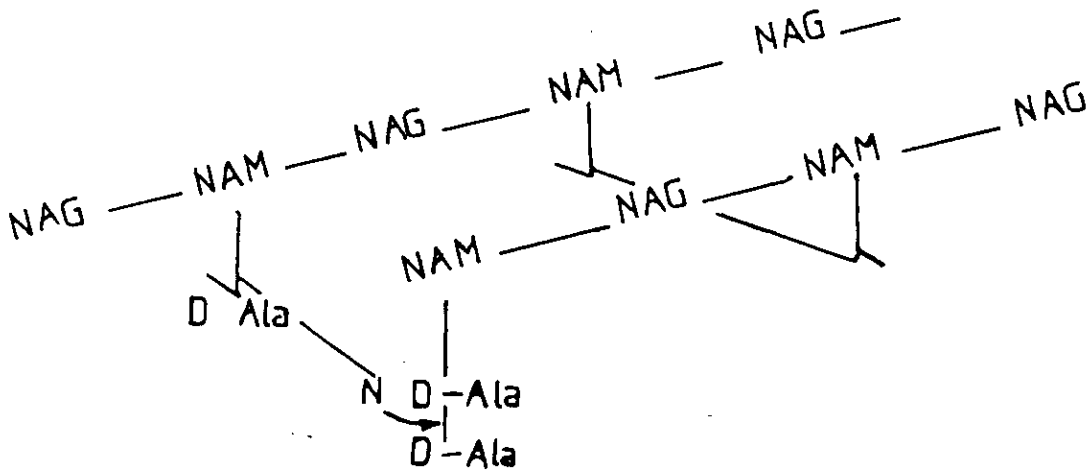


Fig. 1.3: Schematic representation of the structure of the bacterial cell wall. Chains of alternating N-acetyl muramic acid (NAM) and N-acetylglucosamine (NAG) residues are cross-linked by branched chain oligopeptides. The sequence of the oligopeptides differs in different species, but the final cross-link is always made by an acyl transfer from a C-terminal D-Ala-D-Ala bond to an amino group on a neighbouring chain.

amino sugars: N-acetylglucosamine (NAG) and N-acetylmuramic (NAM) acid. The complete 3-dimensional envelope (sacculus) arises from crosslinks of the backbone polysaccharide chains by short peptides borne on the NAM residues (Fig. 1.3.).

Maintenance of cell shape and osmotic balance depends on an intact sacculus. Treatment with lysozyme (N-acetylmuramide glycanohydrolase, which attacks β(1→4) bonds) or other chemical agents which damage the bacterial skeleton, lead to cell leakage and death.

The peptidoglycan sacculus is not static but undergoes extension during cell growth. It is during this process that cells are vulnerable to the effects of penicillins (Lederberg, 1956). Proteins which interact with the antibiotics (penicillin binding proteins, PBP) have been studied. Membrane preparations or whole cells are incubated with a radioactive β-lactam antibiotic. The membranes are solubilized and the proteins are separated by electrophoresis. Labelled proteins are identified by autoradiography. Penicillin sensitive enzymes (PSE) involved in peptidoglycan biosynthesis can be identified by correlating morphological changes in bacteria in the presence of the antibiotic with those produced in specific PBP mutants (Spratt & Pardee, 1975; Spratt, 1975; 1977). Using these techniques, the main PSEs have been identified as transpeptidases and carboxypeptidases (Ghuysen *et al.*, 1979; Spratt *et al.*, 1980).

The natural history of penicillin is complex and poorly understood. They can however be viewed as a weapon devised by soil dwelling fungi against competing bacteria. It has thus been proposed that β-lactams act by mimicking the D-Ala-D-Ala bond involved in the transpeptidase and carboxypeptidase steps (Tipper & Strominger, 1965). The ring strain of the β-lactams makes them good acylating agents, so that it is likely that some of the PBPs become irreversibly acylated (Coulson, 1985).

However, different β-lactams have varying effects on bacterial cells. This has been correlated to the differences in affinity for the many PSEs (Spratt *et al.*, 1980) involved in the different steps of cell wall synthetic reactions. Some penicillin-cell interactions cannot, however, be described as simply inhibiting a particular PSE. In *Staph. aureus*, penicillin G causes the classical accumulation of murein precursors, formation of loose peptidoglycan and subsequent cell death (Tipper & Strominger, 1965). Ampicillin and methicillin in the same organism stimulate an increase in cell wall products released by the cell's autolytic system (Izaki *et al.*, 1968). At the same time, a β-lactam antibiotic added to different bacteria may induce very different physiological effects. *Streptococcus pyogenes* and *Strep. pneumoniae* have about the same minimal inhibitory concentrations. When either strain is treated with benzylpenicillin, *Streptococcus pyogenes* loses viability without any sign of lysis or structural damage to the cell, whereas the pneumococci lose viability and lyse (Tomasz *et al.*, 1970). Analysis of some tolerant pneumococci indicates that resistance in these instances was due to suppression of the cell's autolytic system (Tomasz, 1980).

In conclusion, the Tipper and Strominger hypothesis seems to explain many of the effects of penicillins on bacteria. There are, however, some cases (involving triggering of the cell's autolytic enzymes), for which it may not hold.

1.5.2. Relationship between β-lactamases and PBPs

As β-lactams probably bind to the active sites of PBPs, it is tempting to assume that β-lactamases evolved from cell wall synthetic enzymes. On the basis of active site sequence similarities between carboxypeptidases and class A β-lactamases, it was postulated that the latter enzymes might have evolved from PBPs, having found a way to reverse acylation (Waxman & Strominger, 1980). More sequence information about the whole carboxypeptidase molecule and other PBPs, have shown that, while the active site sequence similarity is a common property (Broome-Smith *et al.*, 1985), homology in other regions is

limited. This supports Ambler's (1980) earlier caution that active site similarity might simply be because they all interact with β-lactam antibiotics. Recent tertiary structural analysis of *B. licheniformis* and *B. cereus* β-lactamases (Kelly *et. al*, 1986; Samraoui *et al*., 1986) suggest extensive homology to *Streptomyces* sp carboxypeptidase folding. This has again been interpreted as evidence of divergence of carboxypeptidases and β-lactamases from the same ancestral gene.

1.5.3. Function of β-lactamases

The *in vitro* reactions of β-lactamases with β-lactam antibiotics show a very high specificity (Pollock, 1960). A striking relationship between degree of resistance to β-lactam antibiotics and β-lactamase activity can be demonstrated in many clinical isolates (Richmond & Sykes, 1973). Complete and permanent genetic loss through deletion of the gene or genes specifying the enzymes does not affect bacterial growth appreciably unless the antibiotics are present in the environment. These observations have led to the proposition that β-lactamases function as antibiotic detoxifiers (Pollock, 1971).

Evidence in support of a β-lactamase function in natural ecological niches was provided by the experiments of Hill (1970). He studied populations of penicillin sensitive and penicillin resistant *Bacillus* sp in competition with penicillin producing fungal isolates in sterile soil. A selective advantage was observed for β-lactamase production. In fact, increase in the use of antibiotics in medicine and agriculture has led to the development of new pathogenic patterns with newly emerged resistance strains (Broda, 1979).

There are, however, some observations which are not in complete agreement with a merely detoxifying role for bacterial β-lactamases. Firstly, the possession of a β-lactamase is not synonymous with β-lactam resistance; e.g. many Gram-negative bacteria have a gene homologous to the *E. coli* amp C and

express it constitutively at low levels (Matthew & Harris, 1976; Normark & Lindberg, 1985). Secondly, induction of the enzymes differs from the classical pathways of inducible enzymes. Whereas, in most inducible enzymes, a very short lag (1-3 min) occurs before enzyme synthesis (Jacob and Monod, 1961), induction of β-lactamase in *B. cereus* 569 is characterised by a long lag period of 5-15 min after addition of inducer and enzyme synthesis continues linearly irrespective of presence or absence of inducer (Citri & Pollock, 1966). A similar situation has been reported for other *Bacillus* sp, *Enterobacter cloacae* (Hennessey, 1967) and *Rps. capsulata* sp 108 (Scahill, 1981) β-lactamases.

These observations have been interpreted as meaning that induction is probably indirect, suggesting that β-lactamases might have an as yet unidentified physiological role in bacteria (Saz & Lowery, 1979). Since β-lactams upset peptidoglycan synthesis (Waxman & Strominger, 1980), the real inducer could well be a peptidoglycan precursor. Induction of β-lactamase synthesis has been described for exogenous low molecular weight peptidoglycan in *B. cereus* (Saz & Lowery, 1979).

There is no doubt that the ability to produce β-lactamase confers a selective advantage on bacteria in both experimental and natural situations. Nevertheless, the diversity in β-lactamase enzymes available does lead to several questions:

- Are β-lactamases long standing enzymes evolved to cope with "poisons" secreted by competing fungi? OR
- Do they represent catabolic enzymes that have been recruited to tackle the problem of β-lactam toxins?

1.6. Scope of this Thesis

In the work reported here, two β-lactamases specified by the *Ps. aeruginosa* RMS 149 plasmid and by *Rps. capsulata* sp 108 were studied. The primary aim was to provide sequence information for further evolutionary analysis of bacterial β-lactamases.

The pseudomonad enzyme was unusual because it has been reported to have a molecular weight of 12,000 daltons (Sawada *et al*, 1974), much smaller than any of those already characterised. It was thus of interest to determine its sequence and find out its relationship (if any) to other β-lactamases. A more detailed account of *Ps. aeruginosa* RMS 149 plasmid and the β-lactamase specified is given in section 3.1.

Most of the research on penicillin-bacterial interactions has concentrated on *Bacillus* spp and pathogenic organisms. The extensive distribution of β-lactamases has already been mentioned in section 1.1. Penicillins and related antibiotics can be expected to be present in physiologically significant levels in soil environments associated with fungi. As a consequence of clinical applications, there are hospital environments where the antibiotic levels are high. It is, however, difficult to imagine that concentrations of antibiotics in water habitats will be sufficiently high to exert a penicillin-selective pressure for β-lactamases. It was considered of interest to analyse the structure of a β-lactamase specified by a photosynthetic bacterium isolated from sewage. This was with a view to studying its evolutionary relationships to clinical β-lactamases.

A more detailed description of the rhodopseudomonad and the β-lactamase is given in section 5.1. In chapters 3, 4, and 5, experiments are described which were aimed at cloning, subcloning and sequencing the DNA encoding the two β-lactamases. Difficulties in isolating and characterising both proteins as produced in their native states, necessitated the indirect method of DNA sequencing.

In the final study reported in chapter 6, several sequence analysis techniques were used to explore the phylogenetic relationships between the two new β-lactamases and other characterised ones.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Chemicals

Acrylamide, N,N'-methylene-bis-acrylamide, sodium dodecyl sulphate (SDS), ethidium bromide, polyethylene glycol-6000 (PEG), ficoll and adenosine 5' -triphosphate were obtained from Sigma Chemical Co. Ltd., London; ethylene diamine tetra-acetic acid (EDTA) and caesium chloride were purchased from Fisons Scientific Apparatus, Loughborough, Leics., England. The deoxy- and dideoxy nucleotides were originally gifts from the Molecular Biology Laboratory in Cambridge and were later obtained from Pharmacia (GB) Ltd., Middlesex.

Dithiothreitol (DTT) and phenol were from Bethesda Research Laboratories (UK) Ltd., Cambridge and the antibiotics- oxacillin, cloxacillin, cephalexin, methicillin, and rifampicin were obtained from Sigma Chemical Co., Ltd. Ampicillin, cephaloridin, benzyl penicillin, chloramphenicol, kanamycin and tetracycline were obtained from the Royal Infirmary, Edinburgh, while carbenicillin and nitrocefin were gifts from Glaxo Group Research Ltd. Synthetic oligonucleotide sequencing primer and the M13 hybridization probe primer were purchased from New England Biolabs, Bishop Stortford, England.

2.1.2. Enzymes

T_4 DNA ligase, (E.C.6.5.11), *E. coli* DNA polymerase I (Klenow large fragment), T_4 polymerase, S1 nuclease (E.C.3.1.30.1) were all bought from Bethesda Research Laboratories (UK) Ltd., Cambridge, while pancreatic RNAse (E.C.3.1.4.22), pancreatic DNAse (E.C.3.14.18) and lysozyme were obtained from Sigma Chemical Co. Ltd., London.

The restriction enzymes were bought from several sources:- NBL enzymes Ltd., Cramlington; New England Biolabs, Bishop Stortford, England; and Boehringer Mannheim, Miles Laboratories, Stoke Poges, England.

2.1.3. Radioactively labelled Compounds

[α-^{32}P] labelled deoxycytosine 5' -triphosphate and [^{35}S]α thio labelled deoxyadenosine 5' -triphosphate were obtained from Amersham International p.l.c., England.

2.1.4. Bacterial strains

2.1.4.1. *E. coli* K12 strains

Strains	Description	Reference/Source
ED 8654	SupE, SupF, hsdR$^-$, M$^+$ S$^+$, Met$^-$, trp^R	Laboratory stock
HB 101	hsdS20, R$^-$, M$^-$, recA13 ara14, proA$_2$, lacY1 galK$_2$, rpsL20 (Sm)^R	Boyer *et al*; 1969; Laboratory stock
TG1	del (lac, pro) supE, thiF' traD36, pro A B lac I ZdM15 hsd5	Gibson, 1984; T.J. Gibson
C_{600}	supE44, tonA21, thr1, thr1, leuB6, thi1, lacY1	Appleyard, 1954; Laboratory stock.
FS 1585	supE, supF, TonA21, thr1, leuB6, thi1, lacY1, recBC	D.R.F. Leach

2.1.4.2. Other Bacterial strains

Species	Strain	Source	Reference
Ps. aeruginosa	1973E (PUE1 with R151)	J. Fleming	Jacoby, 1974
Ps. aeruginosa	1920E (PUE1 with RMS 149)	J. Fleming	Jacoby, 1974
Ps. aeruginosa	PAO5	P. Mulien	
Rps. capsulata	sp 108	J.Fleming	Weaver *et al*, 1975

Rps. capsulata	sp 109	J. Fleming	Weaver *et al.*, 1975
Rps. capsulata	ATCC 23782 (St. Louis)	J. Fleming	Gest, 1974
Rps. sphaeroides	630	J. Fleming	

2.1.5. Plasmids

Plasmid	Genotype	Source/Reference
M13 mp8		Messing & Viera, 1982
M13 mp19		Messing & Viera, 1982
pLG 339	kan^R, Tc^R	Laboratory stock Stoker *et al.*, 1982
pSS20	amp^R (pBR322/BamH1)*	Scahill, 1981
RMS 149	Cb^R, Gm^R,Sm^R, Su^R,	Jacoby, 1977
pRLG 3	kan^R, amp^R (pLG 339/Bam H1)	This thesis
pRLG 300	amp^R (pLG 339/Bam H1-Xho 1)	This thesis
pRLG 301	amp^R (pRLG 300,Sal 1 and Eco R1 del)**	This thesis
pRLG 303	amp^R(pLG 339/[illegible] Eco R1-Xho 1)	This thesis
pRLG 304	amp^R (pLG 339/Cla1-Taq 1)	This thesis
pDR1	amp^R, Tc^R, kan^R (pLG 339/Eco R1)	This thesis
pDR2	amp^R,Tc^R, kan^R (pDR1 with insert in opp. orientation)	This thesis

2.1.6. Phage strains

Phage	Genotype	Source/reference
λcI857	cI857, Sam7	Laboratory stock
λLS20	sb1^0, sb(2-3) ins***, nin 5, red3 amp^R	Scahill, 1981

λLS200	amp^R (λcI857/SalI)	This thesis

* The figures in bracket denote the vector/enzyme site(s) flanking the insert in the recombinant plasmid or phages.

** del=deletion

*** ins=insertion.

2.1.7. Media

The following compositions are per litre of medium.

Luria broth (L-broth; Lennox, 1955)	Bacto-tryptone (Difco), 10 g; yeast extract, 15 g; NaCl, 5 g; D-glucose, 1 g; pH 7.2.
L-agar	L-broth solidified with 15 g/l Difco agar.
Nutrient broth	Oxoid nutrient broth, 25 g; pH 7.2.
R-broth (Sojka *et al*, 1967)	Sodium pyruvate, 1.5 g; sodium hydrogen malate 1.5 g; ammonium chloride, 0.5 g; magnesium sulphate, 0.4 g; calcium chloride, 0.05 g; sodium chloride, 0.4 g; yeast extract, 0.025g. The pH was adjusted to 6.8 with KOH or NaOH before autoclaving and sterile potassium phosphate buffer, pH 6.8 was added to (5mM) after autoclaving.
R-agar	R-broth solidified with 15 g/l of Difco agar.
BBL-broth (Parkinson, 1968)	BBL trypticase (Difco), 10 g; sodium chloride 5 g.
BBL-agar	BBL-broth solidified with 10 g/l agar.
Top agar	BBL-broth solidified with 6.5 g/l agar.

2.1.8. Buffers

The following formulations are per litre of solution.

TE buffer	Tris base,1.21 g; EDTA, 0.37 g; adjusted to pH 8 with 0.6 - 0.8 ml 6 N HCl.
Bacterial buffer	Potassium dihydrogen orthophosphate, 3 g; sodium dihydrogen orthophosphate, 7 g; sodium chloride, 4 g; magnesium sulphate, 0.2 g; pH 7.2.
Phage buffer	Potassium dihydrogen orthophosphate, 3 g; sodium dihydrogen orthophosphate, 7 g; sodium chloride, 5 g; magnesium sulphate, 0.25 g; calcium chloride,

0.015 g; gelatin, 1% (w/v); pH 7.1.

10 x TBE — Tris, 109 g; boric acid, 55 g; EDTA, 93 g; pH 8.3.

20 x SSC — Sodium chloride, 173.3 g; Tri-sodium citrate, 88.2 g; adjusted to pH 7.0 by adding a few drops of 6 N HCl.

Loading buffer — Tris base, 1.21 g; EDTA, 7.5 g; glycerol, 10% (v/v); bromophenol blue, 0.01% (w/v); agarose, 0.2 g; adjust to pH 7.5 with HCl. The mixture is heated to dissolve the agarose and allowed to set. Then the slurry is taken up and forced through a hypodermic needle 5 - 8 times to make a homogeneous slurry.

Universal buffer (for most endonucleases)
Tris, 39.9 g; Magnesium acetate, 21.45 g; potassium acetate, 64.7 g. The solution is adjusted to pH 7.5 with acetic acid and autoclaved. To two ml, are added 10 µl 1 M DTT and 2 mg of Nuclease-free Bovine serum albumin (BRL Ltd., Cambridge).

2.2. Methods

2.2.1. Handling of bacterial cell cultures and general protein methods

2.2.1.1. Maintenance of cultures

All the strains of *E. coli* and *Ps. aeruginosa* used in these investigations were stored on L-agar stab medium and in 25% frozen gelatin solutions. *Rhodopseudomonas* spp were maintained on R-agar stab medium. Strains carrying plasmids were routinely subcultured onto supplemented media for characterization of their phenotype.

2.2.1.2. Growth conditions

i) *E. coli* and *Pseudomonas* spp.

Liquid cultures were grown in 5 ml volumes in 10 ml bottles or in 500 ml volumes in 2 l Erlenmeyer flasks on an orbital shaker incubator. Routinely, *E. coli* and *Pseudomonas* strains were incubated at 37 ^{0}C..

Large scale cultures for β-lactamase preparations were grown in 50 l Biotech fermenters using L-broth. Incubation was at 37 ^{0}C, 5 h for *E. coli* and 18 h for *Ps.*

aeruginosa. The stirrer speed was maintained at 250 - 300 revs/min. Micro-aerophilic conditions were achieved by aerating at 1.5 - 3.5 l of air/min.

The cells were harvested in an Alfa-Laval continous flow centrifuge and yields of 2 - 2.5 g/l (wet weight) were obtained routinely. Harvested cells were stored at -20 ^{0}C until required.

ii) *Rhodopseudomonas* spp

Rhodopseudomonad strains were grown in covered blood bottles filled with 500 ml R-broth, and illuminated with 300W tungsten filament light bulbs. Two lamps were used for every 40 bottles. The cultures were placed 42 -45 cm under the lamps and incubated at room temperature. A perspex screen was used to protect the growth from the intense heat generated by the bulbs.

Typically, 20 - 40 bottles were grown in one batch and the resulting 10 - 20 l cultures were harvested in a continuous flow centrifuge after incubation for seven days. By this time, the cultures were deep red in colour. Approximately, 40 g (wet weight), of cells were recovered from a good growth. The cells were used immediately or stored at -20 ^{0}C until required.

2.2.1.3. Measurement of bacterial growth

Bacterial growth was determined by measuring absorbance of 1 ml culture, or an appropriate dilution, at 650 nm in a Perkin Elmer 320 spectrophotometer. In measurement of enzyme yields, bacterial growth was expressed in terms of mg/dry weight. This was determined by correlating absorbance to its corresponding mg/dry weight value in a predetermined opacity-dry weight curve.

2.2.1.4. Opacity-dry weight curve determination

Standard opacity-mg dry weight curves were determined for *E. coli* ED 8654 and *Ps. aeruginosa* 1920E.

Fig. 2.1: Opacity/mg dry weight curve: *E.coli* ED 8654

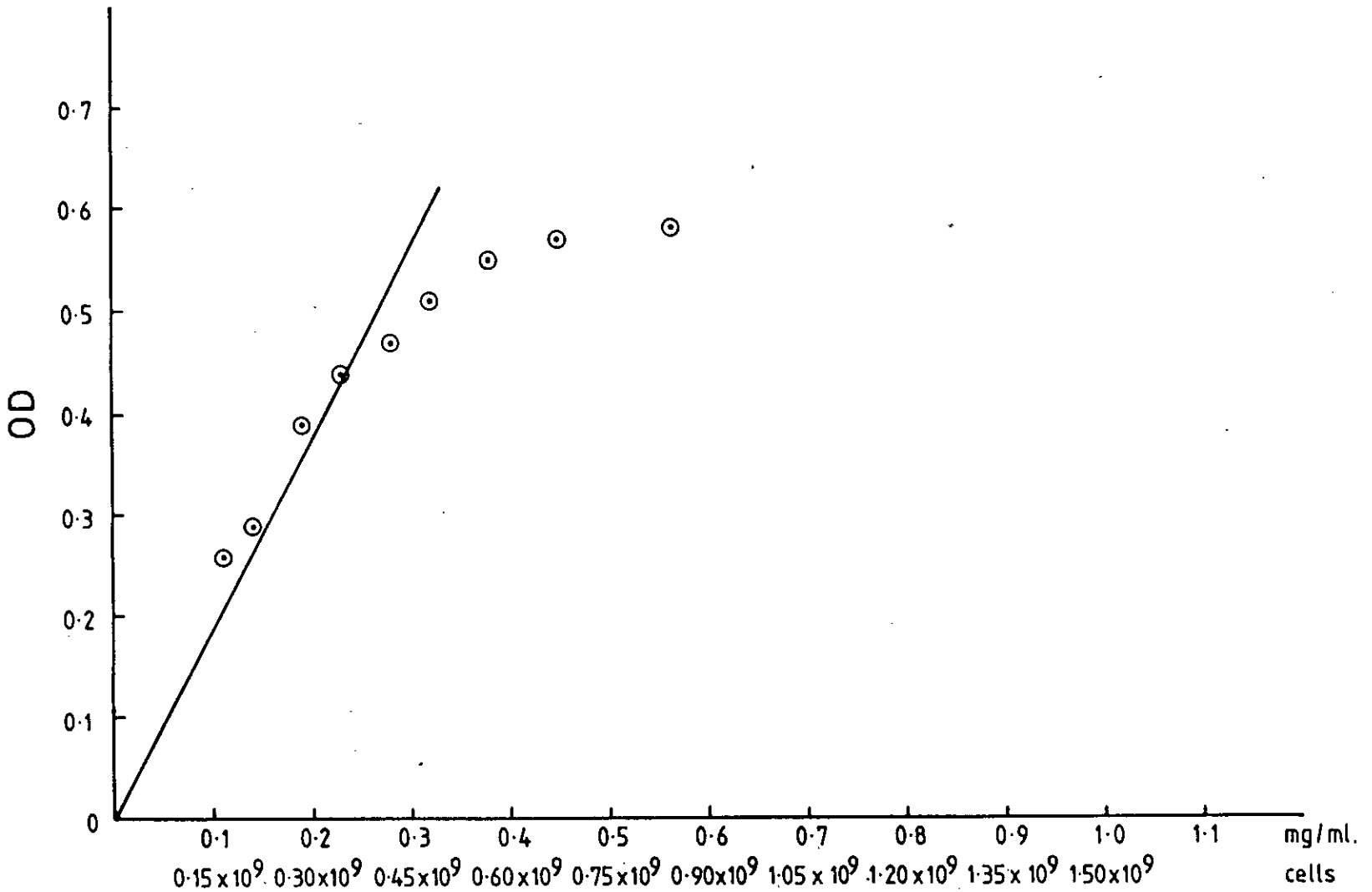
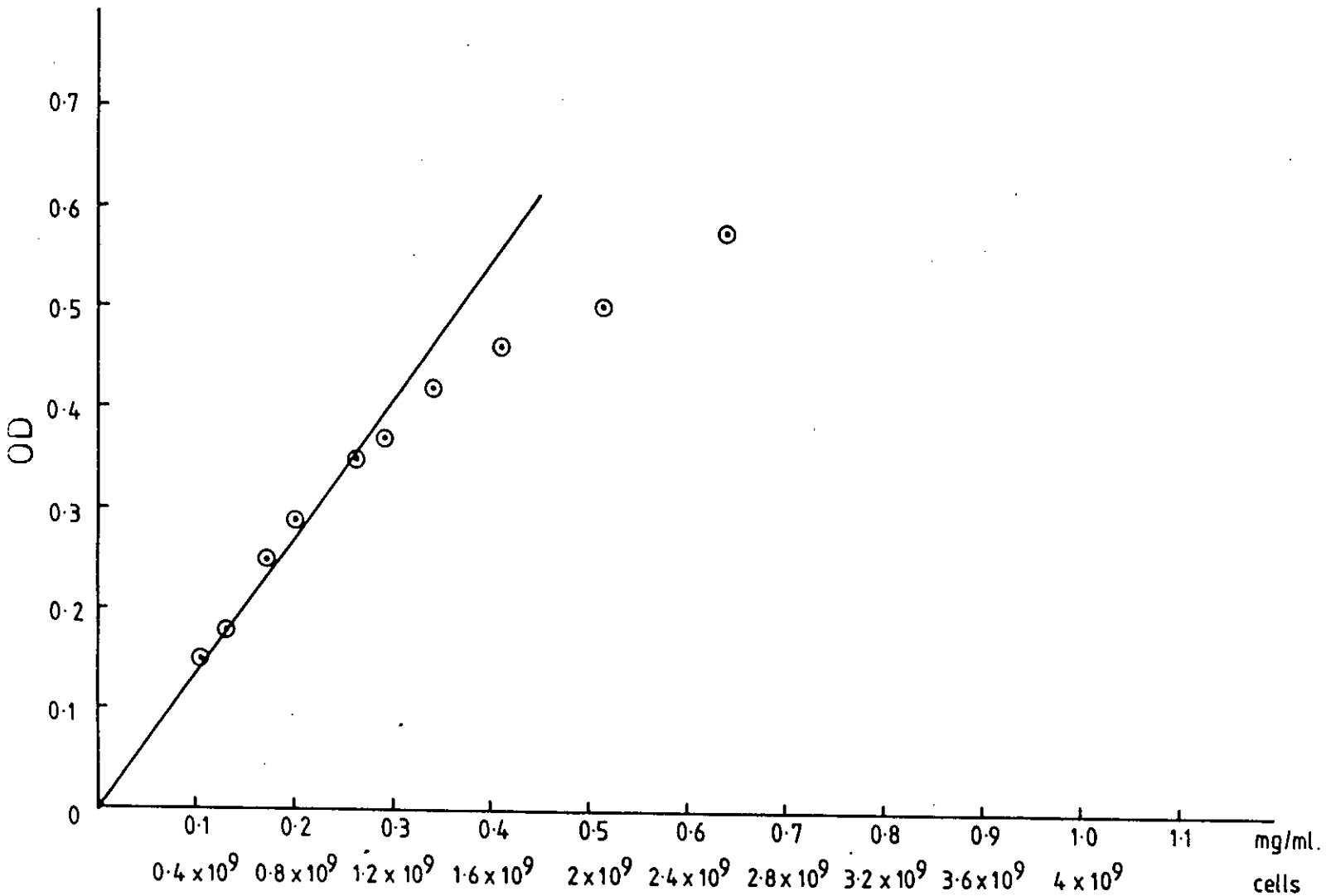


Fig. 2.2: Opacity/mg dry weight curve: *Ps. aeruginosa* 1920E



Nutrient broth cultures of the bacteria (100 ml) were grown overnight. Appropriate dilutions of 1 ml were made and the OD at 650 nm and the viable counts were determined. The remaining 99 ml were spun down and resuspended in 10 ml distilled water. Serial dilutions of the cell suspension were made and the corresponding ODs recorded using a Perkin Elmer spectrophotometer at 650 nm. Portions (3.5 ml) of the original suspension were put in tared beakers, dried in an oven at 70 ^{0}C overnight and weighed after cooling in a desiccator. The beakers were dried again for another 6 h, cooled and weighed. The process of drying and weighing was repeated till a constant weight was achieved in two successive operations.

The mg-dry weight was calculated for the different dilutions. A graph was then drawn by plotting OD against the corresponding mg-dry weight values. The viable count corresponding to a particular OD can also be calculated and correlated on the graph (ref. Figs. 2.1 and 2.2).

2.2.1.5. Application of antibiotics

Antibiotics were applied at the concentrations given below. They were routinely added to the molten agar before pouring plates or directly to broth cultures.

Tetracycline (Tc), 10 µg/ml; ampicillin (Amp), 50 - 100 µg/ml; kanamycin (Kn), 50 µg/ml; carbenicillin (Cb), 1 mg/ml; chloramphenicol (Cm), 50 µg/ml.

2.2.1.6. Determination of Protein concentration (Lowry *et al.*, 1951)

Reagents A (100 ml 2% (w/v) Na_2CO_3 in 0.1 M NaOH + 1 ml 1% (w/v) aqueous NaK tartrate + 1 ml 1% (w/v) aqueous $CuSO_4$. $5H_2O$) and B (10 fold dilution of Folin Ciocalteau's reagent) were made. At time 0, 3 ml of reagent A was mixed with 0.5 ml sample. After 10 min, 0.3 ml of reagent B was added and the mixture was incubated at room temperature for 1 - 3 h.

Standard solutions of serum albumin containing 0.05 - 0.2 µg protein were made and treated alongside the test samples. The extinction at 750 nm was read in both test and standard samples. A standard curve was obtained by plotting known concentration of standards against the corresponding OD_{750} values. The concentration of protein in the sample was then extrapolated from the curve.

2.2.1.7. β-lactamase assay (Perret iodine assay, 1954)

Penicilloic and cephalosporoic acids reduce Iodine, while their parent compounds do not (Groove and Randal, 1955). The Perret assay is based on calculating the amount of substrate destroyed by measuring Iodine used up in a back titration with sodium thiosulphate.

The β-lactam was added in 0.1 M potassium phosphate buffer at a concentration of 7 mM. Five ml of the solution was put into two 50 ml flasks in a shaking water bath at 30 ^{0}C and left to warm up for 5 min. To one flask, 0.05 - 1 ml (v) enzyme solution (depending on the level of β-lactamase activity) was added. After an incubation time of 5 - 20 min (t), 10 ml of 0.0166 N Iodine in sodium acetate buffer, pH 4.2 was added to both test and control flasks. A corresponding volume of enzyme used for the assay in the test flask was also added to the control tube and both were left to stand at room temperature for 10 min.

The residual Iodine was then titrated against 0.0166 N Sodium thiosulphate using starch as indicator. The difference between the control and test flask end point values is the titration difference (TD).

$$\text{International β-lactamase unit/ml} = 2 \times TD/t \times v$$

An international β-lactamase unit is the amount of enzyme which will hydrolyse one µmol of substrate per min at 30 ^{0}C, pH 7 (Davies *et al*, 1975).

2.2.1.8. Isoelectricfocusing

To cross check that a recombinant plasmid encoded a β-lactamase, isoelectric focusing techniques were used as described by Matthew *et al.* (1979). This is based on the fact that proteins can be differentiated on the basis of their pI. Thus, total cell protein are resolved in a pH gradient created by carrier ampholytes. By specifically staining for β-lactamase activity, one can determine the presence or absence of the enzyme.

For a 10 cm x 10 cm gel, sorbitol (1.65 g) and agarose (LKB isoelectrofocusing grade, 0.132 g) were mixed with 15.5 ml water in a screw-capped flask and heated in a boiling water bath until the mixture was homogeneous. The mixture was cooled slightly and carrier ampholyte (0.83 ml) of the required pH range was added. After swirling gently, the gel was poured onto the glass plate and allowed to set. The gel was run on a LKB 2117 "Multiphore" apparatus. The electrode strips were soaked in 1.3 ml, 0.5 M NaOH (cathode) and 1 ml 0.5 M acetic acid (anode) before they were applied. The samples were applied on small pieces of filter paper (0.5 cm x 1 cm, LKB sample application paper). Routinely 5 - 15 µl samples (containing approximately 1 - 2 mg protein) were used. The electrode strips and sample application papers were placed on the gels according to the markings on the LKB template. The gel was run at 400 V (about 50 V/cm) for 1.5 - 2 h or until focusing was complete as judged by a coloured marker protein (usually a cytochrome *c*). If required, the pki of the electrofocussed protein was determined by measuring the pH of the surrounding gel using a pH surface electrode (Russell pH , Ltd., Auchtermuchty).

β-lactamase bands were detected by laying over the gel for a period of 1 min, 3MM Whatmann filter paper soaked in nitrocefin (87/312) (O'Callaghan *et al.*, 1972; 0.5 mg/ml in 0.1 M potassium phosphate buffer, pH 8). The intact substrate is yellow but becomes pink when the β-lactam bond is opened. Thus, the focussed bands with β-lactamase activity appear pink on a yellow background.

2.2.2. Preparation and management of Nucleic acids

2.2.2.1. Plasmid DNA preparations

Small amounts of plasmid DNA for rapid screening of recombinant plasmid were often made from 2 ml overnight broth cultures, while large amounts were made from 500 ml overnight growths.

(i) Birnboim preparation

This refers to the preparation of plasmid DNA from small volumes according to the Birnboim (1979) method. Chromosomal DNA is selectively denatured by alkaline SDS and covalently closed circular DNA is then precipitated from the cleared cell lysate. As small volumes are used throughout, all the operations are carried out in two ml polyethylene microfuge (snapcap) tubes.

1. Grow cells in 5 ml universal bottles with vigorous shaking overnight at 37 ^{0}C.
2. Centrifuge 2 ml of broth culture at 3000g in a bench centrifuge (e.g. Eppendorf Model 3412) for 2 min.
3. Resuspend cells in 100 µl lysis solution (2 mg/ml lysozyme; 25 mM Tris.HCl pH 8, 10 mM EDTA; 50 mM glucose), vortex and leave at 0 ^{0}C (on ice) for 30 min.
4. Add 200 µl alkaline SDS (0.2 N NaOH, 1% SDS) and leave on ice, 5 min.
5. Add 150 µl high salt solution (3 M sodium acetate/acetic acid; pH 4.8) and keep at 0 ^{0}C for 1 h with occasional mixing by inverting.
6. Spin 5 min and carefully avoiding the floating materials, transfer 400 µl clear supernatant to a clean tube.
7. Ethanol precipitate by adding 1 ml ethanol and placing the microfuge tube at -20 ^{0}C for 30 min.
8. Centrifuge 2 min and remove supernatant using a fine drawn out Pasteur pipette.
9. Redissolve precipitate in 100 µl dilute sodium acetate (0.1 M sodium acetate/acetic acid pH 6).
10. Precipitate DNA by adding 200 µl ethanol, spin, discard supernatant, vacuum dry pellet and resuspend in 30 µl TE.

This DNA solution can be used for restriction enzyme analysis.

(ii) Large scale plasmid DNA preparations from *E. coli*

1. Grow 500 ml cells overnight, harvest and resuspend in 6 ml sucrose solution (25% sucrose; 50 mM Tris.HCl pH 8; 40 mM EDTA).
2. Add 1 ml 10 mg/ml lysozyme and leave on ice for 5 min.
3. Add 1 ml 0.5 M EDTA, pH 8 and 13 ml Triton lysis mix (2 ml 10% Triton X-100; 25 ml 0.5 M EDTA pH 8.1; 10 ml 1 M Tris.HCl pH 8.1; and water to 200 ml). When lysis is complete (as determined by clearing of the cell suspension), spin for 30 min (rcf=36,000g).
4. Decant the supernatant and make up to a density of 1.56 by adding 0.96 g CsCl and 0.1 ml (5 mg/ml) EtBr, per ml of supernatant.
5. Load the plasmid suspension into suitable ultracentrifugation tubes (e.g. heat sealed tubes for 50 Ti rotor, OTD 50 B Sorvall ultracentrifuge), and centrifuge at 80,000g for 36 h at 20 ^{0}C.
6. Visualize plasmid DNA bands in the dark under longwave UV and remove with a syringe.
7. Extract EtBr by shaking DNA suspension with an equal volume of NaCl saturated isobutanol, 3 - 4 times.
8. Dilute CsCl by adding 2 - 3 times the volume of water. Precipitate DNA by adding twice the new volume of ethanol. Redissolve precipitate in 500 µl TE.
9. Phenol extract twice, phenol/chloroform extract, chloroform extract and ethanol precipitate. Vacuum dry DNA pellet and resuspend in TE (to a final concentration of 5 mg/ml).

The plasmid DNA is usually ready for use at this stage. Occasionally (due to failure to achieve good endonuclease digestion), a dialysis step was included. Thus, the DNA solution was dialysed against two changes of TE in the cold (4 ^{0}C) for 1 - 2 days. This was enough to remove any residual phenol.

(iii) Large scale plasmid preparations from *Ps. aeruginosa* and *Rps. capsulata*

These strains have large natural plasmids which also occur in low copy numbers. The best plasmid yields were obtained by first selectively concentrating

the plasmids in several cell lysates before the CsCl/EtBr ultracentrifugation step.

1. Grow five to ten 500 ml L-broth or R-broth late stationary phase cultures.

2. Harvest and resuspend 25 g cell pellet (wet weight) in 6 ml sucrose solution (see above in ii).

3. Add 1 ml each of lysozyme (10 mg/ml) and 0.25 M EDTA.

4. Lyse by adding 7 ml alkaline SDS (same as in Birnboim preparation). Add 8 ml 3 M sodium acetate pH 5.6 and chromosomal DNA clumps visibly. Leave on ice for 1 h.

5. Centrifuge at 17,000g for 1 h. Decant supernatant. Add equal volume isopropanol. Leave at -20 ^{0}C for 30 min.

6. Spin (16,000g for 10 min) and redissolve plasmid pellet in 2 ml TE. Pool this with 5 - 10 similar preparations and set up CsCl/EtBr gradients as in *E. coli* preparations. Ultracentrifugation, plasmid band harvest and cleaning were as outlined in the preceding section.

Pre-treatment of the crude plasmid preparation with heat-treated RNase (50 µg/ml for 20 min at 37 ^{0}C), gave a cleaner plasmid preparation.

2.2.2.2. Bacterial chromosomal DNA preparation

This was according to the method of Bellard *et al.* (1973), which is an adaptation of Marmur (1961).

1. Spin a 100 ml fresh overnight growth culture and resuspend cells in 6 ml sucrose solution. For *Rhodopseudomonas* spp, 4 g cells (wet weight) were used.

2. Add 1 ml lysozyme solution (10 mg/ml). Leave on ice for 10 min.

3. Add 1.5 ml 0.5 M EDTA pH 8.1 and keep on ice 10 min.

4. Add 1 volume Triton lysis mix (same as in 2.2.2.1 (ii)).

5. When the cells have lysed, add SDS to a 0.5% concentraton followed by RNAse treatment (100 µg/ml) for 30 min at room temperature. Add proteinase K (50 µg/ml) and leave at 37 ^{0}C overnight with gentle shaking.

6. Extract cell lysate with an equal volume of phenol. * Spin and layer the water phase onto two volumes ethanol in a small beaker.

7. Wind the precipitating DNA around a glass rod and dissolve in half original volume 0.1 x SSC; make up to 1 x SSC by adding the appropriate amount of 10 x SSC and shake with an equal volume of chloroform – isoamylalcohol (24:1) mixture. Spin. **

Repeat * to ** 5 – 8 times until there is little or no protein left in the interphase. The final clear DNA wound round the glass rod is dissolved in 1 ml 0.1 x SSC. This is phenol extracted twice and dialysed against two changes of TE for 24 h at 4 ^{0}C. Finally the DNA is precipitated and redissolved in 500 µl TE. The DNA concentration was usually determined at this stage.

2.2.2.3. λ Phage DNA preparation

λ cI875 phage DNA was used as marker DNA in all the restriction digest analyses in these investigations as well as a vector in a subcloning step. In preparing phage DNA, phage particles were first produced in a large liquid lysate, concentrated and cleaned in two CsCl–centrifugation steps and then phenol extracted. Briefly, the steps are outlined as follows:–

(i) Plate lysates

1. Resuspend a fresh plaque in 1 ml phage buffer with a drop of CH_3Cl and leave for 10 min.
2. Adsorb 0.1 ml of phage suspension onto 0.1 ml fresh plating cells containing 10 mM $MgSO_4$ for 10 min at room temperature.
3. Add 3 ml top agar and plate on fresh L–agar plates. Incubate for 5 – 8 h at 37 ^{0}C.
4. When confluent lysis has occured, add 4 ml L–broth to the plate and harvest after 1 h. Add a drop of CH_3Cl, titre the supernatant and leave till next day.

(ii) Liquid lysates

1. Dilute an overnight broth culture of sensitive bacteria (e.g. ED 8654) 1:20 in 200 ml fresh L–broth + 10 mM $MgSO_4$.
2. Monitor OD every 30 min at 650 nm. When OD is 0.5, add

phages from (i) to a multiplicity of infection 2 - 5. Check OD which will rise, fall when cells begin to lyse and then rise again.

3. Add 0.5 ml Chloroform as OD begins to rise and shake for a further 10 min.

4. Add NaCl (0.5 M) and DNAse (1 µg/ml). Keep shaking for another 20 min.

5. Spin, 18,000g for 12 min to remove bacterial debris and titre cleared supernatant to ensure lysate was successful.

(iii) Phage concentration

1. Dissolve PEG-6000 (10% w/v) in the liquid lysate to precipitate the phages. Leave overnight at 4 ^{0}C.

2. Centrifuge 16,000g for 10 min and resuspend phages in phage buffer, 0.04 - 0.02 the original volume.

3. Layer onto a CsCl step gradient (adapted from S. Bruce). The three solutions of CsCl used here have densities 1.3, 1.5 and 1.7 mg/ml (i.e. 31%, 45% and 56% w/v CsCl). They were made up as follows: first a saturated CsCl solution (1.9 mg/ml) is made and diluted to obtain the different concentrations as follows:

Density required	Volume (1.9mg/ml)	Volume Phagebuffer	Ratio	Amount in gradient
1.9	1 ml	0 ml		0.5 ml
1.6	1 ml	0.5 ml	2:1	0.5 ml
1.48	1 ml	1 ml	1:1	1.0 ml
1.32	1 ml	2 ml	1:2	1.5 ml

The solutions are overlaid from above and the phage suspension is added last. A maximum of 10 ml sample can be loaded on such a gradient. Spin in a Beckman ultracentrifuge 79,000g for 2 h at 20 ^{0}C in a 6 x 14 rotor.

4. Collect whitish phage band with a syringe, mix with 41.5% preclarified CsCl solution and centrifuge in a 6 x 5 Ti swing-out rotor, 79,000g for 24 h at 4 ^{0}C.

5. Remove phage band with a Pasteur pipette and dialyse against TE for 1 h to remove CsCl.

(iv) Phage DNA

1. Dilute phage suspension with 0.5 volume TE and phenol extract three times; phenol/chloroform extract once; chloroform extract and dialyse the water phase containing phage DNA against three changes of TE at 4 ^{0}C to remove any residual organic solvents.

2. Determine DNA concentration, ethanol precipitate and reconstitute in TE at a final concentration of 5 mg/ml.

2.2.2.4. RNA preparation

This was done by the method of Shaw and Guest (1982) which is a modification of Salsier *et al.* (1967). Briefly,

1. Grow a 500 ml culture in L-broth until OD at 650 nm is 1.

2. Spin 16,000g for 8 min, wash cells in 50 ml TE.

3. Spin as in (2), resuspend cells in 6.5 ml RNA preparation solution A (5 mM $MgCl_2$; 10 mM Tris.HCl pH 7.3; 10 mM KCl), and transfer to a sterile universal bottle.

4. Add 2 mg lysozyme, freeze cells (-70 ^{0}C for 30 min), and thaw.

5. Add 0.9 ml 10% SDS and incubate at 60 - 65 ^{0}C until turbidity drops (5 - 10 min).

6. Add 0.33 ml 2.5 M sodium acetate pH 5.2 (preparation goes cloudy).

7. Extract with 1 volume water saturated phenol at 64 ^{0}C. Shake the mixture in a 65 ^{0}C water bath for 4 min; spin 5 min (16,000g); remove and discard lower phenol layer using a Pasteur pipette.

8. Repeat (7), 6 - 8 times until phenol layer clears.

9. Add 1 g solid NaCl and divide the preparation between two siliconized corex tubes. Add 1.5 volumes cold ethanol to each and leave at -20 ^{0}C for 1 h.

10. Spin for 10 min, wash pellet three times with RNA preparation solution B (70% ethanol; 10 mM Tris.HCl pH 7.5; 10 mM NaCl).

11. Vacuum dry pellet and redissolve in 1 ml water.

Determine RNA concentration in the same way as it is done for DNA (see below). Routinely, the yield was 16 mg/ml. For *Rps. capsulata*, 3 g (wet weight) cells

were used.

2.2.2.5. Determination of DNA concentration

DNA was usually resuspended in water or TE. The OD at 260 and 280 nm was determined for a 1/20 -1/200 dilution using a water/TE blank.

OD_{260} = 1 is equivalent to 50 µg/ml DNA
Conc of DNA (µg/ml) = OD_{260} x 50 x dilution factor

$OD_{260/280}$ is a measure of the purity of DNA (Maniatis *et al*, 1982). A value within the range 1.8 - 2.0 is good and such DNA was used in the cloning steps.

2.2.2.6. DNA purification

Standard methods (Maniatis *et al*, 1982) were used in the purification of DNA. Organic solvents were used to clean and concentrate DNA while contaminating proteins were removed by extraction with distilled phenol.

(i) Phenol distillation

Redistilled phenol was used in these investigations. Solid phenol was distilled in an atmosphere of nitrogen and collected under 1 M Tris.HCl pH 7.5 containing 0.1% of 8-hydroxyquinoline (AR). 8-hydroxyquinoline is an antioxidant, an inhibitor of RNAse and a weak chelating agent (Kirby, 1956). In addition, it imparts a yellow colour which provides a convenient way of identifing the phenol phase.

Routinely, four kg were distilled in a single preparation to yield 1 l liquid phenol. This was divided into 20 ml volumes and stored frozen at -20 ^{0}C after saturation with 1 M Tris.HCl pH 7.5. Before use, the 20 ml volume was thawed and pre-equilibrated with TE or water as required.

(ii) Phenol extraction

To phenol extract a DNA preparation, 1 volume phenol is added and the tube/bottle is rolled gently at room temperature for 5 min. Spin at low speed to

layer. Remove the phenol layer with a Pasteur pipette. Repeat procedure twice or until both layers are clear and there is no protein visible in the interface. Sometimes (especially when working with very small volumes of DNA solutions), the phenol was back extracted; i.e. extracting used phenol with an equal volume of water or TE. This was passed through the other volumes used and later pooled with the original DNA solution.

Generally, the phenol extraction was followed by a 1 volume phenol/chloroform (1:1) extraction of the DNA solution and a chloroform extraction to remove residual phenol. A dialysis against TE with 2 - 3 changes in 24 h at 4 ^{0}C was sufficient to remove all organic solvents. (dialysis tubing 8/32, was boiled for 10 min in 1 mM EDTA before use.)

(iii) Ethanol precipitation

This serves to clean and concentrate DNA in a solution. It was also used when a change in the surrounding buffer was required. The technique is rapid and very efficient, allowing recovery of short DNA fragments (e.g. less than 1 kb, Maniatis *et al*, 1982).

1. Add 1/10 volume 2 M sodium acetate pH 5.6 to the DNA solution
2. Add 2 volumes cold (-20 ^{0}C) ethanol, keep 15 s in a dry ice/methanol bath or for 30 min at -20 ^{0}C.
3. Spin 5 min (12,000 g), and pour off the supernatant.
4. Vacuum dry pellet and resuspend in TE/water.

When DNA solutions were very dilute or fragments too small (0.1 - 0.5 kb), efficiency of recovery of DNA was enhanced by adding carrier tRNA (50 µg/ml) before using the above procedure.

(iv) Isopropanol precipitation

Isopropanol precipitation serves the same purpose as (iii) and was preferred when there was not enough space for addition of 2 volumes ethanol to the tube/bottle.

1. Add 1/10 volume 2 M sodium acetate pH 5.6.
2. Add 0.6 - 1 volume isopropan-2-ol to the DNA solution and leave on ice for 30 min or 15 s in a dry ice/methanol bath.
3. Spin 5 min, wash pellet in 1 ml ethanol.
4. Spin, vacuum dry pellet and resuspend in TE/water.

2.2.3. *In vitro* DNA manipulations

2.2.3.1. Restriction analysis

The nomenclature for restriction enzymes is as proposed by Smith and Nathan (1973).

Routinely, 0.5 µg DNA was digested in a small volume (20 - 40 µl) with 0.5 unit of enzyme in the appropriate restriction buffer (Table 2.1) and incubated at 37 ^{0}C for 1 h. The reaction was stopped by 10 min incubation at 70 ^{0}C.

Stock solutions (10 x) of the restriction buffers (shown below) were stored at -20 ^{0}C. When double digests were carried out, the buffer containing the lowest salt concentration was used. If digests could not be carried out simultaneously, DNA was restricted with one enzyme, ethanol precipitated and reconstituted before digesting with the second enzyme. Universal buffer (section 2.1.8) was found to be sufficient for most restriction endonucleases in my experience.

Table 2.1: Restriction assay buffers.

Enzyme	EDTA	Tris-HCl	pH	NaCl	$MgCl_2$	2-Me	BSA (µg/µl)
Bam H1	-	6	7.5	50	6	6	1
Eco R1	-	100	7.5	50	5	-	1

Hind III	–	60	7.5	7	–	–	1
Xho 1	–	500	7.5	500	100	–	–
Sal 1	0.2	8	7.6	150	6	–	–
Cla 1	–	10	8.0	–	10	–	0.1
Sau 3A	–	6	7.5	50	6	–	0.1
Nci 1	–	6	7.5	6	6	6	0.1
Taq 1	–	500	7.5	500	100	–	0.1
Dde 1	–	100	7.5	100	–	–	–

(except those for BSA)

All concentrations are in mM.

2.2.3.2. Gel electrophoresis

DNA fragments resulting from enzyme digestions were separated using agarose gel electrophoresis. All the gel apparatus used in these investigations were constructed in this laboratory from perspex.

(i) Agarose gels

These were run according to the recommendations of Maniatis *et al.* (1982). For gel analysis of restriction digests, 0.75% agarose gels (200 ml) were run overnight at 15 mA (LKB, Multiphor 2117) in horizontal perspex plates (160 cm x 250 cm) set between two buffer tanks. Muslin wicks were used to connect the gel to the gel buffer (40 mM Tris.HCl pH 8.2; 20 mM NaAc; 1 mM EDTA) in the tanks. Minigels (1% agarose gels) were used for rapidly checking the efficiency of digestion as well as the range of the resulting fragments. They were run at 40 mA for 40 - 50 min in the "concorde" apparatus in which the gel is totally submerged in the gel buffer.

Agarose gels were made by adding the appropriate amount of agarose to the gel buffer and dissolved by boiling for 5 min. The cleared solution was then cooled to about 50 ^{0}C and poured. Samples were mixed with 1/3 - 1 volume loading buffer (section 2.1.8) and applied to preformed wells in the gels. Routinely, the sample was run into the gel at a high voltage (500 V) and when the blue dye had run into the gel, the voltage was lowered to 70 - 100 V, so as to achieve the appropriate current flow. For the large horizontal gels which were

often run overnight, the gel was then covered with Saran wrap.

The gels were stained with the fluorescent dye ethidium bromide (Sharp *et al*, 1973) either by direct incorporation into the gel (0.5 µg/ml) or by soaking electrophoresed gel in 0.5 µg/ml EtBr solution for 20 - 30 min. The gel is then destained for 15 min in water to remove nonspecifically bound dye.

The gels were viewed under UV and photographed using Ilford FP4 film through a red filter (Kodak Wratten filter number 9).

(ii) Purification of DNA fragments from gels

It was sometimes necessary to isolate a specific DNA fragment from a mixture, especially as this gave more specific recombinants.

The DNA was digested with the required enzyme(s) and run in the submerged minigel apparatus for 20 min to 2 h depending on the relative sizes of fragments being fractionated. Low gelling temperture (LGT) agarose (1%) was used in these steps and was usually poured 8 - 16 h before use and kept at 4 ^{0}C.

The gel was stained, visualized and photographed as outlined in section (i) above. The agarose surrounding the target fragment was then cut out, placed in a sterile microfuge tube and left at 65 ^{0}C for 10 min to melt. The tube was cooled down to about 40 ^{0}C at room temperature, phenol extracted three times, back extracted and isopropanol precipitated. The pellet was washed with 1 ml ethanol, vacuum dried and resuspended in 20 µl TE.

2.2.3.3. Ligation

Target and vector DNA (0.2 µg each) were restricted separately with the appropriate enzyme(s), pooled and ethanol precipitated. The mixed fragment pellet was resuspended in a 10 - 20 µl ligation "cocktail" (Tris.HCl pH 7.5, 660 mM; EDTA pH 9, 10 mM; $MgCl_2$, 100 mM; 2-mercaptoethanol or DTT, 100mM;

ATP, 10 mM; NaCl, 400 mM; T_4 ligase, 400 - 800 units).

The ligation mixes were then incubated overnight at 16 ^{0}C. Controls containing target DNA and no vector, and vice versa, were also routinely included.

2.2.3.4. Recombinant plasmid/phage recovery

The *in vitro* recombination reactions mediated by T_4 ligase generate a range of recombinants depending on the fragments present originally. With a knowledge of the phenotypic characters coded for by the various DNA fragments, one can select for these characters after transformation or transfection of the recombinant DNA into the appropriate hosts (competent cells).

(i) Making competent cells

In the investigations reported in this thesis, the competent cells were bacteria treated such that they could accept foreign DNA upon subsequent transformation. The procedure used was an adaptation from Mendel and Higa (1970).

1. Dilute a fresh overnight growth of a competent strain (HB 101, ED 8654, etc.) 1:50 in L-broth and grow till OD_{650}=0.55.
2. Chill culture on ice and spin, 17,000g for 10 min.
3. Wash cells in equal volume of cold (0 ^{0}C) 100 mM $MgCl_2$.
4. Spin and wash cells again in 1/2 volume 100 mM $MgCl_2$.
5. Spin and resuspend cells in 1/10 volume 100 mM $CaCl_2$.
6. Keep on ice till use.

Routinely, 100 µl competent cells were tested by transforming with 0.2 µg plasmid DNA of known phenotype and spreading on appropriate media. The rest of the competent cell preparation was kept at 4 ^{0}C overnight, awaiting the test results.

(ii) Transformation/transfection

For transformation, 10 - 20 µl DNA solution was mixed with 100 - 200 µl competent cells, left on ice for 30 min, heat shocked at 42 ^{0}C for 2 min and put back on ice for a further 30 min. L-broth (1 ml) was added and the tube was incubated at 37 ^{0}C for 30 min to allow expression of any recently acquired gene(s), before 0.1, 0.3, and 0.5 ml portions were spread on L-agar plates containing the relevant antibiotics.

During transfection, the recombinant DNA solution was diluted by adding an equal volume of TE. Then, 10 - 20 µl samples were mixed with competent cells as outlined above in transformation. Plating cells (100 µl) were added to the transfected cells before spreading on L-agar plates containing ampicillin.

While the appearance of stable colonies signified positive transformants during selection for the presence of recombinant plasmids, the recombinant phage carrying an amp gene was identified as a distinct plaque surrounded by a halo of bacterial cells.

2.2.4. DNA sequencing

2.2.4.1. Introduction

The principles involved in the dideoxy sequencing methods have been thoroughly described (Sanger *et al*, 1977; 1980; Bankier & Barrell, 1983), so only an outline of the various steps is given here.

In general, the DNA fragment to be sequenced is purified, ligated on itself (to minimise the availability of loose ends) and sonicated. The resulting random population of DNA fragments are treated with T_4 DNA polymerase to fill in any single stranded ends, fractionated to select fragments of sizes between 300 - 600 bp, and then cloned into the Sma 1 site of a double stranded M13 vector. The recombinant DNA is transfected into *E. coli* host cells, which are

subsequently spread on L-agar plates containing inducer and indicator molecules. Successful recombinants are identified by the white plaque phenotype as against the unchanged vector blue plaques.

Single-stranded templates are made from packaged phages. Primed synthesis across the insert is effected by addition of Klenow DNA polymerase and four precursor deoxynucleotides, one of which is radioactively labelled. Specific termination at each of the four deoxynucleotides is brought about by including in the deoxynucleotide mixes, dideoxynucleotide triphosphates (ddNTPs). Four separate reactions are carried out and the four sets of reaction products are fractionated on denaturing polyacrylamide gels. The corresponding sequence is then read off the ladder of terminated fragments.

2.2.4.2. The random cloning strategy

Three main ways have been used in generating DNA segments for sequencing in M13 vectors:

- By the use of specific restriction endonucleases.
- By cloning a large DNA fragment into double stranded M13 DNA and in time course digests with exonuclease III (Anderson, 1981), generating a random range of deletions spanning the whole insert.
- Mechanical shearing of DNA.

The generation of discrete DNA fragments with specific restriction enzymes has the disadvantage that the relevant site(s) might not be present, especially when dealing with sizes which are 3 kb or less. M13 vectors do not grow well with large inserts, with the consequent production of reduced yields. Deletions also arise at low frequency by aberrant homologous recombination and the shortened mutants tend to outgrow the full length phage. Prior to sequencing, this anomaly cannot be easily detected. Thus, shearing DNA by using mechanical force seems a more practical way to generate random DNA fragments for sequencing. By

Fig. 2.3: Rate of Sequence accumulation

(Culled from Bankier & Barrell, 1983)

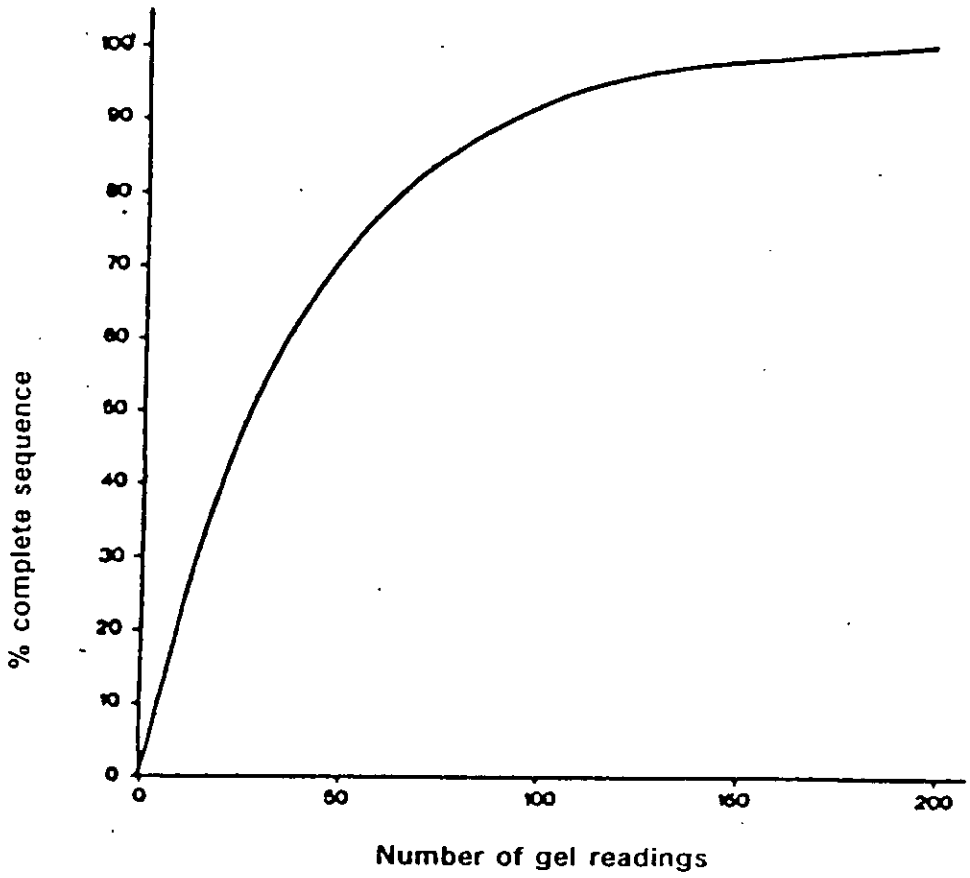


Fig. 2.3: Rate of sequence accumulation. The percentage sequence accumulation is plotted against the theoretical number of gel readings required to determine the sequence of a 5 kb fragment on both strands, assuming an average number of 250 bp per clone.

introducing the fractionation step, a range of fragments tolerated by M13 vectors, and allowing for competition during the subsequent ligation step, is selected. The latter is deduced from the observation that smaller fragments tend to be picked up repeatedly.

In random sequencing, each base must be determined twice on each strand. All gel readings must also overlap with two other gel readings; i.e. one on either side in order to ensure that no data are missed. The lowest practical limit for determining a sequence would be to determine a base on average three times.

Thus theoretically, for an average gel reading of "a" bp, one would require:

$$N \text{ clones} = 6 \times L/a$$

where, L = length of DNA to be sequenced in bp

a = average reading in bp/gel

Nevertheless, the sonication protocol proved less efficient than was expected. As illustrated in Fig. 2.3; (Bankier and Barrell, 1983), the theoretical number of gel readings of 250 bp each, required to complete a sequence of 5 kb DNA on both strands, shows a great deal of redundancy towards the completion of the sequence. There is, however, an unquantified but valuable increase in confidence in the completed sequence as a result of the extra accumulated data.

Hence, after more than 100 gel readings, about 95% of the sequence was obtained and a 100% increase in gel readings did not significantly make any difference. The project could be completed by extending readings of gels adjacent to gaps. Where this is not practicable, clone turn arounds of gels in the region can supply the missing data (Bankier & Barrell, 1983).

The random sequencing protocol has another advantage in that it is quick as all cloning and preparation of templates can be completed in less than one week.

2.2.4.3. Cloning into M13

1. Purify the fragment to be sequenced as outlined in section 2.2.3.2.(iii).

2. Self ligate 5 - 10 µg fragment in a 30 µl volume ligation "cocktail".

3. Check ligated DNA on minigel and sonicate if percentage of re-ligation is about 80%. Where re-ligation was too efficient, 10% total DNA of linear fragment was added before sonication. This was necessary as religated junctions containing inverted repeat regions do not clone into M13 vectors. In these experiments, a Cup-Horn sonicator made by Heat Systems Ultrasonics, Inc., model W-375, was used. The sample is given two bursts of 80 s each and spun in between treatments and afterwards.

4. End repair fragments by adding 2 µl 0.25 mM dNTPs (chase) and 2 µl T_4 polymerase. Leave at room temperature for 30 min.

5. Size fractionation:- Add 6 µl loading buffer and apply sample in a 1 cm slot of a 40 ml 1.5% agarose gel and run at 30 mA until the blue dye has run 1 - 2 cm into the submerged minigel. View under UV, cut a 1 cm by 2 cm slot at the lower limit of the required size by comparing with an adjacent track containing pBR322 Sau 3A fragments (1.5 kb, 650 bp and 350 bp). Fill slot with 800 µl TBE and electroelute DNA into buffer by running the gel at 20 mA for 45 s, 15 times (i.e. remove the buffer and refill after every 45 s, 15 times). Phenol extract DNA solution once, isobutanol clean and isopropanol precipitate. Ethanol wash, vacuum dry and reconstitute in 40 µl water.

6. Vector preparation:- M13 mp8 (and later M13 mp19) double-stranded plasmid DNA was digested with Sma 1, calf intestinal phosphatase enzyme treated and phenol extracted. It is then isopropanol precipitated, ethanol cleaned and vacuum dried before resuspending in water (1 µg/ml).

7. Ligation of the sonicated and end-repaired target DNA fragments into M13 vectors was as outlined in section 2.2.3.3. in 10 µl mixes.

8. Transfection:- The *E. coli* (TG1) host cells were made competent using the Hanahan method (1983) as follows:

    a. Dilute an overnight broth culture of TG1 (1:100) in L-broth and grow at 37 ^{0}C until OD_{650} is 0.4 - 0.6.

    b. Spin 16,000g at 4 ^{0}C, resuspend cells in 3 ml Hanahan transformation buffer (10 mM MES, pH 6.2; 100 mM KCl; 45 mM $MnCl_2.4H_2O$; 10 mM $CaCl_2.2H_2O$; 3 mM hexaminecobaltic chloride)

c. Add 100 µl DMF, leave 10 min on ice, add 100 µl DTT solution (2.25 M DTT, 40 mM KAc pH 6), keep on ice for a further 10 min. Add 100 µl DMF, leave on ice for 5 min and cells can now be used. Transformation efficiency was observed to be reduced to zero if cells were not used within 1 h of preparation.

9. To 200 µl competent cells, add 1 - 2 µl portions of the 10 µl ligation mixture and keep on ice for 45 min. Heat shock at 42 ^{0}C for 2 - 5 min and add 3 ml top agar, 25 µl 2% BCIG (Xgal) in DMF and 25 ul 2.5% IPTG.

10. Add 100 µl plating cells and spread on L-agar plates. Incubate at 37 ^{0}C overnight.

2.2.4.4. Template DNA preparation

1. Put two ml of a 1:100 dilution of an overnight culture of TG1 into a 5 ml glass culture tube and pick a white plaque into each tube.

2. Grow with vigorous shaking at 37 ^{0}C for 4 - 5 h. Spin culture in a microfuge tube, 5min at 16,000g, and transfer supernatant to a clean tube.

3. Add 150 µl PEG solution (20% PEG; 2.5 M NaCl), and leave at room temperature for 10 min.

4. Spin 5 min, suck off PEG solution, respin for 30 s and remove residual PEG solution.

5. Resuspend the phage pellet in 100 µl 0.3 M NaAc pH 5.6 and phenol extract (100 µl phenol), isopropanol precipitate, ethanol wash and vacuum dry the phage DNA.

This is the template DNA solution and was stored at -20 ^{0}C until required.

2.2.4.5. Sequence reactions

All sequencing reactions were carried out in microtitre plates. Wells are marked out for T, C, G and A reaction mixes.

1. Dispense 2 µl template DNA into each well in sample rows of 4 wells/clone.

2. Add 2 µl of primer mix (0.2 pmol, 17mer primer; 100 mM Tris.HCl pH 8; 50 mM $MgCl_2$), to each well.

3. Make the different nucleotide mixes for T, C, G and A reactions (see below) and dispense in 2 µl portions to the appropriate well.

4. Mix the above solutions by covering microtitre plate with Saran wrap and spinning for 30 s in an IEC Centra-3 bench centrifuge.

5. Incubate in a 55 ^{0}C oven for 30 min and add 2 µl of the enzyme/label "cocktail" (0.1 mM DTT; Klenow polymerase, 2 units; [α-^{35}S] ATP, 4 µCi; 0.07 mM Tris.HCl pH 8). Spin to mix and incubate at room temperature for 20 min.

6. Add 2 µl chase (0.25 mM cold dNTPs) and leave at room temperature for another 15 min.

7. Add 2 µl formamide dyes (100 ml deionized formamide; 0.1 g xylene cyanol FF; 0.1 g bromophenol blue; 2 ml 0.5M EDTA), spin and keep sample at 80 ^{0}C for 15 min.

This encourages DNA denaturation as well as reduction in the water content. Load onto denaturing polyacrylamide gels. If samples were not electrophoresed immediately, they were stored for up to one week at -20 ^{0}C. The solutions in the sequencing reactions were dispensed using a Hamilton PB600 repetitive dispenser fitted with a 1710 LT syringe.

Nucleotide mixes

The nucleotide mixes were made up by mixing the appropriate amounts in µl volumes as shown below:

Stock solution	T	C	G	A
0.5 mM dTTP	25	500	500	500
0.5 mM dCTP	500	25	500	500
0.5 mM dGTP	500	500	25	500
10 mM ddTTP	50	-	-	-
10 mM ddCTP	-	8	-	-
10 mM ddGTP	-	-	16	-
10 mM ddATP	-	-	-	3
TE (10 mM, 0.1 mM EDTA pH 8)	1000	1000	1000	1000

dNTP chase

0.25 mM dTTP; 0.25 mM dCTP; 0.25 mM dGTP; 0.25 mM dATP.

2.2.4.6. DNA sequencing gels

(i) Pouring the gel

Routinely, 50 x 50 cm glass plates were used for running the sequencing gels. The plates were thoroughly cleaned with soap and water and dried. The top (notched) plate was siliconised and both plates were cleaned with alcohol. The plates were then taped together, using yellow 2.5 cm wide Selotape, separated by 1 cm wide plastikard spacers.

For each gel, 5 x TBE gel mix (7 ml, 5 x TBE gel mix; 14 µl AMPs; 14 µl TEMED) and 0.5 x TBE gel mix (45 ml, 0.5 x TBE gel mix; 70 µl AMPs; 70 µl TEMED), were prepared. In a 25 ml pipette, 6 ml 0.5 x TBE gel mix and all 7 ml of the 5 x TBE gel mix were taken up. With the taped plates held diagonally, the gel mix was carefully poured down one edge and when nearly all the gel mix in the pipette had gone in, the plates were lowered to the horizontal level to stop the flow. The rest of the 0.5 x TBE gel mix was then put in using a pre-filled 30 ml plastic syringe. The slot-former (also made of plastikard) was put in place and both top and bottom edges of the plates were clamped. The gel was left lying horizontally for 20 - 30 min to set.

(ii) Loading the sample

The slot former was removed and the glass plates holding the gel were clamped onto the vertical gel apparatus used for running sequencing gels. TBE buffer (single strength) was poured into the top and bottom tanks. Denatured DNA sample (from section 2.2.4.5) was loaded onto marked slots using drawn out capillary polypropylene tubing. The gel was electrophoresed at 37 - 40 mA for 3 h (or 15 min after the deep blue dye had run out of the bottom of the gel).

The glass plates were disengaged from the gel apparatus and the notched plate was carefully prised off. The gel, still on the whole glass plate was fixed in 10% acetic acid and 10% methanol for 15 min. After draining the gel, it was transferred to 3MM Whatman paper, covered in Saran wrap and dried on a gel drier at 80 ^{0}C for 30 min to 1 h.

The Saran wrap was taken off and the dried gel was placed in direct contact with an X-ray film overnight in a cassette.

Gel solutions

40% acrylamide stock solution	38% acrylamide, 2% bis-acrylamide. The solution is made up in distilled water and deionized by adding 20 g Amberlite MB1 resin (Hopkin and Williams). The solution is then filtered and stored at 4 ^{0}C until required.
0.5 x TBE gel mix	150 ml 40% acrylamide; 50 ml 10 x TBE; 460 g urea, made up to 1 l with water.
5 x TBE gel mix	150 ml 40% acrylamide; 500 ml 10 x TBE; 460 g urea; 50 mg bromophenol blue, made up to 1 l with water.

2.2.4.7. Clone turn around

After sequencing more than twice the theoretical number of clones required (see section 2.2.4.2), some regions were still sequenced only on one strand in the *Rps. capsulata* derived fragment (section 5.5). A clone in M13 mp8 was then turned around in order to determine the complementary sequences (Bankier & Barrell, 1983). Essentially, a template likely to contain the required region (because it starts upstream of it and because it has the desired length as observed on the sequencing gels), is primed and made double stranded by addition of Klenow and four cold deoxynucleotides. The insert is cut out by the use of two enzymes having sites on both sides of the polylinker (Eco R1 and Bam H1). The excised DNA fragment is then cloned into M13 mp9 or M13 mp19 which have the cloning sites of the polylinker in the opposite orientation. Thus, a

single stranded template generated from the resulting clone is complementary to the original M13 mp8 or M13 mp18 clone. Briefly;

1. Anneal 8 µl template DNA, 4 µl sequencing primer and 2 µl Tris.HCl pH 8; $MgCl_2$ buffer in a 60 ^{0}C oven for 30 min.
2. Add extension "cocktail" (6 µl chase; 1.5 units Klenow) and leave at room temperature for 30 min.
3. Phenol extract once, isopropanol precipitate and ethanol wash. Dry and reconstitute pellet in 14 µl water.
4. Digest double stranded DNA by adding 2 µl each of Bam H1 and Eco R1 enzymes. Incubate at 37 ^{0}C for 1 - 2 h.
5. Add colourless loading buffer (25% sucrose solution) and load into 1% LGT agarose gels using the 1 cm slot. Run gel at 35 - 40 mA for 15 - 20 min. Cut out small insert band and purify as outlined in section 2.2.3.2.(iii).

Ligation of purified fragment to M13 mp19, template DNA preparation and sequencing procedures were executed in the same way as outlined for M13 mp8 clones (sections 2.2.4.3 - .6). The templates prepared in M13 mp19 were usually checked for hybridization to M13 probes (section 2.2.5.3 below), made from the original M13 mp8 clones before sequencing.

2.2.4.8. Computer methods for analysis of sequences

A Vax 11/780 computer (Digital Equipment Corp. Marynard. M.A. USA) was used for compilation and analysis of all the randomly generated DNA sequences. In addition, a Graf/bar GP7 sonic digitizing device (Science Accessories Corp. Boston, USA) and VT640 enhanced graphics terminal (Digitizing Engineering Inc. Sacramento, USA) were also used.

Most of the sequence analyses were carried out using the Staden programs which are written in standard Fortran and are mostly interactive. The program GELIN (Staden, 1984) was used to enter new sequence data into the computer by reading directly from an autoradiograph using the Graf/bar digitizer. The gel readings were checked for the presence of the restriction enzyme sequences

used in excising the original fragment and for homology to any M13 sequences using the program SCREENV (Staden, 1982). Any vector sequences were rejected by the program. DBAUTO (Staden, 1982a) was then used to compile the sequence database. Problem gel readings, editing the database and other functions were dealt with by using the general database handling program DBUTIL (Staden, 1982a).

The major analysis programs used for interpreting the sequences were ANALYSEQ (Staden, 1984) which has options to search for most of the properties of nucleic acids and proteins; and DIAGON (Staden, 1982a) for the diagonal matching of homologous sequences. Many of the results are presented graphically on the VT640 VDU.

For interpretations of sequence relationships (chapter 6), some programs of the University of Wisconsin Genetics computer package (UWGCG) were used. Protein sequences were matched using BESTFIT and GAP. Protein databases were searched for sequences homologous to open reading frames using WORDSEARCH. CHOUFAS was used to obtain secondary structure predictions and PRETTY served to align nucleotide and protein sequences.

2.2.5. *In vitro* labelling of DNA and hybridization techniques

2.2.5.1. Strategy

Hybridization studies were carried out to ensure that recombinant DNA originated from target DNA. Firstly, the DNA (restricted and run on agarose gels or from plaques) was transferred and fixed onto nitrocellulose filters. The immobilized DNA was exposed to the probe in conditions which permit hybridization and then washed in a series of buffers to remove all non-specific binding. The filters were later dried and autoradiographed so that related DNA could be identified as those which produced dark spots/bands.

2.2.5.2. Blotting DNA & RNA onto nitrocellulose filters

(i) Southern transfers

DNA bands were transferred from agarose gels onto nitrocellulose filters according to the method of Southern (1975).

1. Digest DNA samples to be blotted and run gel using large horizontal plates.
2. Visualize gel under UV (long wavelength) and photograph. If necessary, cut gel to minimum size.
3. Soak gel in denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 40 - 60 min, depending on the size of gel.
4. Wash in distilled water for 1 - 2 min. Neutralize in 1 M Tris.HCl pH 5.5; 3 M NaCl for 1 h.
5. Wash in 2 x SSC for 10 min.
6. Place a perspex or glass plate slightly bigger than the gel on two big bungs in a tray. Fill tray 2 cm lower than the plate with 20 x SSC. Wet two layers of blotting paper with 2 x SSC and lay over the plate so that they dip into the solution (these are the wicks). Now, place the gel on the wicks and lay nitrocellulose filter (cut to same size as gel), on top of the gel after thoroughly wetting it with 2 x SSC. Carefully remove air bubbles.
7. Soak two Whatman 3MM filter papers (cut to same size as gel) in 2 x SSC and lay on top of filter. Cut ordinary blotting paper to same size as gel, enough to make a wad 15 - 20 cm thick. Place wad of blotting paper as well as a weight about 1 kg on the filter paper and allow transfer to continue for 16 - 48 h.
8. Remove filter and soak in 2 x SSC for less than 1 min.
9. Air dry for 30 min at 37 ^{0}C; vacuum dry at 80 ^{0}C for 2 h.

Filters were stored in sealed plastic bags until required. Some filters were reused by boiling in distilled water for 10 - 15 min to disengage bound probe.

(ii) Plaque blotting

This was done according to the method of Benton and Davies (1977). Petri dishes (in duplicates) containing plaques were chilled at 4 ^{0}C to harden the top

agarose. Nitrocellulose filters cut with a circular template were marked and placed on the plates for 1 min. Next, the filters were carefully removed and:

- placed on a pad of Whatman 3MM filter paper soaked in 0.5 M NaOH; 1.5 M NaCl for 2 min (the surface which had been in contact with the plaques being placed uppermost),
- submerged in 0.5 M Tris.HCl pH 7.4; 3 M NaCl for 10 min, and
- rinsed in 2 x SSC for less than 1 min.

The filters were then air dried and vacuum baked at 80 ^{0}C for 2 h.

(ii) Dot blotting of DNA/RNA

Purified nucleic acid solutions were used for testing M13 clone turn around templates (refer to section 2.2.4.7) and RNA preparations, in order to ensure that the RNA preparation contained sequences that would interact with the probe.

For M13 templates

Template DNA (5 µl) was spotted onto dry nitrocellulose filters which were then placed on wads of blotting paper soaked with the following solutions:-

1. 0.5 M NaOH, 3 min;
2. 0.5 M NaOH, 1.5 M NaCl, 4 - 5 min;
3. 3 M NaCl, 1 M Tris.HCl pH 7.5, 2 min;
4. 3 M NaCl, 1 M Tris.HCl pH 7.5, 2 min.

The filter was washed in 2 x SSC for 4 min, air dried at 37 ^{0}C and baked at 80 ^{0}C under vacuum.

RNA dot blotting

Nitrocellulose filters cut to the appropriate size, were pre-equilibrated in 3 M NaCl, 0.3 M sodium citrate pH 7 (20 x SSC) for 10 min and air dried at 37 ^{0}C for 30 min. Samples of RNA (10 µg) were spotted on, dried at 37 ^{0}C for 30 min and

the filter was baked in vacuo for 2 h.

2.2.5.3. Probes

(i) Nick translation

Labelling of DNA by nick translation was carried out according to the method of Rigby *et al.* (1977) in 2 ml microfuge tubes containing:

1. 3 µl (0.25 - 1 µg) double stranded DNA sample (can be fragment, plasmid or phage).
2. 20 µl 1 x buffer + cold dNTPs (see below),
3. 1.5 µl [α-^{32}P]dCTP (20 µCi),
4. 1 µl DNAse I stock solution.

The mixture was left at room temperature for 1 - 1.5 min and 0.5 µl *E. coli* polymerase I (1 unit/µl) was added. The tube was then kept at 15 ^{0}C for 3 h.

The reaction mix was passed through a sterile Sephadex G-50 column (15 cm x 0.7cm) to separate labelled DNA from unincorporated nucleotides. The column was eluted with sterile water and Cerenkov counts were determined for 5 µl fractions. A successful nick translation reaction is characterised by two radioactive peaks - the first corresponding to labelled DNA and the second one to unincorporated label. The first peak was usually pooled and used as probe in the hybridization reactions.

Solutions used

1 x buffer + cold dNTPs	100 µl 4 x buffer; 4 µl 2 mM dTTP, dGTP, dATP; 1 µl β-mercaptoethanol; 290 µl water.
4 x buffer	210 mM Tris.HCl pH 7.5; 21 mM $MgCl_2$; 20 µg/ml BSA.
DNAse I stock solution (2 x 10^{-5} mg/ml DNAse in Polymerase I buffer)	1 µl 2 x polymerase dilution buffer; 4 µl 1 mg/ml DNAse I; 1 ml glycerol. Mix 20 µl of 2 x 10^{-3} mg/ml DNAse solution (from above) with 1 ml 2 x polymerase I dilution buffer, and 1 ml glycerol to get

working DNAse I stock solution.

2 x Polymerase I dilution buffer
2 ml 1 mg/ml BSA or gelatin; 100 µl 2 M ammonium sulphate; 20 µl 1 M β-mercaptoethanol; 100 µl 1 M Tris.HCl pH 7.5.

The solutions were dispensed into 250 µl aliquots and stored at -20 ^{0}C. The yield was routinely 1.5 x 10^{6} cpm (Cerenkov).

(ii) M13 probe

The primer extension method of Hu and Messing (1982) as modified by Brown (1982) was used to make single stranded template DNA, radioactive. In this instance, the primer (PH M 235) is complementary to a region on the 5′ side of the insert. Thus, while the M13 vector is made double stranded, the insert remains unpaired and can be hybridized to related DNA. In short:-

1. Mix 5 µl M13 template, 1 µl 10 x Hin buffer (Tris.HCl pH 8, 1 mM; NaCl, 6 mM; $MgCl_2$, 0.66mM); 1 µl primer PH M 235; and 3 µl water. Boil for 3 min.

2. Cool to room temperature and add 3 µl 500 µmol dATP, dCTP, dTTP mix; 6 µl water; 1.5 µl [α-^{32}P]dCTP (10 µCi); 1 µl Klenow (1 unit).

3. Incubate at room temperature for 90 min and terminate the reaction by adding 100 µl 10 mM EDTA. This preparation was then used in hybridization reactions without any further treatment.

(ii) Prime-cut probes (Farrell *et al*, 1983)

Single-stranded prime-cut probes were used in S1 mapping experiments. The probes were generated by replicating the relevant template DNA and purifying the resulting complementary strand. Briefly,

1. Mix 8 µl template DNA, 4 µl (0.2 pmol) sequencing primer and 2 µl 100 mM Tris.HCl pH 8; 50 mM $MgCl_2$ buffer. Incubate at 55 - 60 ^{0}C for 30 min.

2. Initiate polymerisation by adding 50 mM dTTP, dGTP, dATP mix, 10 µCi [α-^{32}P]dCTP and 1 unit DNA polymerase (Klenow) to a final volume of 20 µl. Leave at room temperature for 20 min.

3. Add chase of 1 µl (50 mM cold dNTP mix), and continue reaction for 5 min.

4. Add restriction enzyme buffer and a restriction enzyme which cuts at a chosen site. Eco R1 which cuts on the 5' end of the insert was routinely used in these investigations.

After the digest, the sample is desalted by passage through a Sephadex G-50 spun column (Maniatis *et al*, 1982). A one ml syringe is filled with sterile Sephadex G-50 equilibrated in TE and spun at 16,000g for 2 min in a Centra-3 centrifuge (IEC Dunstable, Bedfordshire). Next, the sample is loaded onto the column and spun at 16,000g for 10 min. The eluate (15 - 20 µl) is then mixed with 20 µl formamide dye mix (section 2.2.4.6), placed in a boiling waterbath for 3 min and loaded onto a 3 cm wide slot in a sequencing type gel.

After electrophoresis, the gel (still on the whole glass plate) is covered with Saran wrap and autoradiographed for 5 min to locate the labelled DNA. The region corresponding to the appropriate band is cut out, and crushed into 500 µl elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% SDS, 10 µg/ml tRNA; Maxam & Gilbert, 1980). The gel slurry is kept at 37 ^{0}C for 16 - 24 h before spinning in 1 ml blue tips stoppered with nylon wool.

The resulting probe was either used directly in hybridization reactions or ethanol precipitated with carrier tRNA to concentrate the DNA. The yield was generally 1 - 3 x 10^{6} cpm (Cerenkov)

(iv) End-labelled probes

End-labelled digests of plasmid DNA were used as markers in the S1 mapping gels.

1. Plasmid DNA (1.5 µg) is digested with a restriction enzyme

that creates recessed 3′ termini at suitable positions. Sau 3A was used to digest pBR322 in these investigations as it gave a range of desired sizes (37 – 358 bp).

2. Add 10 µCi of [α-^{32}P]dCTP and 2 µl cold 500 mM dATP, dGTP, and dTTP mix.

3. Add 1 – 2 units of Klenow DNA polymerase I fragment and incubate at room temperature for 30 min.

The DNA fragments are labelled by polymerisation of the staggered ends generated by Sau 3A digestion. Labelled DNA is separated from unincorporated nucleotides by passage through the Sephadex G-50 spin column (see section 2.2.5.3.iii).

2.2.5.4. Conditions for hybridization of probes

The conditions used for Southern and dot hybridizations as well as plaque hybridizations were as described by Maniatis *et al.* (1982).

1. Prehybridize filter in hybridization solution (4 x SSC; 50% formamide; 1 x Denhardts solution; 0.1% SDS), for 1 h at 37 ^{0}C.

2. Using bag sealer, make a plastic bag round filter(s).

3. Mix 100 µl denatured Salmon sperm or Calf thymus DNA (5 mg/ml), 10 ml hybridization solution and probe. Add the mixture to the bag and seal after making sure that there are no enclosed air bubbles.

4. Shake bag in a perspex box overnight at 37 ^{0}C.

5. Remove filter and wash for 2 h in 2 x SSC, 0.1% SDS; at 37 ^{0}C.

6. Wash filter in 2 x SSC for a further 2 h.

7. Blot filter dry for 10 min at room temperature and cover with Saran wrap before autoradiography.

To increase stringency, the post-hybridization wash was scaled down to 0.5 x SSC. The probe was routinely denatured by boiling for 3 – 5 min before use (except with M13 probes).

Solution used

10 x Denhardts	0.2% BSA; 0.2% Polyvinylpyrolidine (PVP); 0.2% Ficoll. The solution was sterilized by filtration and stored at -20 ^{0}C.

2.2.5.5. S1 nuclease mapping

DNA/RNA hybrids were digested with S1 nuclease and analysed for protection of labelled DNA by total cell RNA. This was in order to determine the presence and nature of any transcript, arising from the deduced protein coding regions of sequenced DNA.

1. Mix RNA (200 - 300 µg) and single stranded labelled DNA (4 µg) with 10 x S1 hybridization buffer (0.1 M Tris.HCl pH 8; 1.5 M NaCl; 0.1 M $MgCl_2$; 1 mM EDTA), to a final volume of 100 µl. Incubate at 68 ^{0}C for 1 h to denature any secondary structures in the nucleic acids.

2. Allow tube to cool to room temperature and add 1 µl 3 M sodium acetate pH 4.5; 1 µl 0.1M $ZnSO_4$ and S1 nuclease (2 units).

3. Ethanol precipitate DNA/RNA hybrids, redissolve in 10 - 20 µl TE and add formamide loading dyes (15 µl). Boil for 3 - 5 min and electrophorese on denaturing gels.

Autoradiography was as described for sequencing gels (section 2.2.4.6)

CHAPTER 3

PS. AERUGINOSA; RMS 149 AND ITS ENCODED B-LACTAMASE

3.1. Introduction

3.1.1. *Ps. aeruginosa*

The name *Ps. aeruginosa* is given to a well defined group of bacterial isolates, examples of which are sometimes found as pathogens to animals and plants, but more commonly as inhabitants of soil or fresh water. *Ps. aeruginosa* is a small (0.5 - 1 µm by 1.5 - 4 µm), Gram-negative, polarly flagellated bacterium, classified into family I Pseudomonadaceae, section 4 of Bergy's Manual of Systematic Bacteriology (1984) . The G + C content is 67%. The bacteria are strict aerobes and, in artificial media, they develop a fluorescent diffusible greenish pigment. In humans, pseudomonads are particularly notorious in ear infections and in secondary infections of burns. In both situations, the bacteria are difficult to kill because of an inherent resistance to most chemotherapeutic agents.

Ps. aeruginosa and related species have been extensively studied because of their unusual antibiotic resistance, complex genetic organization and biochemical versatility (Holloway, 1969; 1973; Bryan *et al*, 1974; Jacoby, 1977; 1979; etc.). There is a book that covers many aspects of the organism (Clarke and Richmond, eds., 1975).

A characteristic feature of pseudomonad species is that they harbour several plasmids and bacteriophages. The phenotypic properties specified by the various plasmids include antibiotic resistance, UV protection, cell envelope functions and metabolic enzymes. Many of the plasmids are highly promiscuous as they can be transferred to widely differing bacterial species. Jacoby (1977) has grouped *Ps. aeruginosa* plasmids based on plasmid size, DNA composition and homology, susceptibility or resistance of plasmid host to certain bacteriophages as well as

plasmid-plasmid incompatibility relationships.

Gram-negative β-lactamases which are basically cephalosporinases (ref. section 1.3.3) were first characterized in *Ps. aeruginosa* (Sabath *et al*, 1965). Other types of β-lactamases specified by plasmids in pseudomonad sp have since been reported by several workers (Bryan *et al*, 1974; Jacoby, 1979; etc.).

3.1.2. Plasmid RMS 149

The RMS 149 plasmid was originally isolated by Jacoby (1977) while studying gentamicin resistant pseudomonads in a burns unit. It was characterised as Tra^+, Fi^-, (RP_1, RP_2), Phi($\overset{F}{\underset{\wedge}{1}}$16), *bla*, *aacC*$^+$ and assigned to the IncP-6 group. The plasmid is not transmissible to *E. coli*, but can be transferred to other pseudomonad spp, e.g. *Ps. fluorescens*, albeit with reduced efficiency as compared to transfer to *Ps. aeruginosa*. Its presence confers resistance to gentamicin, carbenicillin, spectinomycin, streptomycin and sulphonamide. However, the plasmid can be mobilized into *E. coli* by the use of a helper plasmid (Hedges and Matthew, 1979). It was stably maintained, conferred carbenicillin resistance but low gentamicin protection.

In a study of pseudomonad natural plasmids, Moore (1980) investigated the occurrence of restriction enzyme sites present on the RMS 149 plasmid and reported 3 Bam H1, 9 Bgl II, 6 Hind III, 14 Cla 1, 9 Pst 1, 12 Pvu II, 20 Sal 1, and 4 Sma 1 sites. By analysing the sizes of fragments generated in restriction enzyme digests, a molecular weight of 63 kb was deduced.

3.1.3. RMS 149 plasmid encoded β-lactamase

The *Ps. aeruginosa* plasmid specifies a β-lactamase which is unusual in hydrolysing carbenicillin much faster than other β-lactam antibiotics (Sawada *et al*, 1974). The enzyme was released by ultrasonication and partially purified using CM Sephadex chromatography and isoelectrofocusing. An isoelectric point of pH 7 was observed and the estimated molecular weight as determined by thin layer

gel filtration using Sephadex superfine, was 12,000 daltons.

Compared with the published data, the RMS 149 enzyme studied here, differs from already characterised β-lactamases because of its significantly lower molecular weight. Studies on the reversible unfolding of the staphylococcal β-lactamase molecule led to the suggestions that the enzyme might consist of three domains (Carrey and Pain, 1978; Carrey *et al*, 1980). The approximate molecular weight ascribed to one of these domains is about 12,000 d. Carbenicillin is a semi-synthetic β-lactam used to combat opportunistic pseudomonad infections which are prevalent in hospitals (Hamilton-Miller, 1982). Since carbenicillin had been effectively used against *Ps. aeruginosa*, an interesting question was: had the pseudomonad lost one of its domains in developing the ability to hydrolyse a previously resistant β-lactam?

Amino acid sequence provides a satisfactory method of comparing related proteins (Ambler, 1975; 1984). Initial attempts at protein purification of the RMS 149 encoded β-lactamase were not successful as the enzyme was unstable in cell lysates. Thus, with the great improvements in recombinant DNA techniques and DNA sequencing (Sanger *et al*, 1977; Maxam & Gilbert, 1977), it was considered feasible to determine the nucleotide sequence encoding the pseudomonad enzyme and thus investigate its relationship to other β-lactamases.

In the work reported here, some protein chemistry was done to confirm the properties reported in the literature for the pseudomonad β-lactamase. The gene coding for the enzyme was cloned and, in several subcloning steps, limited to a 3 kb region of the RMS 149 plasmid. The DNA sequence of this region was determined and the complete amino acid sequence of the enzyme was deduced. The similarity of the deduced sequence to that of the class A β-lactamases (Ambler, 1980) provided reassurance that the correct gene had been studied.

TABLE 3.2: MIC OF BENZYL PENICILLIN AND AMPICILLIN FOR *PS. AERUGINOSA* STRAINS

Strain	Penicillin	0.5	1	2	3	4	5	6	7	8 (mg/ml)
1920E	B.pen	+	+	+	+	+	+s	+s	+s	-
1973E	B.pen	+	+	+	+	+	+s	+s	+s	-
PA05	B.pen	+	+	+	+	+	+s	+s	-	-
1920E	Amp	+	+s	+s	-	-	-			
1973E	Amp	+	+	+s	+s	-	-			
PA05	Amp	+	+s	+s	-	-				

Notes:

B.pen = benzyl penicillin, Amp = ampicillin.

+ = growth (normal sized colonies, i.e. same size as growth without antibiotic).

+s = growth producing smaller colonies.

- = no growth.

TABLE 3.1: ANTIBIOTIC SENSITIVITY PROFILE OF *PS. AERUGINOSA* STRAINS

Antibiotic	1920E	1973E	PA05	*E. coli* ED 8654	*E. coli* ED 8654 + RMS 149*
ampicillin	R	R	R	s	R
gentamycin	R	R	s	s	r
spectinomycin	R	R	R	R	s
streptomycin	R	R	s	s	R
kanamycin	R	R	R	s	s
tetracycline	R	R	R	s	s
chloramphenicol	R	R	R	s	s
rifampicin	R	R	s	s	s
carbenicillin	R	R	s	s	R
trimethoprim**	–	–	–	–	–
sulphonamide**	–	–	–	–	–

Notes:

R=resistant; s=sensitive; r=resistance.

* *E. coli* transformed with purified RMS 149 plasmid DNA.

** Trimethoprim and sulphonamide sensitivity could not be tested as minimal plates were required and *Ps. aeruginosa* PUE1 (the host strain) is ile$^-$, val$^-$ and rif^R.

Finally, the nature of the β-lactamase gene m-RNA transcripts in *Ps. aeruginosa* harbouring the plasmid was investigated.

3.2. Antibiotic sensitivity tests

Ps. aeruginosa plasmids - RMS 149 and R151 code for a wide range of antibiotic resistance genes (Jacoby, 1977). In order to investigate the resistance profile of the plasmid strains, a known concentration of cells (100 cfu) were spread on L-agar plates containing various antibiotics. The strain PAO5 was included as a plasmid-free pseudomonad control. The results of the tests are shown in Table 3.1. Both *Ps. aeruginosa* 1920E and 1973E (carrying plasmids RMS 149 and R151 respectively) were apparently resistant to all the antibiotics tested. The high resistance level of PAO5 that was observed, conforms to other reports of the innate antibiotic resistance of *Ps. aeruginosa*.

3.3. Determination of Minimal Inhibitory Concentrations (MIC)

The MIC is defined as the lowest concentration of antibiotics which prevents visible bacterial growth. MIC values were investigated to determine the level of protection from β-lactam antibiotics conferred on *Ps. aeruginosa* 1920E by its plasmid encoded β-lactamase. Strains 1973E and PAO5 were included as reference points for protection by another β-lactamase and intrinsic cell resistance, respectively.

Cells from mid-exponential phase cultures grown in the absence of any β-lactam, were diluted in bacterial buffer and samples containing 50 - 100 cfu were spread onto L-agar plates containing a range of ampicillin concentrations. The plates were incubated overnight at 37 ^{0}C and scored for growth (+) or no growth (-).

MIC values were very high for all the pseudomonads (Table 3.2), though there seems to be no significant difference in the resistance levels of strains 1920E and PAO5. As the concentration of antibiotic increased, the colonies tended to

get smaller. This phenomenon was not observed with *E. coli* cells that were treated alongside.

3.4. Plate detection of β-lactamase activity

Routinely, $\underline{amp}^R$ colonies or plaques were selected on ampicillin plates. As bacteria can be $\underline{amp}^R$ without producing any β-lactamase (ref. section 3.2), any putative $\underline{amp}^R$ colonies were picked onto duplicate ampicillin plates, incubated overnight at 37 ^{0}C and stained for β-lactamase activity on one of the plates. Nitrocefin solution (5 mM in 0.1 M KH_2PO_4 buffer; O'Callaghan *et al.*, 1972) was spread on the petri-dish and β-lactamase producing colonies were observed to be surrounded by pink wheels against a yellow background.

3.5. Release of β-lactamase activity

In Gram-negative bacteria, the β-lactamases are located in the periplasmic space and are commonly released by ultrasonic disintegration (Dale, 1979). With thick cell suspensions however, cleared solutions were not obtained and the need to cool the suspension restricted the amount that could be sonicated at one time. In order to find an alternative method, treatments used for disrupting cells such as freeze-pressing in a French Press and lysozyme/detergent treatment (according to the method for making pseudomonad sphaeroplasts; Mizuno & Kageyama, 1979), were compared with ultrasonication. The efficiency of the different methods were compared by measuring the amount of activity released into cleared supernatants with each treatment, as well as percentage lysis achieved.

Cells harvested from a 2 l overnight growth were resuspended in 24 ml lysis cocktail (0.01 M Tris.HCl pH 8; 0.01 M EDTA; 0.005 M benzamidine; 0.005 M PMSF; Oegema *et al.*, 1975). The RMS 149 encoded β-lactamase was observed to be unstable in cell lysates. The protease inhibitors were therefore included to protect the β-lactamase in crude cell extracts from proteolysis. Seven ml

TABLE 3.4: YIELDS OF B-LACTAMASE ACTIVITY (UNITS/MG DRY WEIGHT)

Organism	Enzyme yields (units/mg dry weight) b.penicillin	carbenicillin
E. coli TG1	0.00094	n.d.*
E. coli ED 8654	0.0012	n.d.
Ps. aeruginosa PA05	0.0009	n.d.
Ps. aeruginosa 1920E	0.008	0.02
E. coli TG1 + (PRLG 304**)	0.04	0.1

Notes:

* n.d. = not detected.

** *E. coli* TG1 carrying the recombinant plasmid PRLG 304 (this thesis, section 3.9).

TABLE 3.3: RELEASE OF B-LACTAMASE ACTIVITY IN *PS. AERUGINOSA* 1920E

Treatment	Activity in units/ml b.penicillin	carbenicllin	% lysis	% effectiveness*
No treatment	n.d.	n.d.	< 2%	0
Ultrasonication	1.1	1.6	97.5%	100
French press	0.4	0.75	81.2%	36.4
Sphaeroplasts	0.6	1.1	68.5%**	54.5%

Notes:

* % effectiveness was derived by expressing the β-lactamase activity released in the cell lysates as percentages of the activity released by ultrasonication.

** This value represents the percentage of cells in a a population unable to grow, as fully formed sphaeroplast cannot divide, though, they have not lysed.

portions of the cell suspension were used for each treatment. During sphaeroplasts formation, sucrose was incorporated in the lysis cocktail to 25%. Sonication was at a wavelength of 3 - 4 microns for 30 s bursts (MSE 100W Ultrasonic Disintegrator), 4 - 5 times until there was considerable sample clearing. A viable count was determined before and after each treatment, percentage lysis was determined by expressing the number of non-viable cells as a percentage of original viable count.

The results are presented in Table 3.3. Generally, the better the lysis, the greater the enzyme activity released. Thus, ultrasonication proved to be the best method for lysing cells at the concentration used in these tests. Sphaeroplasts formation also released a significant amount of enzyme activity which, from its reaction profile (see section 3.8 below), suggests it is the plasmid encoded enzyme. This is consistent with a periplasmic location for the RMS 149 β-lactamase.

Lysis with the French press gave the lowest yields. However, it is a simpler process which can be applied to large samples without much dilution. It was thus, the process used in the rest of these investigations to release β-lactamase activity from cells.

3.6. Yield of β-lactamase activity

The yield of β-lactamase activity released in *Ps. aeruginosa* 1920E, as determined per mg dry weight was compared to that produced in strains PAO5, *E. coli* TG1 and ED8654 which do not carry any plasmids.

The presence of the RMS 149 plasmid made a difference to the amount of enzyme activity detectable in the cell lysates, as illustrated in Table 3.4. The ratio of 1:2 for the benzyl penicillin/carbenicillin activities was also observed here.

TABLE 3.5: SUBSTRATE PROFILE OF B-LACTAMASE IN *PS. AERUGINOSA* 1920E

Sample	Relative rate of hydrolysis							
	b.pen	amp	carb	met	oxa	clox	ceph	cephal
1920E lysate	100	100	240	0	12	0	32	32
1973E lysate	100	210	63	260	515	1010	35	0
TEM (50 µg/ml)	100	133	36	20	20	2	8	0
TG1 (PRLG 304*)	100	93	219	7	25	0	7	30

Notes:

* *E. coli* TG1 carrying the recombinant plasmid PRLG 304 (this thesis).

b.pen, benzyl penicillin; amp, ampicillin; carb, carbenicillin; met, methicillin; oxa, oxacillin; clox, cloxacillin; ceph, cephaloridine; cephal,cephalexin.

3.7. Isoelectricfocusing of *Ps. aeruginosa* 1920E β-lactamases

Matthew *et al.* (1979) demonstrated that the β-lactamases found in Gram-negative bacteria could be classified into six (6) groups based on their isoelectric points (pI).

For this test, 1.5 l cells of *E. coli* TG1, *Ps. aeruginosa* strains 1920E and 1973E were grown overnight, spun down, resuspended in 20 ml lysis cocktail (section 3.5) and passed through the French press. DNAse (1 mg/ml) was added to the cell lysate and the mixture was left at room temperature for 10 - 20 min. The clarified cell lysate cell was dialysed overnight against water at 4 ^{0}C, lyophilized and resuspended in a tenth of its original volume of distilled water. Samples of 10 - 15 µl containing 1 - 2 mg protein were then applied to isoelectric focusing gels.

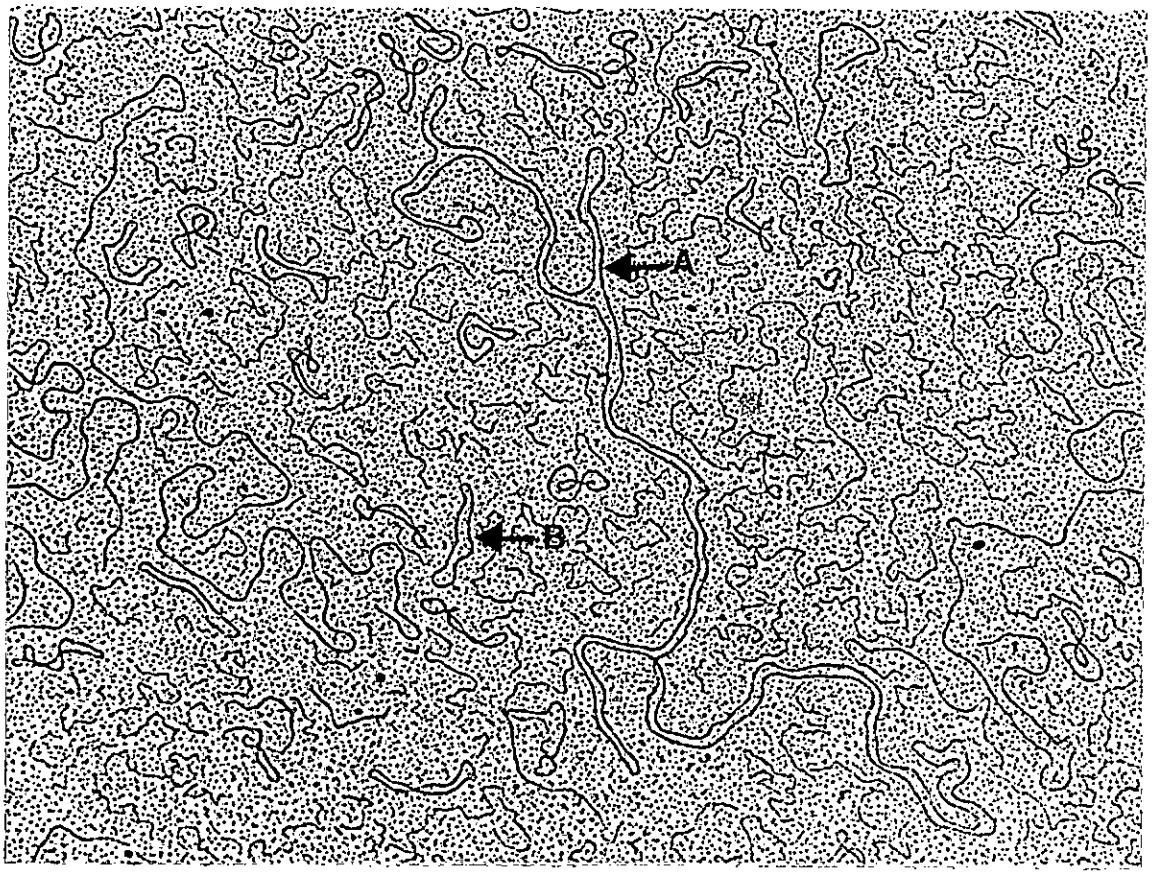
Samples from 1920E and 1973E each produced two β-lactamase bands at positions corresponding to pI values 7 and 8, and 6.3 and 8, respectively. The activity at pH 8 (also present in *E. coli* TG1 and *Ps. aeruginosa* PAO5 lysates) is the chromosomally mediated β-lactamase while the other bands are those specified by the RMS 149 and R151 plasmids respectively (Matthew, 1978).

3.8. Substrate profile of the *Ps. aeruginosa* RMS 149 β-lactamase

This was determined by assaying the β-lactamase activity in a clarified cell lysate prepared as outlined in section 3.7. above. All the substrates were used at a concentration of 7 mM and the profile was calculated relative to an arbitrary value of 100 for the rate of hydrolysis of benzyl penicillin. The *E. coli* TEM enzyme was included in the test as control. It was applied at a concentration of 50 µg/ml which gave approximately 17 international β-lactamase units/ml.

As can be deduced from the results presented in Table 3.5, the β-lactamase coded for by the RMS plasmid has a preference for penicillins but is unique in being twice as active against carbenicillin as compared to its activity against

Fig. 3.1: Electron micrograph of the RMS 149 plasmid molecule



A — RMS 149
B — pAT 153

FIG. 3.2: RESTRICTION ENZYME ANALYSIS OF RMS 149 PLASMID

Track no.	DNA	Restriction endonuclease
1	RMS 149 plasmid	undigested
2	-do-	Bam H1
3	-do-	Eco R1
4	λcI357	Hind III
5	RMS 149 plasmid	Hind III
6	-do-	Xho 1
7	-do-	Pvu II
8	-do-	Sal 1
9	λcI857	Hind III
10	RMS 149 plasmid	Sma 1
11	-do-	Pst 1

Notes:

1. Fragments in marker track (S) are derived from a Hind III digestion of λcI857.

2. Most of the fragments generated from digesting RMS 149 plasmid DNA were in the high molecular weight range so that, they could not be properly sized using electrophoretic methods.

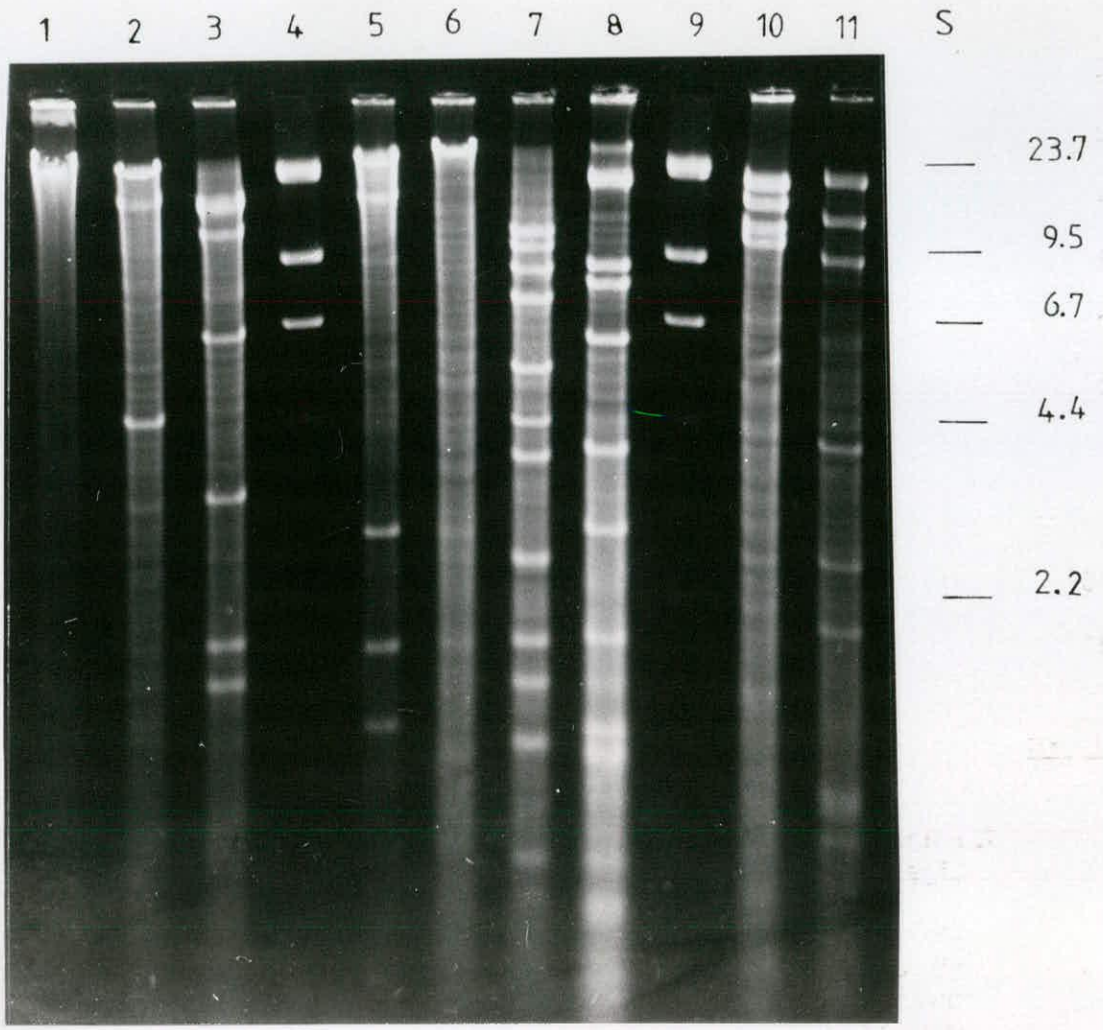


Fig. 3.2

benzyl penicillin.

3.9. RMS 149 plasmid structure

3.9.1. Isolation and sizing of the plasmid DNA

Plasmid DNA was prepared by the alkaline denaturation method from *Ps. aeruginosa* 1920E as outlined in section 2.2.2.1. The yield was routinely 0.3 mg DNA from 2.5 l cultures. The purified plasmids were visualized by electron microscopy e.g. Fig. 3.1. Since agarose gel electrophoretic sizing techniques are inaccurate in the large size range (Southern, 1979), electron microscopic methods were adopted for sizing the plasmid molecules instead. Twenty molecules were measured using a Ferranti Cetec Digitizer and the average size was 58.9 kb (1.4). (The standard deviation is shown in brackets).

3.9.2. Restriction enzyme analysis of RMS 149 DNA

The restriction fragments from Bam H1, Sal 1, Xho 1, Eco R1 and Pvu II digests of the RMS 149 plasmid were separated by agarose gel electrophoresis (Fig. 3.2).

The approximate sizes of some of the fragments were determined by the method of Southern (1979) and, as several of these were too large for proper measurement, the plasmid size could not be deduced by summing the sizes of individual fragments.

At this stage, only a rudimentary physical map of the RMS 149 could be deduced, as only the Bam H1 and Xho 1 sites could be unambiguously assigned (Fig. 3.3a). No experiments were performed to determine the copy number of the RMS plasmid. However, one would expect it to be single copied, judging from the large size.

Fig. 3.3: Cloning strategy

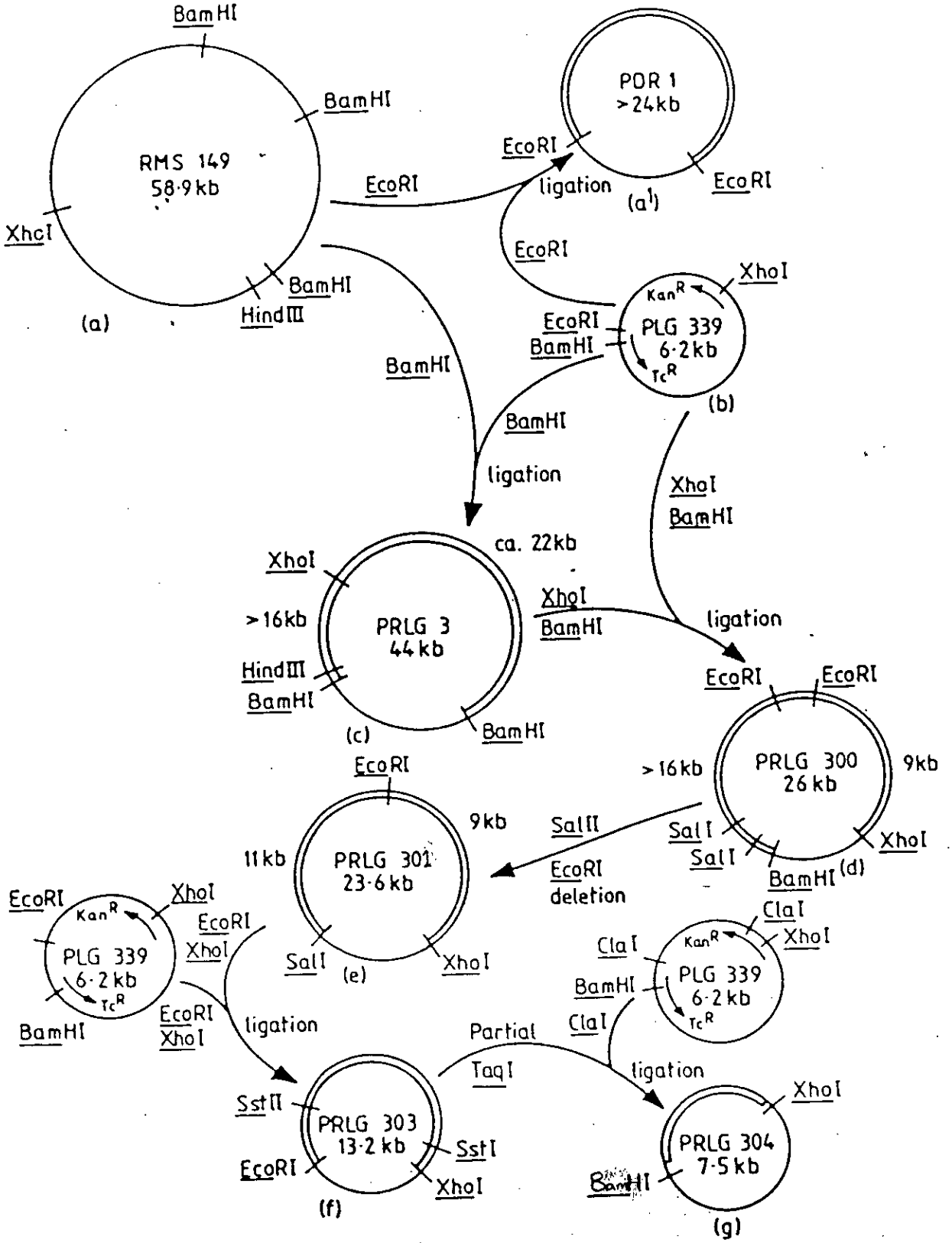


Fig. 3.3: Schematic diagram showing the construction of the recombinant plasmid pRLG 304, which carries a portion of the RMS 149 DNA. The several steps are explained in the text.

3.10. Cloning the RMS 149 plasmid β-lactamase gene

3.10.1. Choice of vector

Multiple copies of genes coding for regulatory or membrane proteins have been observed to be unstable in *E. coli* cells (Murray and Kelly, 1979; Hansen & Von Meyenberg, 1974; Spratt *et al*, 1980). As the β-lactamase gene codes for a periplasmic enzyme (Matthew, 1979), pLG 339 was chosen as the vector because it is a low copy number plasmid (6 – 8 copies/cell) and has been found to be suitable for some difficult genes (Stoker *et al*, 1982). pLG 339 (6.2 kb) carries genes conferring resistance to kanamycin and tetracycline. It also has several unique sites for hexanucleotide cutting enzymes, five of which occur within a drug resistance gene (Fig. 3.3b).

The pLG 339 plasmid does not code for ampicillin resistance, so that selection for $\underline{amp}^R$ recombinant plasmids is possible.

3.10.2. Cloning strategy

The RMS 149 β-lactamase gene was localized by recombinant DNA techniques according to the scheme in Fig. 3.3. A shotgun cloning of Bam H1 digested target and vector DNA produced the first recombinant plasmid (pRLG 3, Fig. 3.3c) with an insert of about 42 kb. The choice of Bam H1 was arbitrary, although it made it possible to relate later recombinant plasmids to the original RMS 149 physical map.

Next, a directional cloning using Xho 1 and Bam H1 double digest of vector DNA was carried out in order to determine which arm of the 42 kb Bam H1 insert carried the β-lactamase gene. The resulting plasmid pRLG 300 showed it was the longer arm (about 20 kb), which lacks the Hind III site (ref. Figs. 3.3c and 3.3d).

After careful mapping of the plasmid pRLG 300, multiple Sal 1 fragments and a 2 kb internal Eco R1 fragment were observed (Fig. 3.3d). In separate deletion

FIG. 3.4: PI OF B-LACTAMASES SPECIFIED BY RECOMBINANT PLASMID - PRLG 304

Track no.	Sample
1	*Ps. aeruginosa* 1920E cell lysate
2	*E. coli* (ED 8654) lysate
4	*Rps. sphaeroides* c_2 cytochrome
3	*E. coli* ED 8654 with pRLG 304 lysate

Notes:

1. Isoelectricfocusing was carried out as described in the text. The pH gradient was generated by application of pH 3-10 Ampholines (LKB).

2. Track 3 indicates that the *E. coli* host strain has acquired the β-lactamase band (pI, 7) by virtue of carrying the recombinant plasmid.

Fig. 3.4: pI of pRLG 304 derived β-lactamase

◀ ampC β-lactamase (pI 8)

◀ RMS 149 β-lactamase (pI 7)
◀ *Rps. sphaeroides* c_2 cytochrome

experiments which involved digestion with the appropriate enzyme, re-ligation and transformation steps, the Sal 1 and Eco R1 fragments were observed to be unnecessary for the β-lactamase gene expression. A further Eco R1 and Xho 1 directional cloning yielded pRLG 303 (Fig. 3.3f).

Attempts to generate amp^R colonies by cloning the 3.5 kb Eco R1/Sst 2 fragment of pRLG 303 were not successful. A partial Taq 1 digest of pRLG 303 plasmid was then made and cloned into the pLG 339 vector digested with Cla 1. The recombinant plasmid obtained from the resulting amp^R transformants which carried the smallest insert (3.3 kb) was later designated pRLG 304 (Fig. 3.3g).

During all these steps, plasmids were prepared as outlined in section 2.2.2.1, digested, ligated and recombinants were transformed into competent cells as described earlier (section 2.2.3). Amp^R transformants were selected by spreading the bacterial cells on LB-amp plates and transferred (with toothpicks) onto Tc or Kn plates to check for loss of any antibiotic resistance depending on the restriction enzyme being used. The resulting recombinant plasmids from bacteria having the expected phenotype, were made by the Birnboim method (section 2.2.2.1.i) and checked for the presence of the cloned DNA by separating restricted fragments on agarose gels.

In another experiment where RMS 149 was digested with Eco R1 and the fragments were ligated to the pLG 339 Eco R1 site, the recombinant plasmids pDR1 and pDR2 (Fig. 3.3a') were obtained. The two plasmids carry the approximately 18 kb Eco R1 fragment of RMS 149 in opposite orientation. This helped to determine the Eco R1 sites within the original 42 kb Bam H1 fragment.

The β-lactamase coded for by the recombinant plasmid pRLG 304 in *E. coli* was extracted as outlined earlier in section 3.6 and subjected to isoelectricfocusing (ref. Fig. 3.4). There are two β-lactamase staining bands - one corresponding to the RMS 149 plasmid encoded β-lactamase at pI 7 and another

at pI 8 which is the chromosomal enzyme. The substrate profile (Table 3.5) also agrees well with that previously determined for *Ps. aeruginosa* 1920E lysate. The yield of the enzyme activity per mg dry weight was 0.04 units/mg which is five times (5 x) the value obtained from *Ps. aeruginosa* carrying the native plasmid (Table 3.4). This correlates approximately with the increase in copy number of 6 - 8 genes per cell expected from the vector (Stoker *et al*, 1982).

3.11. Summary

The RMS 149 plasmid encoded β-lactamase is remarkable for its high activity against carbenicillin. The yield of the enzyme in *Ps. aeruginosa* 1920E was not high enough for protein purification to be successful or for molecular weight determination by SDS-PAGE. The gene coding for the β-lactamase was located to a 3.3 kb DNA fragment using recombinant DNA techniques as the β-lactamase encoded in this region had similar properties to that produced in the native plasmid.

Fig. 4.1: RMS 149 Xho 1/Bam H1 fragment, nucleotide sequence

```
TCGAGAGAGCCGATCCCGATGCAGACCATCGACCTGGCGGCCACGAGTTGCAAGGACAGT
        10                  30                  50
AGCTCTCTCGGCTAGGGCTACGTCTGGTAGCTGGACCGCCGGTGCTCAACGTTCCTGTCA

ATCACTGGCCCCGAAGACCCGACGCTCTGGAACTTCAAGAACCTCGACGACTTGTGGATG
        70                  90                  110
TAGTGACCGGGGCTTCTGGGCTGCGAGACCTTGAAGTTCTTGGAGCTGCTGAACACCTAC

GAATAGCCCCCACAAAACGACGGCCCCGTGAGGGAGCCGTTGAACATCGTCGAGCGATGC
        130                 150                 170
CTTATCGGGGGTGTTTTGCTGCCGGGGCACTCCCTCGGCAACTTGTAGCAGCTCGCTACG

CCGTCCAGGGAGGGATGGCCGAGCTCGATTCTCCAGAAGGTAGCTGGTGACACGAGTGTT
        190                 210                 230
GGCAGGTCCCTCCCTACCGGCTCGAGCTAAGAGGTCTTCCATCGACCACTGTGCTCACAA

GAGCCCACCCAACTAGGGTGACGGAAATTTCTGGGGTTCCGGCTTACATCCCCTAAACGG
        250                 270                 290
CTCGGGTGGGTTGATCCCACTGCCTTTAAAGACCCCAAGGCCGAATGTAGGGGATTTGCC

TGCCCCCCTACGGGCTACCCCCTTTATTCCACGGTTCGCTTGACCCCCTCTCCACTAGTA
        310                 330                 350
ACGGGGGGATGCCCGATGGGGGAAATAAGGTGCCAAGCGAACTGGGGGAGAGGTGATCAT

GGCGGCCCGTGGCGGCGTTTTCTCGACGTGGCCCTGCCTGCCCCCTTTCTTACCTTCTCA
        370                 390                 410
CCGCCGGGCACCGCCGCAAAAGAGCTGCACCGGGACGGACGGGGGAAAGAATGGAAGAGT

AAACGCGCTGTAGGCCCGTTTCTGGCCGGAGCCGGGCGGCGGGCGGAAAGCTCCGCTTTT
        430                 450                 470
TTTGCGCGACATCCGGGCAAAGACCGGCCTCGGCCCGCCGCCCGCCTTTCGAGGCGAAAA

CGTTCGACGGCTGGCCGACAACGAGCCGAAGGCGAGGCGTCAGCACGGTGTTTCACCAAA
        490                 510                 530
GCAAGCTGCCGACCGGCTGTTGCTCGGCTTCCGCTCCGCAGTCGTGCCACAAAGTGGTTT

AACGAGCAGCCGACCGGGCAGCGGTACGTTACGAAAACCGCTCGCATATCGTATCTGTAT
        550                 570                 590
TTGCTCGTCGGCTGGCCCGTCGCCATGCAATGCTTTTGGCGAGCGTATAGCATAGACATA

TGCGTATCATATCTGTATTGATACGATTTACACTGATACGGTACGAAAGGAGCAACGATC
        610                 630                 650
ACGCATAGTATAGACATAACTATGCTAAATGTGACTATGCCATGCTTTCCTCGTTGCTAG
                    *   V   S   V   T   R   F   S   C   R   D
```

URF-1 ⟶

```
M  A  R  Q  G  I  T  F  E  Q  V  A  A  V  A  D  A  L  A  G
ATGGCACGTCAGGGCATCACTTTTGAGCAGGTCGCCGCCGTCGCTGATGCGCTGGCAGGC
       670                       690                       710
TACCGTGCAGTCCCGTAGTGAAAACTCGTCCAGCGGCGGCAGCGACTACGCGACCGTCCG
  H  C  T  L  A  D  S  K  L  L  D  G  G  D  S  I  R  Q  C  A

E  G  Q  Q  P  T  I  R  A  V  R  E  K  L  G  D  T  G  S  P
GAGGGTCAGCAGCCGACCATCCGGGCGGTGCGGGAGAAGCTGGGCGACACCGGCAGCCCG
       730                       750                       770
CTCCCAGTCGTCGGCTGGTAGGCCCGCCACGCCCTCTTCGACCCGCTGTGGCCGTCGGGC
  L  T  L  L  R  G  D  P  R  H  P  L  L  Q  A  V  G  A  A  R

N  T  I  H  K  H  L  T  A  W  R  E  A  R  P  V  A  A  A  A
AACACTATTCACAAGCACTTGACCGCGTGGCGCGAGGCCCGCCCGGTGGCCGCAGCCGCT
       790                       810                       830
TTGTGATAAGTGTTCGTGAACTGGCGCACCGCGCTCCGGGCGGGCCACCGGCGTCGGCGA
  V  S  N  V  L  V  Q  G  R  P  A  L  G  A  R  H  G  C  G  S

A  P  E  L  P  Q  A  L  T  A  A  I  A  A  E  I  E  R  A  A
GCGCCGGAACTGCCGCAGGCGCTGACCGCTGCCATTGCCGCCGAGATCGAGCGGGCCGCA
       850                       870                       890
CGCGGCCTTGACGGCGTCCGCGACTGGCGACGGTAACGGCGGCTCTAGCTCGCCCGGCGT
  R  R  F  Q  R  L  R  Q  G  S  G  N  G  G  L  D  L  P  G  C

S  R  A  R  G  E  I  E  G  R  L  V  Q  A  Q  A  E  A  A  E
TCGCGGGCACGCGGCGAGATCGAGGGCCGTCTGGTGCAGGCGCAGGCGGAGGCCGCCGAA
       910                       930                       950
AGCGCCCGTGCGCCGCTCTAGCTCCCGGCAGACCACGTCCGCGTCCGCCTCCGGCGGCTT
  R  P  C  A  A  L  D  L  A  T  Q  H  L  R  L  R  L  G  G  F

L  A  A  A  G  E  V  L  E  G  E  R  D  E  L  A  E  Q  V  A
CTGGCCGCCGCTGGCGAGGTGCTGGAAGGCGAACGCGACGAGCTGGCCGAGCAGGTGGCC
       970                       990                      1010
GACCGGCGGCGACCGCTCCACGACCTTCCGCTTGCGCTGCTCGACCGGCTCGTCCACCGG
  Q  G  G  S  A  L  H  Q  F  A  F  A  V  L  Q  G  L  L  H  G

V  L  T  T  E  R  D  T  L  A  G  K  A  A  Q  Q  A  A  D  L
GTGCTGACCACTGAGCGCGACACGCTGGCAGGCAAAGCCGCCCAGCAGGCCGCCGACTTG
      1030                      1050                      1070
CACGACTGGTGACTCGCGCTGTGCGACCGTCCGTTTCGGCGGGTCGTCCGGCGGCTGAAC
  H  Q  G  S  L  A  V  R  Q  C  A  F  G  G  L  L  G  G  V  Q

A  E  A  Q  Q  R  I  E  E  V  F  K  G  R  L  L  M  P  K  A
GCCGAGGCGCAGCAGCGCATCGAAGAAGTTTTTAAAGGCCGCTTGCTGATGCCGAAGGCA
      1090                      1110                      1130
CGGCTCCGCGTCGTCGCGTAGCTTCTTCAAAAATTTCCGGCGAACGACTACGGCTTCCGT
  G  L  R  L  L  A  D  F  F  N  K  F  A  A  Q  Q  H  R  L  C
```

Fig. 4.1 (continued)

```
L  L  E  R  D  T  G  H  V  I  E  P  V  K  F  A  G  A  F  D
CTGCTGGAGCGGGACACAGGACACGTCATTGAGCCAGTCAAATTCGCCGGAGCGTTTGAT
      1150                 1170                 1190
GACGACCTCGCCCTGTGTCCTGTGCAGTAACTCGGTCAGTTTAAGCGGCCTCGCAAACTA
  Q  Q  L  P  V  C  S  V  D  N  L  W  D  F  E  G  S  R  K  I


A  G  Q  F  G  I  G  L  N  V  A  D  F  F  L  A  L  I  E  G
GCGGGTCAGTTCGGCATTGGCTTGAATGTAGCCGATTTTTTCCTGGCGCTGATAGAAGGC
      1210                 1230                 1250
CGCCCAGTCAAGCCGTAACCGAACTTACATCGGCTAAAAAAGGACCGCGACTATCTTCCG
  R  T  L  E  A  N  A  Q  I  Y  G  I  K  E  Q  R  Q  Y  F  A


I  G  A  P  T  K  D  V  I  V  D  K  A  R  T  A  E  R  L  G
ATCGGTGCGCCAACGAAGGATGTGATTGTAGACAAAGCGCGTACAGCCGAACGTCTGGGC
      1270                 1290                 1310
TAGCCACGCGGTTGCTTCCTACACTAACATCTGTTTCGCGCATGTCGGCTTGCAGACCCG
  D  T  R  W  R  L  I  H  N  Y  V  F  R  T  C  G  F  T  Q  A


K  Q  L  G  L  F  R  C  W  V  K  A  V  F  V  G  A  F  C  H
AAGCAACTGGGCTTGTTCAGGTGTTGGGTAAAAGCGGTATTTGTAGGCGCGTTTTGTCAT
      1330                 1350                 1370
TTCGTTGACCCGAACAAGTCCACAACCCATTTTCGCCATAAACATCCGCGCAAAACAGTA
  L  L  Q  A  Q  E  P  T  P  Y  F  R  Y  K  Y  A  R  K  T  M
                                                 ←  URF-2


V  S  H  F  I  S  F  N  V  K  I  A  I  V  *
GTATCACATTTTATTTCTTTTAATGTGAAGATAGCAATAGTTTAAAGGAGGAGCGAGAAC
      1390                 1410                 1430
CATAGTGTAAAATAAAGAAAATTACACTTCTATCGTTATCAAATTTCCTCCTCGCTCTTG


AGGGGCGGCTAACGCCGCCGCGCTATCCCTCCCCGGCATGAATGCCGGGGTTTCTCGCGC
      1450                 1470                 1490
TCCCCGCCGATTGCGGCGGCGCGATAGGGAGGGGCCGTACTTACGGCCCCAAAGAGCGCG

β-lactamase →
    M  M  K  F  Q  C  H  F  L  S  V  P  V  A  I  L  G  C  V
ATAAATGATGAAATTTCAATGTCATTTTCTCTCTGTCCCTGTGGCGATACTCGGATGTGT
      1510                 1530                 1550
TATTTACTACTTTAAAGTTACAGTAAAAGAGAGACAGGGACACCGCTATGAGCCTACACA


 G  L  I  C  T  S  A  Y  A  M  D  T  G  I  L  D  L  A  V  T
TGGCCTCATCTGCACATCCGCATATGCCATGGATACGGGAATCCTCGACCTCGCTGTAAC
      1570                 1590                 1610
ACCGGAGTAGACGTGTAGGCGTATACGGTACCTATGCCCTTAGGAGCTGGAGCGACATTG


 Q  E  E  T  T  L  Q  A  R  V  G  V  A  V  I  D  T  D  S  G
CCAAGAGGAGACCACGCTTCAGGCTAGGGTCGGAGTGGCGGTGATCGATACGGATTCCGG
      1630                 1650                 1670
GGTTCTCCTCTGGTGCGAAGTCCGATCCCAGCCTCACCGCCACTAGCTATGCCTAAGGCC
```

Fig. 4.1 (continued)

```
 L  T  W  Q  H  R  G  D  E  R  F  P  L  N  S  T  H  K  A  F
CCTGACGTGGCAGCATCGTGGCGACGAACGCTTCCCGCTGAACAGCACGCATAAAGCCTT
      1690                  1710                  1730
GGACTGCACCGTCGTAGCACCGCTGCTTGCGAAGGGCGACTTGTCGTGCGTATTTCGGAA

 S  C  A  A  V  L  A  Q  A  D  R  H  K  L  N  L  E  Q  A  I
TTCCTGCGCGGCCGTTCTGGCCCAGGCCGACCGCCACAAGCTGAACCTGGAGCAGGCGAT
      1750                  1770                  1790
AAGGACGCGCCGGCAAGACCGGGTCCGGCTGGCGGTGTTCGACTTGGACCTCGTCCGCTA

 P  I  E  R  T  A  L  V  T  Y  S  P  V  T  E  R  V  P  P  G
ACCGATCGAGCGCACAGCGCTGGTCACATACTCACCCGTGACGGAAAGGGTGCCACCTGG
      1810                  1830                  1850
TGGCTAGCTCGCGTGTCGCGACCAGTGTATGAGTGGGCACTGCCTTTCCCACGGTGGACC

 G  T  L  T  L  R  E  L  C  R  A  A  V  S  I  S  D  N  T  A
CGGCACGCTGACCCTGCGTGAGCTGTGCAGGGCCGCCGTCAGTATCAGTGACAACACAGC
      1870                  1890                  1910
GCCGTGCGACTGGGACGCACTCGACACGTCCCGGCGGCAGTCATAGTCACTGTTGTGTCG

 A  N  L  A  L  D  A  I  G  G  A  R  T  F  T  A  F  M  R  S
GGCCAATTTGGCGTTGGATGCAATCGGCGGGGCACGGACATTCACCGCGTTCATGCGGTC
      1930                  1950                  1970
CCGGTTAAACCGCAACCTACGTTAGCCGCCCCGTGCCTGTAAGTGGCGCAAGTACGCCAG

 I  G  D  D  K  T  R  L  D  R  R  E  P  E  L  N  E  A  T  P
TATCGGTGACGATAAGACACGCCTGGATCGGCGAGAACCCGAACTCAACGAGGCCACGCC
      1990                  2010                  2030
ATAGCCACTGCTATTCTGTGCGGACCTAGCCGCTCTTGGGCTTGAGTTGCTCCGGTGCGG

 G  D  A  R  D  T  T  T  P  I  A  A  A  R  S  L  Q  T  L  L
GGGGGATGCACGCGACACGACAACGCCAATTGCGGCAGCGCGGAGCCTGCAAACACTGTT
      2050                  2070                  2090
CCCCCTACGTGCGCTGTGCTGTTGCGGTTAACGCCGTCGCGCCTCGGACGTTTGTGACAA

 L  D  G  V  L  S  A  P  A  R  N  E  L  T  Q  W  M  L  G  D
GCTCGACGGTGTCCTCTCCGCTCCGGCTCGGAACGAACTGACACAATGGATGCTCGGGGA
      2110                  2130                  2150
CGAGCTGCCACAGGAGAGGCGAGGCCGAGCCTTGCTTGACTGTGTTACCTACGAGCCCCT

 Q  V  A  D  A  L  L  R  A  G  L  P  R  D  W  Q  I  A  D  K
TCAAGTTGCCGATGCCTTGCTACGCGCTGGCTTGCCGAGGGATTGGCAAATTGCGGACAA
      2170                  2190                  2210
AGTTCAACGGCTACGGAACGATGCGCGACCGAACGGCTCCCTAACCGTTTAACGCCTGTT

 S  G  A  G  G  H  G  S  R  S  I  I  A  V  V  W  P  P  K  R
GTCGGGAGCAGGTGGTCACGGATCACGTTCCATAATCGCCGTTGTCTGGCCGCCCAAGCG
      2230                  2250                  2270
CAGCCCTCGTCCACCAGTGCCTAGTGCAAGGTATTAGCGGCAACAGACCGGCGGGTTCGC
```

Fig. 4.1 (continued)

```
 S  A  V  I  V  A  I  Y  I  T  Q  T  A  A  S  M  S  A  S  N
TTCAGCCGTCATTGTGGCGATCTACATCACCCAAACCGCAGCGTCTATGTCGGCAAGCAA
      2290                 2310                 2330
AAGTCGGCAGTAACACCGCTAGATGTAGTGGGTTTGGCGTCGCAGATACAGCCGTTCGTT

 Q  A  V  S  R  I  G  S  A  L  A  K  A  L  Q  *
CCAGGCGGTGTCCAGAATCGGATCAGCCTTAGCAAAGGCGTTGCAATGAGCGTTCCCGCC
      2350                 2370                 2390
GGTCCGCCACAGGTCTTAGCCTAGTCGGAATCGTTTCCGCAACGTTACTCGCAAGGGCGG

TCAGTGCACACTTAGCGTGCTTTATTTTCCGTTTTCTGAGACGACCCCATCGACGAACTG
      2410                 2430                 2450
AGTCACGTGTGAATCGCACGAAATAAAAGGCAAAAGACTCTGCTGGGGTAGCTGCTTGAC

AACAAAACCGTGCAAGGCGTTTGGGACGACCTCAAGCCAACCGAAACGCAAGGCAAGAAG
      2470                 2490                 2510
TTGTTTTGGCACGTTCCGCAAACCCTGCTGGAGTTCGGTTGGCTTTGCGTTCCGTTCTTC

AAGGGCCAGCCGAAAGCACTACCAGTCTTTTCCGTCCATGCGGACGGCCTGTTGCTCTCC
      2530                 2550                 2570
TTCCCGGTCGGCTTTCGTGATGGTCAGAAAAGGCAGGTACGCCTGCCGGACAACGAGAGG

GTCGAGACATGGAAGAACCCACGCCGGGTGCTCAATCCGCTTTGCACACTTCAGCACGGA
      2590                 2610                 2630
CAGCTCTGTACCTTCTTGGGTGCGGCCCACGAGTTAGGCGAAACGTGTGAAGTCGTGCCT

AAAATAGCCCCATTCTGGATGCACGAAGCGTGGTTAGAAGCCAGGTGTGGGATGCGAATA
      2650                 2670                 2690
TTTTATCGGGGTAAGACCTACGTGCTTCGCACCAATCTTCGGTCCACACCCTACGCTTAT

AAGATTCATCCGTACAAATAAAGTCTCATTCTTTGAAATAGAGTTTCATCCTAAAACAGA
      2710                 2730                 2750
TTCTAAGTAGGCATGTTTATTTCAGAGTAAGAAACTTTATCTCAAAGTAGGATTTTGTCT

GGGCAAAGCCTCCTGCGGGACGGTGGCCACTTTCGGCCCTGCCGGTCGCCGTCCGTTCGT
      2770                 2790                 2810
CCCGTTTCGGAGGACGCCCTGCCACCGGTGAAAGCCGGGACGGCCAGCGGCAGGCAAGCA

GCGGCGCTGCAAGTATCGCAGCACACCCAGCGTCATGCGGGCGTCGGCAAGGGCGCGGTG
      2830                 2850                 2870
CGCCGCGACGTTCATAGCGTCGTGTGGGTCGCAGTACGCCCGCAGCCGTTCCCGCGCCAC

TGCGCCGTCTTCTGCCACGCCGCAGTCGGTCGCAGCCTCGATCAGCTTCAACCATCGCCA
      2890                 2910                 2930
ACGCGGCAGAAGACGGTGCGGCGTCAGCCAGCGTCGGAGCTAGTCGAAGTTGGTAGCGGT
```

Fig. 4.1 (continued)

```
GCGGTCGCGGCTCCGGTCGTACTCTCCATGCCACTCGGCATACATCAGCATCGCGCAAGC
      2950                 2970                 2990
CGCCAGCGCCGAGGCCAGCATGAGAGGTACGGTGAGCCGTATGTAGTCGTAGCGCGTTCG

GTCGTCAACACTGGCGCGGTGATTCCGTGTATCTGATAGGTCTGCCGCAGCAAGCGCGCA
      3010                 3030                 3050
CAGCAGTTGTGACCGCGCCACTAAGGCACATAGACTATCCAGACGGCGTCGTTCGCGCGT

TCTAAAAGCGACGTTGTAGGCAACGACGGTTCGGCCAGCAACGGCGGCGGCGTACTGCTC
      3070                 3090                 3110
AGATTTTCGCTGCAACATCCGTTGCTGCCAAGCCGGTCGTTGCCGCCGCCGCATGACGAG

GGCCACCTCGGGCCAGCTCGGCGCACTGGCCACCATGGCGTCGTGATTTTGTGAATCGCT
      3130                 3150                 3170
CCGGTGGAGCCCGGTCGAGCCGCGTGACCGGTGGTACCGCAGCACTAAAACACTTAGCGA

GTCGCCTCGACTACGCGATCATGGCGACCACACCCGTCCTGTGGATC
      3190                 3210
CAGCGGAGCTGATGCGCTAGTACCGCTGGTGTGGGCAGGACACCTAG
```

Fig. 4.1: Nucleotide sequence of the *Ps. aeruginosa* RMS 149 β-lactamase gene and flanking DNA sequences. The boxes indicate the mapped β-lactamase promoter and Shine/Dalgarno sequences. The deduced amino acid sequences of URF-1, URF-2 and the β-lactamase gene are indicated.

Fig. 4.1 (continued)

CHAPTER 4

SEQUENCE DETERMINATION OF THE RMS 149 DNA FRAGMENT

4.1. Introduction

The nucleic acid sequence of the Xho 1/Bam H1 fragment of pRLG 304 encompassing the Taq 1/Cla 1 cloned piece of RMS 149 DNA was determined. The method used for gene sequencing was the M13 dideoxy sequencing method (Sanger *et al*, 1977; Messing & Viera, 1982; Biggin *et al*, 1983) as outlined in section 2.2.4. The DNA sequences were analysed for their coding content and three large open reading frames (ORF) were identified with candidate signals for transcription and translation. The sequence of the β-lactamase gene was identified using the program DIAGON (Staden, 1982a) which compared the ORFs to the *B. licheniformis* β-lactamase I sequence held in the EMBL library.

4.2. Compilation of sequences

Individual 150 - 350 bp long gel readings of the M13 clones, generated from the pRLG 304 Xho 1/Bam H1 fragment, were transcribed into a computer and compiled as outlined in section 2.2.4.8. The consensus sequence of the major contiguous fragment is presented in Fig. 4.1. This represents DNA sequence from the Xho 1 site within the pLG 339 vector to the Bam H1 site also on the vector but on the other side of the Taq 1 generated insert. The junctions of the Taq 1/Cla 1 insert are underlined at either end. After sequencing 105 M13 clones, all the bases in the main contiguous fragment were determined on both strands, an average of 3.7 times.

4.3. Identification of genes

Advances in recombinant DNA and gene sequencing techniques have led to an accumulation of the knowledge and rules governing the protein coding content and expression signals in DNA sequences. The analysis of the sequence

Fig. 4.2: ORFs & stop codons in the RMS 149 plasmid DNA fragment

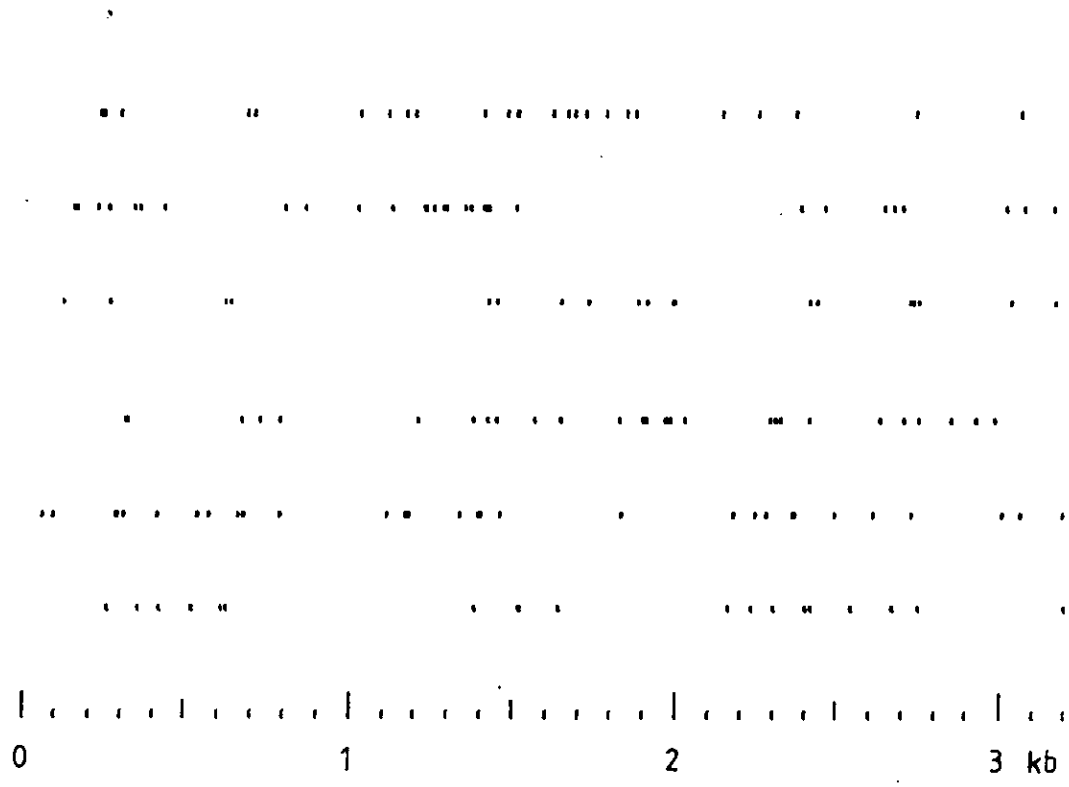


Fig. 4.2: Open reading frames predicted for the RMS 149 DNA fragment. The stop codons are marked as vertical bars along the base of each of the six reading frames. The first three are in the forward orientation, while the others represent the frames in the opposite orientation.

data from the RMS 149 DNA fragment is presented below in the context of the current state of prokaryotic gene expression.

4.4. Open reading frames (ORFs)

Potential protein coding regions were identified by searching for long ORFs i.e. reasonably long stretches of nucleotide sequences which code for amino acids without any internal stop codons. To this end the program ANALYSEQ (Staden, 1984), which has many options, including plotting stop and start codons, Shine/Dalgarno (1974) sequences and searching for *E. coli*-like promoters, was used (ref. Fig. 4.2).

4.5. Confirmation of protein coding regions

The presence of an ORF is not conclusive evidence that this stretch of DNA encodes a protein. The ORFs presented in Fig. 4.1 as translations were further studied by the use of other ANALYSEQ options (Staden, 1984).

Examination of the nucleic acid content of known coding and noncoding regions reveal that coding for a protein confers some regular properties on DNA sequences (Dayhoff, 1969; Staden, 1984a). These arise by:

- the uneven use of amino acids (some occur more frequently in proteins than others),
- the unequal number of codons for different amino acids (except for tryptophan and methionine) and
- the unequal use of the different codons for a particular amino acid.

In practice the task of coding for a protein in frame 1, for example, will influence: the amino acid content in frame 1 as well as in frames 2 and 3; the codon composition in all 3 frames; and the frequency with which each of the 4 bases occurs in each of the three positions in the codons (positional base frequencies). ANALYSEQ has the gene search by content menu which predicts coding regions using:

Table 4.1: Codon usage in *Ps. aeruginosa* RMS 149 DNA fragment

F	TTT	1.	S	TCT	2.	Y	TAT	0.	C	TGT	0.
F	TTC	3.	S	TCC	5.	Y	TAC	2.	C	TGC	2.
L	TTA	1.	S	TCA	4.	*	TAA	0.	*	TGA	0.
L	TTG	6.	S	TCG	2.	*	TAG	0.	W	TGG	4.
L	CTT	1.	P	CCT	1.	H	CAT	2.	R	CGT	4.
L	CTC	6.	P	CCC	3.	H	CAC	2.	R	CGC	6.
L	CTA	1.	P	CCA	2.	Q	CAA	6.	R	CGA	1.
L	CTG	13.	P	CCG	6.	Q	CAG	5.	R	CGG	5.
I	ATT	3.	T	ACT	0.	N	AAT	1.	S	AGT	2.
I	ATC	10.	T	ACC	6.	N	AAC	6.	S	AGC	3.
I	ATA	2.	T	ACA	8.	K	AAA	1.	R	AGA	1.
M	ATG	4.	T	ACG	10.	K	AAG	5.	R	AGG	4.
V	GTT	3.	A	GCT	5.	D	GAT	10.	G	GGT	4.
V	GTC	6.	A	GCC	13.	D	GAC	8.	G	GGC	6.
V	GTA	1.	A	GCA	8.	E	GAA	5.	G	GGA	5.
V	GTG	6.	A	GCG	14.	E	GAG	6.	G	GGG	3.

TOTAL CODONS= 265.

	T	C	A	G	
1	23.19	26.67	38.60	41.87	
2	48.55	37.08	34.50	20.33	
3	28.26	36.25	26.90	37.80	
1	12.08	24.15	24.91	38.87	
2	25.28	33.58	22.26	18.87	
3	14.72	32.83	17.36	35.09	
%	17.36	30.19	21.51	30.94	OBSERVED, OVERALL TOTALS
%	20.82	26.77	24.38	28.03	EXPECTED, EVEN CODONS PER ACID

Fig. 4.3: Positional base preference plot of RMS 149 sequence

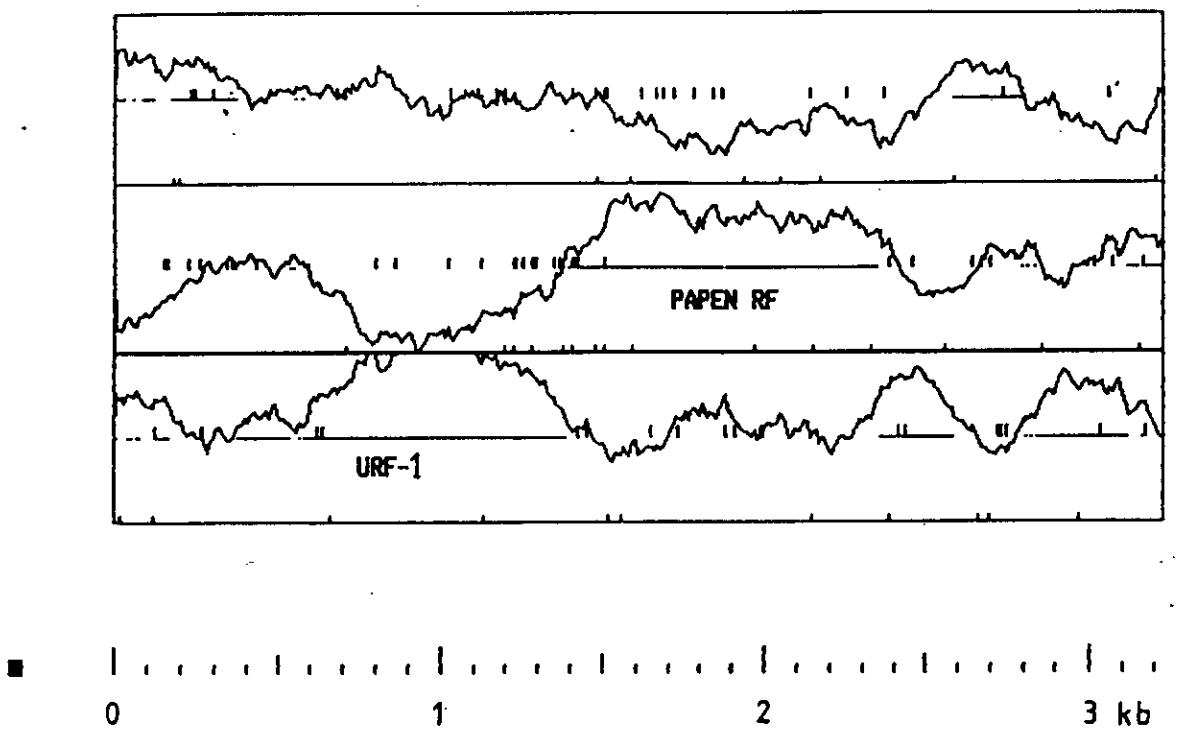


Fig. 4.3: Gene predictions for the RMS 149 plasmid DNA using the base preference method (Staden, 1984). The abscissa represents the DNA in the same orientation as presented in Fig. 4.1. Probabilities of coding are calculated by sliding a window of 23 codons along the sequence, one codon at a time. For every position of the window, the codons found in each reading frame are compared with an average calculated by Dayhoff (1969) for 314 known coding regions. A continuous line at the mid-point of a reading frame indicates which of the 3 frames is most likely to be coding. The initiation codons, ATG or GTG, are marked as vertical bars along the base of each plot and the termination codons as vertical bars along the 50% level.

Positional base preference methods

Here the effects produced in base, codon and amino acid compositions are measured in known coding sequences and the average obtained is compared to the frequencies in a new sequence. This allows for the prediction of potential coding regions. The average used is that calculated by Dayhoff (1969) for 314 protein families.

Uneven positional base frequencies method

This does not depend on an expected frequency but measures the relative abundance of each of the four bases in each of the three positions. Non-randomness in base frequency is characteristic of a DNA segment encoding a protein (Staden, 1984a).

To perform the positional base preference method, a specified window length is slid along the sequence, one codon at a time. For every position of the window, the codons found in each of the reading frames are compared with the Dayhoff (1969) standard and the corresponding probabilities of coding is thus calculated. A probability level above the 50% median line is considered to be significant (Fig. 4.3). The broken line shows possible coding regions. The base preference plot is run independently of the gene search by signal menu (which searchs for ORFs), although they have been superimposed on one another in the figure.

On the whole, the gene search by signal and content menus of ANALYSEQ strongly predicted two potential coding regions in the forward orientation. The entire RMS 149 derived fragment showed a bias in all ORFs for codons with a G and, to a lesser extent, an A in the first position (Table 4.1). This agrees with the characteristic distribution of bases in a DNA segment encoding a protein (Staden, 1984a).

4.5.1. Codon usage

The codon usage in all ORFs is summarized in Table 4.1. As was expected from a high G + C content of 61.1% (T, 17.4%; C, 30.2%, G, 30.9%, A, 21.5%), there was a preference for codons with a G or C in the third position.

FIG. 4.5: PLASMID LOCATION OF *PS. AERUGINOSA* B-LACTAMASE GENE

(A) 0.75% agarose gel electrophoresis of RMS 149 and associated recombinant plasmids.

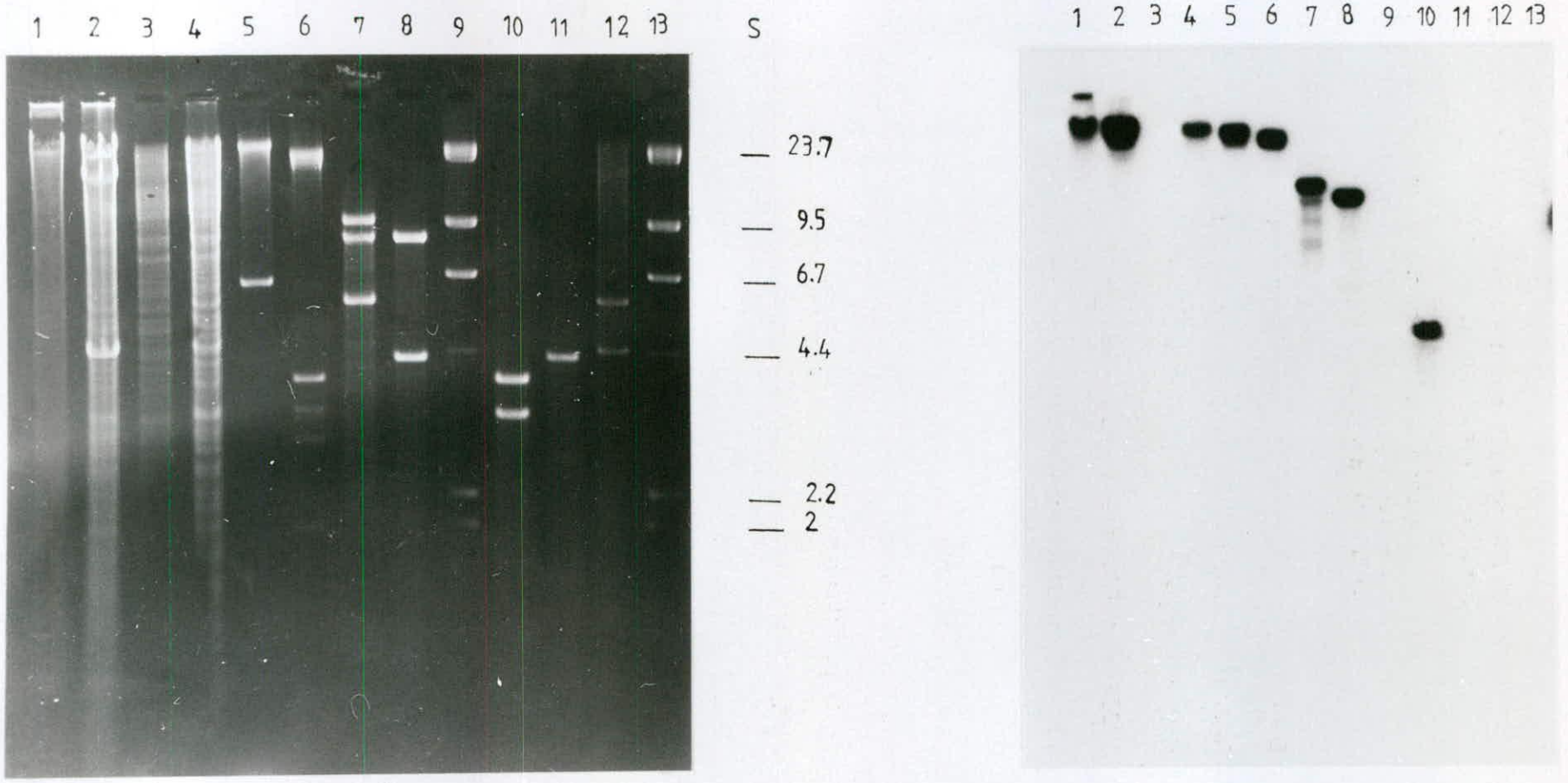
(B) Hybridization analysis of Southern transfers of DNA from gel (A) using an M13 template carrying DNA within the β-lactamase gene as probe.

Track	DNA	Restriction endonuclease
1	RMS 149 plasmid	undigested
2	RMS 149 plasmid	Bam H1
3	1973E Total DNA preparation	Bam H1
4	1920E Total DNA preparation	Bam H1
5	pRLG 3 plasmid	Bam H1
6	pRLG 300 plasmid	Bam H1/Xho 1
7	pRLG 301 plasmid	Eco R1/Sal 1
8	pRLG 303 plasmid	Eco R1/Xho 1
9	λcI857	Hind III
10	pRLG 304 plasmid	Xho 1/Bam H1
11	pACYC 187 plasmid	Bam H1
12	pss20 plasmid	Bam H1
13	λcI857	Hind III

Notes:

1. Hybridization in tracks 1, 2, 4, 5, 6, 7, 8 and 10 of Fig. 4.5B, confirm that the β-lactamase gene studied in these investigation originated from the pseudomonad RMS 149 plasmid.

2. The fragment sizes in the (S) marker tracks are derived from a Hind III digest of λcI857 phage DNA.



A

B

Fig. 4.5

Fig. 4.4: Diagonal match between *B. licheniformis* β-lactamase & papen

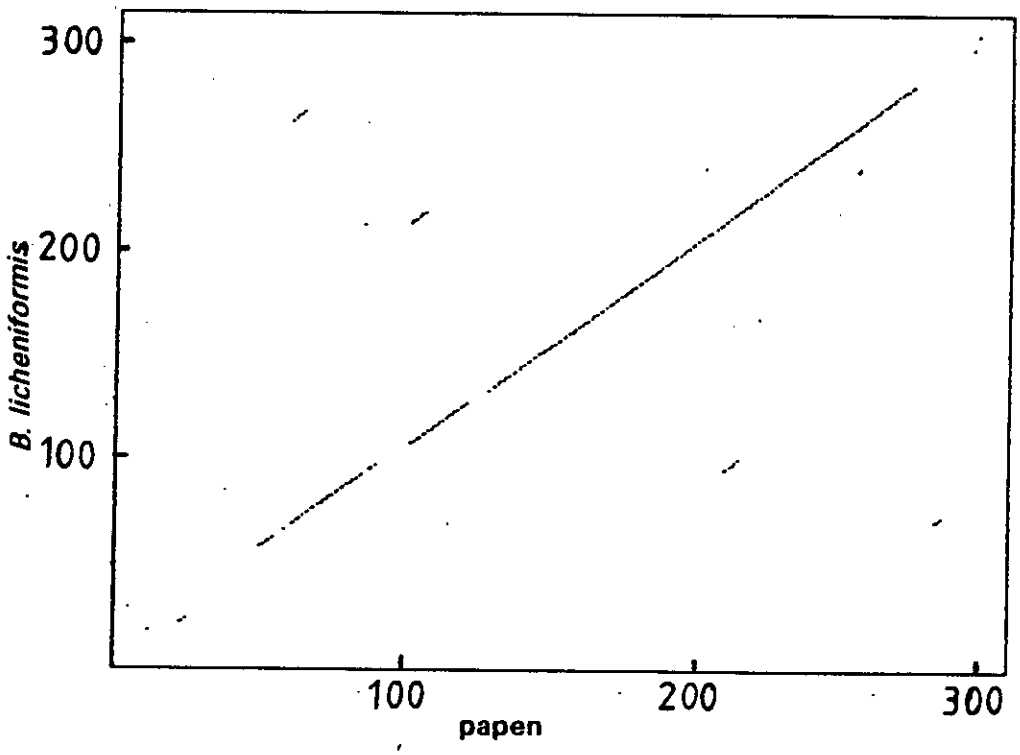


Fig. 4.4: Comparison of the protein sequences deduced for papen and *B. licheniformis* 749/C using the program DIAGON (Staden, 1982). DIAGON produces a diagram which contains a representation of all the matches between a pair of sequences. The *Bacillus* enzyme is on the vertical axis while the pseudomonad enzyme is on the horizontal axis. Each dot represents sequences, 21 amino acids long which have a double matching probability more than a calculated maximum value for random occurence of such matches.

4.5.2. Identification of the ORFs

4.5.2.1. The β-lactamase gene

The translation of the Xho 1/Bam H1 fragment in all 6 reading frames was compared with the *B. licheniformis* β-lactamase I amino acid sequence held in the EMBL library using the diagonal match program DIAGON (Staden, 1982a). There was a match within the reading frame 2 and the ORF in this region is named papen. Papen codes for a protein of molecular weight 31,283. Dot plot analysis of papen sequence and the *Bacillus* β-lactamase showed that they share extensive regions of homology and are proteins of about the same size (Fig. 4.4). The coding region starts from double Met codons at nucleotide 1505. A search for a putative Shine/Dalgarno sequence revealed one 73 bp upstream of the putative start Met codon. The potential ribosome binding site was well placed from a good *E. coli*-like promoter (Fig. 4.1).

To confirm that the β-lactamase gene comes from the RMS 149 plasmid, a M13 hybridization probe was made from a template carrying an insert which extends from nucleotide 1623 to 1871 (a segment within the β-lactamase gene), and used to probe *Ps. aeruginosa* 1920E total DNA, RMS 149 plasmid and several recombinant plasmids obtained during the subcloning procedure. The appropriate RMS 149 derived fragments were hybridized (Fig. 4.5a & b) and the location of the gene on the RMS 149 plasmid was confirmed.

4.5.2.2. Other ORFs

In the course of these investigations, some other open reading frames were encountered. They are designated below as unidentified open reading frames (URFs), as an extensive search using the program WORDSEARCH (UWGCG) against the National Biomedical Research Foundations (NBRF) protein databases failed to reveal any significant homology between them and the stored sequences.

Fig. 4.6: Graphical representation of RMS 149 β-lactamase gene

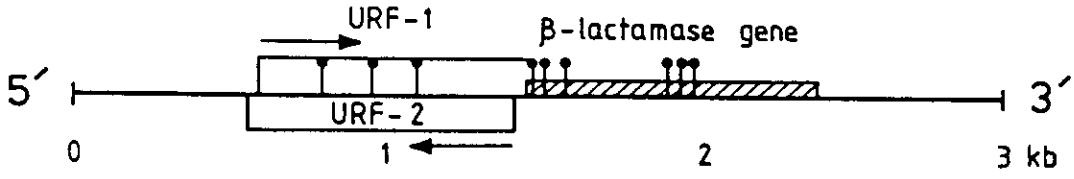


Fig. 4.6: Graphical representation of the RMS 149 β-lactamase gene and flanking sequences. The arrows indicate the direction of transcription of the genes. The symbol ⍑ stands for possible stem-loop forming sequences.

URF 1

This reading frame was convincingly predicted by ANALYSEQ (Fig. 4.3). It starts from the potential initiation codon at position 661 and extends up to nucleotide 1421, thus ending at the putative Shine/Dalgarno sequence of the downstream β-lactamase gene. URF 1 has a good Shine/Dalgarno sequence (Fig. 4.1) which is almost identical to that of the β-lactamase gene. It codes for a protein of molecular weight 27,003.

URF 2

Another long URF in the RMS 149 fragment runs in the opposite direction from nucleotide 1380 to 628. It has a poor putative Shine/Dalgarno sequence and codes for a protein of 250 amino acids (molecular weight, 27,707).

It was not, however, clear what significance these URFs had. URF 1 and URF 2 occur in the same phase, so that either of them could give rise to the other. Coding on one DNA strand affects its complementary strand (Staden, 1984a). A common observation is that, on the complementary strand in phase with the coding strand, there are often fewer stop codons than in the other two frames. The result is that one gets long open reading frames on both strands of the DNA that are exactly in phase with one another when using the uneven positional base preference method to predict coding regions. From studies with the *E. coli* unc operon and the unc-54 myosin gene of nematode (Macleod *et al*, 1981; Staden, 1984a), it has been postulated that this might be a feature of highly expressed genes.

The gene arrangement as suggested by the presence of ORFs around the *Ps. aeruginosa* RMS 149 β-lactamase gene is shown in Fig. 4.6.

TABLE 4.2: POSSIBLE SECONDARY STRUCTURES IN RMS 149 DNA TRANSCRIPTS

Stem sequences	loop size	Location	$\triangle G$ (kcal mol^{-1})
138 CGACGGCCCC 162 GUUGCCGAGG	4	outside coding region	-10.7
829 GCCGCAGC 845 CCGCGUCG	1	within URF-1	-14.6
945 GGCGGAGGCCG 972 UCGCCGCCGGU	6	within URF-1	-20.2
1083 CGAGGCGC 1102 GCUACGCG	4	within URF-1	-11.6
1442 GGGGCGGC 1461 CGCCGCCG	4	first attenuator seq.	-16.2
1471 CCCCGGCAU 1490 GGGGCCGUA	2	2nd attenuator seq.	-32.2
1552 CGGAUGUGU 1580 GCCUACACG	13	within β-lactamase gene	-6.8
1850 GUGCCACC 1866 CACGGCGG	1	first middle loop	-16.4
1869 UGACCCUGCGUG 1895 GCCGGGACGUGU	3	2nd middle loop	-9.4
1891 GGCCGCCGU 1924 CCGGCGACA	16	third middle loop	-14.4
2778 GGACGGUGGCC 2814 CCUGCCGCUGG	15	outside coding region	-13.6
2819 GUGCGGCG 2844 CACGACGC	10	outside coding region	-8.6
2875 GCGGUGUG 2903 CGCCGCAC	13	outside coding region	-6.6

Notes:

- The putative secondary structures of RNA transcribed from the RMS 149 encoded β-lactamase DNA sequence and surrounding region have been predicted using the UWGCG program STEMLOOP (it is explained in the text).
- The free energies of pairing have been calculated using the method of Tinoco *et al.* (1973), which takes into consideration stem sequences as well as nearest neighbour and loop sequence effects.

4.6. Possible secondary structures

A search for possible stem-loop structures in the DNA sequence using the program STEMLOOP (UWGCG) revealed 13 potential structures with eight or more base pairs in the stem (Table 4.2). Only four of them are in the flanking non-coding regions (these are not shown in the figure). Three occur at regular intervals of about 120 bp along URF 1, while six occur within the β-lactamase gene in two main clusters (Fig. 4.6). Two of these lie in the 73 bps between the putative Shine/Dalgarno sequence and the putative initiation codon, and the others are grouped between the nucleotides 1850 and 1924, in the middle of the β-lactamase gene. The sixth one is situated 50 bps downstream from the initiation codon. The β-lactamase gene however, ends in a poor 6 bp long stem. The latter has only two G-C bonds though it precedes a run of thymine residues.

Application of the program STEMLOOP to other DNA sequences suggests that stem-loop structures occur randomly (results not shown). Hence, any interpretations of their relevance has to be in the context of other neighbouring sequences. Secondary and tertiary structure of RNA have been implicated in the control of gene expression. Rosenberg *et al.* (1978) observed pausing by RNA polymerase at the stem-loop of λR1 terminator. This role has been proposed as the general function for the RNA hairpin structures (Adyha & Gottesman, 1978; Pribnow, 1979; Rosenberg & Court, 1979; Yanofsky,1981; etc.). Actual termination, however, requires a sufficient stretch of uridines downstream of the pause sites (Platt, 1981).

Thus, the two hairpin structures occurring after the putative Shine/Dalgarno sequence of the β-lactamase gene have the characteristics of attenuator sequences: short G-C rich palindromes followed by A-T rich regions, occurring before the main body of the structural gene (Yanofsky, 1981). The other cluster in the middle of the gene raises interesting implications. The third loop is preceded by two adenines and three thymines. It would seem likely, that although the

FIG. 4.7: S1 MAPPING REACTIONS OF RMS 149 DNA PROBES AND 1920E TRANSCRIPTS

1. Tracks TCGA are dideoxy-sequencing tracks of a template used as marker DNA. Size markers (M) are fragments derived from a Sau 3A digest of pBR322 DNA, end-labelled with Klenow fragment of DNA polymerase.

2. The numbers 5, 56 and 127 refer to probes generated from templates 5, 56 and 127 respectively. They are described in the text.

3. Tracks a contain untreated probe; tracks b, no hybridization and S1 digestion of probe; tracks c, hybridization to 1920E total RNA preparation followed by S1 digestion.

4. Arrows indicate protected fragments.

Notes:

- Probe 5 (the termination region probe) was completely protected.

- The start region probe (127; Fig. 4.7b), was largely fully protected, but a small proportion was reduced by about 100 bases. The shorter probe 127a (Fig. 4.7a) was reduced by about 19 bases.

- Probe 56 was partly fully protected while the remaining part was reduced by about 35 bps.

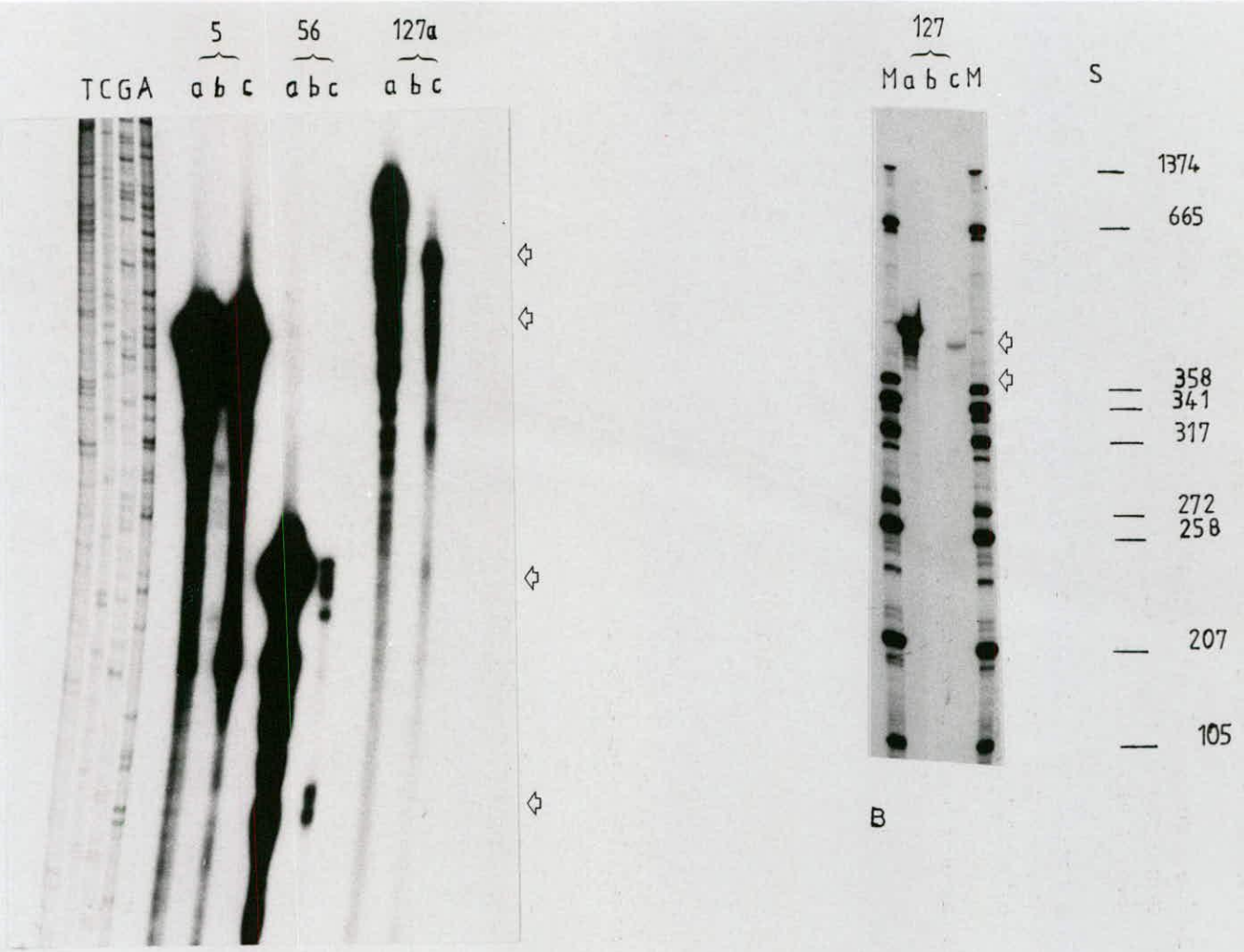

5
56
127a
TCGA
a b c
a b c
a b c
127
M a b c M
S
1374
665
358
341
317
272
258
207
105
A
B

Fig. 4.7

β-lactamase gene codes for a bigger 'normal' sized enzyme, transcription/translation may get aborted in the middle of the gene.

4.7. Repeats

A computer search of the sequenced DNA with the program REPEAT (UWGCG) did not reveal any major repetitive sequences.

4.8. Transcription of the β-lactamase gene

In order to determine the transcriptional start and stop sites of the β-lactamase gene, S1 mapping experiments with internally labelled probes were carried out as outlined in section 2.2.5.5. Three M13 templates covering different parts of the gene were used to make the prime cut probes.

Probe 127	covers DNA 128 bases upstream of the Shine/Dalgarno sequence from the Rsa 1 site at nucleotide 1302, down past the attenuator sequence to the N-terminal region of the β-lactamase at position 1782. Probe 127a is a shorter version starting 55 bases downstream the Shine/Dalgarno sequence.
Probe 56	consists of DNA occurring completely within the gene and spans the putative middle region stem-loops. It extends from the Bgl I site at nucleotide 1756 to 1918. This region was used, so as to investigate the effect of the multiple loop-forming structures in the middle of the β-lactamase gene sequence.
Probe 5	starts from within the gene at nucleotide 2269 and spans the stop codon before ending at nucleotide 2514.

All the probes were complementary to the expected mRNA sequence. The probes were hybridized to total RNA preparations from *Ps. aeruginosa* 1920E. After S1 digestions of the RNA/DNA hybrids, the protected fragments were examined on sequencing type gels for any changes in length. The results are presented in Fig. 4.7.

Two protected fragments were observed for the initiation site probe 127 (Fig.

4.7b). One of these is the fully protected probe while the other is shortened by about 100 bases (the sizes were determined by interpolation of a plot $\log_{10}$[fragment length] against distance of migration of pBR322 Sau 3A digested markers). This observation suggests that the β-lactamase gene is sometimes read through from an upstream promoter, possibly, that of URF-1 as well as from its own promoter. As the fully protected band is stronger, (ref. Fig. 4.7b), most of the transcripts are initiated from the upstream promoter. The shorter probe 127a (Fig. 4.7a) is shortened by about 19 bases (sizes here, are determined by counting the number of bands in the marker dideoxynucleotide sequencing tracks). The reduction in bases can be accounted for by the loss of the 17mer sequencing primer used in generating the probe.

Probe 56 (Fig. 4.7a) yielded two fragments – one that is full size and another which is 35 bases shorter. This observation is indicative of transcript abortion in the middle of the β-lactamase gene around the second loop.

The termination probe 5 was completely protected in these S1 mapping experiments. This would seem to suggest that β-lactamase gene transcription in the RMS 149 plasmid does get past the middle termination region, even though, final termination is not at the putative stop site.

In summary, there is preliminary evidence in these investigations, that the pseudomonad β-lactamase gene is expressed from its own promoter mapped in Fig. 4.1. Most of the transcripts, however, arise from an upstream promoter. There is termination of transcription around the middle of the gene but, the full gene sequence is also transcribed.

4.9. Conclusion

The nucleic acid sequence of the 3.3 kb Xho 1/Bam H1 fragment of pRLG 304 revealed the sequence of a 2.9 kb portion of the RMS 149 plasmid which codes for a β-lactamase, one definite URF and possibly another URF. The enzyme

closely resembles the *B. licheniformis* and TEM penicillinases which are class A enzymes (see chapter 6). The 2.9 kb insert can have arisen either by cloning a partial Taq 1 digested fragment or by multiple cloning of smaller fragments. The good sequence alignment of the β-lactamase gene to other characterised enzymes, and the presence of an internal Cla 1 site suggest that, the sequenced DNA was derived by cloning a partial Taq 1 digested fragment.

The gene sequence implies that the β-lactamase gene is not as small as reported by Sawada *et al.* (1974). This observation does not however rule out the possibility of post-translational processing, though the method used for estimating the molecular weight (thin layer gel filtration) is not reliable (R. Ambler, Pers. comm.). The presence of a cluster of putative stem-loop forming structures in the middle of the gene, suggests that there could be abortions during the transcriptional process. S1 nuclease mapping of RNA transcripts support this suggestion. It is not clear at this stage, if the truncated message produces a smaller enzyme or not. The fact that the full message is also available would suggest that the overall effect is to reduce the level of translation rather than to produce a smaller enzyme.

The low yields observed in the expression of the β-lactamase gene (section 3.6) may be due in part to the presence of an attenuator sequence. Premature termination of transcription is thought to result if significant amounts of mRNA are left unprotected by ribosomes (Pribnow, 1978). Therefore, it is possible that, the effect of the attenuator sequences is to reduce translation which in turn, cause abortions in transcription. It was not, however, clear why a constitutive enzyme should have been placed under controls normally characteristic of inducible systems.

The leader transcript of the β-lactamase cistron is unusual in lacking a second Shine/Dalgarno sequence downstream from the attenuator sequence. The

result is that the Shine/Dalgarno sequence is separated from the start codon by 73 nucleotides in the putative mRNA. This separation is rarely less than 5 or longer than 9 nucleotides (Scherer *et al*, 1980). Exceptions have been reported in literature. Thr leader sequence has a Shine/Dalgarno to Met codon space that is 12 nucleotides long (Stormo *et al*, 1982), suggesting that the ribosome can accommodate some extra nucleotides between the Shine/Dalgarno region and the initiation codon. An extreme example is the situation in T_4 gene 38 mRNA (Gold *et al*, 1981). In this case, the Shine/Dalgarno - initiation codon separation is 23 nucleotides long and an intramolecular hairpin has been hypothesised to reduce this to 5 nucleotides. Thus, secondary structure formation of the putative stem-loops in the *Ps. aeruginosa* RMS 149 β-lactamase attenuator region, may act to bring the Shine/Dalgarno sequence closer to the start codon. AUG is not absolutely required for translation (Gold *et al*. 1984). The use of GUG, UUG, AUU, and maybe UUU have been observed (Gold *et al*, 1981; 1984). Consequently, some other codon apart from the first Met may act as initiation codon.

CHAPTER 5

RHODOPSEUDOMONAS CAPSULATA SP 108 B-LACTAMASE

5.1. Introduction

5.1.1. *Rps. capsulata* sp 108

Rhodopseudomonas spp are non-sulphur, purple, Gram-negative bacteria, commonly found in stagnant bodies of water and mud. They are photosynthetic by means of bacteriochlorophyll in the presence of extraneous hydrogen donors such as alcohols, fatty acids, hydroxy- and keto-acids. Photosynthesis proceeds without the evolution of oxygen. Although most of them grow optimally as anaerobic organotrophs, rhodopseudomonads can also be cultivated in darkness under aerobic conditions. Representative species have been characterised using DNA/DNA hybridizations (Weaver *et al.*, 1975), comparative cytochrome *c* sequence analysis (Ambler *et al.*, 1979), and 16S RNA catalogues (Gibson *et al.*, 1979). The findings suggest that photosynthetic bacteria are a phylogenetically diverse group.

Rps. capsulata (Van Niel, 1944) are small (1 µm by 0.5 µm) polarly flagellated bacteria. They are closely related to two other *Rhodopseudomonas* spp - *Rps. sphaeroides* and *Rps. gelatinosa* (Pfenning, 1977). A generalized transducing phage (GTA) capable of interstrain transfer of rifampicin resistance has been reported in *Rps. capsulata* (Marrs, 1974). Wall *et al.* (1975) observed that the system was highly species specific as transfer could not be effected to other rhodopseudomonads. There have been few reports of naturally occurring plasmids (Gibson & Niedman, 1970; Saunders *et al.*, 1976; Scahill, 1981). These are very large and of several types per cell. Some *Rps. capsulata* phages confer pseudo-lysogeny (Wall *et al.*, 1975), thus raising the possibility that certain plasmids may be phage genomes.

Photosynthetically incompetent (Pho$^-$) strains were isolated after *Rps. sphaeroides* 2.4.1 had been incubated with lauryl sulphate or exposed to N-methyl-N-nitro-N'-nitrosoguanidine (Saunders *et al*, 1976). The known curing abilities of these compounds led to the suggestion that the Pho$^-$ phenotype might be due to loss of a plasmid. Electron microscopic studies revealed that the original classes of plasmids still existed though one of them had increased in size. No conjugative ability has been observed for these natural plasmids. Sistrom (1977) was able to transfer a resistance factor (R68.45) from *Ps. aeruginosa* PAO25 to *Rps. sphaeroides* and *Rps. gelatinosa*. While the plasmid conferred neomycin and carbenicillin resistance to *Rps. gelatinosa*, only neomycin resistance was observed in *Rps. sphaeroides*. Interestingly, *Rps. sphaeroides* could transfer the acquired R factor to *Rps. capsulata*, although carbenicillin resistance was not conferred (J.D. Wall, quoted in Scahill, 1981).

Rps. capsulata sp 108 was originally isolated from sewage and, in a routine antibiotic sensitivity screen, it was identified as a highly penicillin-resistant strain (Weaver *et al*, 1975). Its penicillin resistance was found to be due to a high level of β-lactamase production as a mutant was derived which no longer produced the enzyme (Weaver *et al*, 1975).

Most *Rps. capsulata* strains are highly penicillin sensitive (Weaver *et al*, 1975). There is experimental evidence that *Rhodopseudomonas* spp can receive resistance transfer factors from other Gram-negative bacteria (Sistrom, 1977; J.D. Wall, quoted in Scahill, 1981). It was thus interesting to carry out structural studies on this β-lactamase produced in photosynthetic bacteria which have no clinical significance and live in an aquatic environment. As previously stated, (section 1.6), it is unlikely that a high concentration of β-lactam antibiotics become accummulated in such niches by penicillin-producing fungi, as to exert a selective pressure on bacteria for β-lactamase production.

5.1.2. *Rps. capsulata* sp 108 β-lactamase

Preliminary studies on the β-lactamase production in the photosynthetic bacteria were carried out by Scahill (1981). With the use of isoelectricfocusing, cell extracts from strains sp 108 and sp 109 were shown to produce the same β-lactamase. They differed in that, whereas enzyme synthesis is inducible in sp 108, it is non-inducible in sp 109 and the β-lactamase is produced at a very low basal cell level. Essentially, the enzyme is a penicillinase with isoelectric points corresponding to pH values 4.5 and 4.7. These represent the cytoplasmic enzyme and the periplasmic derivative as determined by comparing isoelectricfocusing gels of whole cell lysates and cold osmotic shock (COS) released materials. The COS process consists of incubating bacterial cells in concentrated sucrose/EDTA solutions, followed by sedimentation and sudden exposure to cold water or dilute magnesium solutions. The treatment acts to release periplasmic proteins without affecting cell viability (Neu & Chou, 1967).

The structural gene was cloned from total DNA digests of *Rps. capsulata* strain sp 108 into plasmid (pMB9) and phage λ781-0 (an Eco R1 replacement vector; Murray *et al*, 1977) vectors. The gene was expressed only when on the phage. Thus, it was suggested that transcription was probably via the powerful phage promoters. Both rhodopseudomonad strains were analysed for their plasmid content and were shown to possess identical plasmid complements (four classes). Hybridization studies, using the cloned fragment as a probe, indicated that the β-lactamase gene was probably on one of these plasmids. The results of similar hybridization studies in the investigations reported here do not agree with this observation (see section 5.9). The actual coding sequences were located to a 5.8 kb Bam H1 fragment by subcloning the original 12 kb Eco R1 fragment into λNM570BV2 (a Bam H1 vector, Klein & Murray, 1979) as a preliminary step to nucleotide sequencing. An internal Bgl II site, mapped in the 5.8 kb Bam H1 fragment, was implicated as probably occurring within the gene as ampicillin

Fig. 5.2: Physical map of the *Rps. capsulata* 5.8 kb DNA fragment

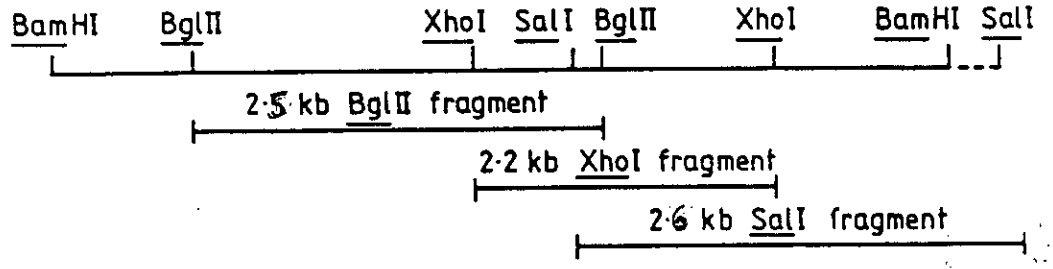


FIG. 5.1: RESTRICTION ENZYME ANALYSIS OF PSS20 PLASMID

Track	DNA	Restriction endonuclease
1	pSS20 plasmid DNA	Xho 1/Sal 1
2	-do-	Bam H1/Xho 1
3	λcI857	Hind III
4	pSS20 plasmid DNA	Xho 1
5	-do-	Bgl II/Xho 1
6	-do-	Bgl II
7	-do-	Bam H1/Bgl II
8	λcI857	Hind III
9	pSS20 plasmid DNA	Bam H1
10	-do-	Bam H1/Sal 1
11	-do-	Sal 1
12	-do-	Bgl II/Sal 1

Note:

The fragments in the marker track (S) are derived from digesting λcI857 with Hind III.

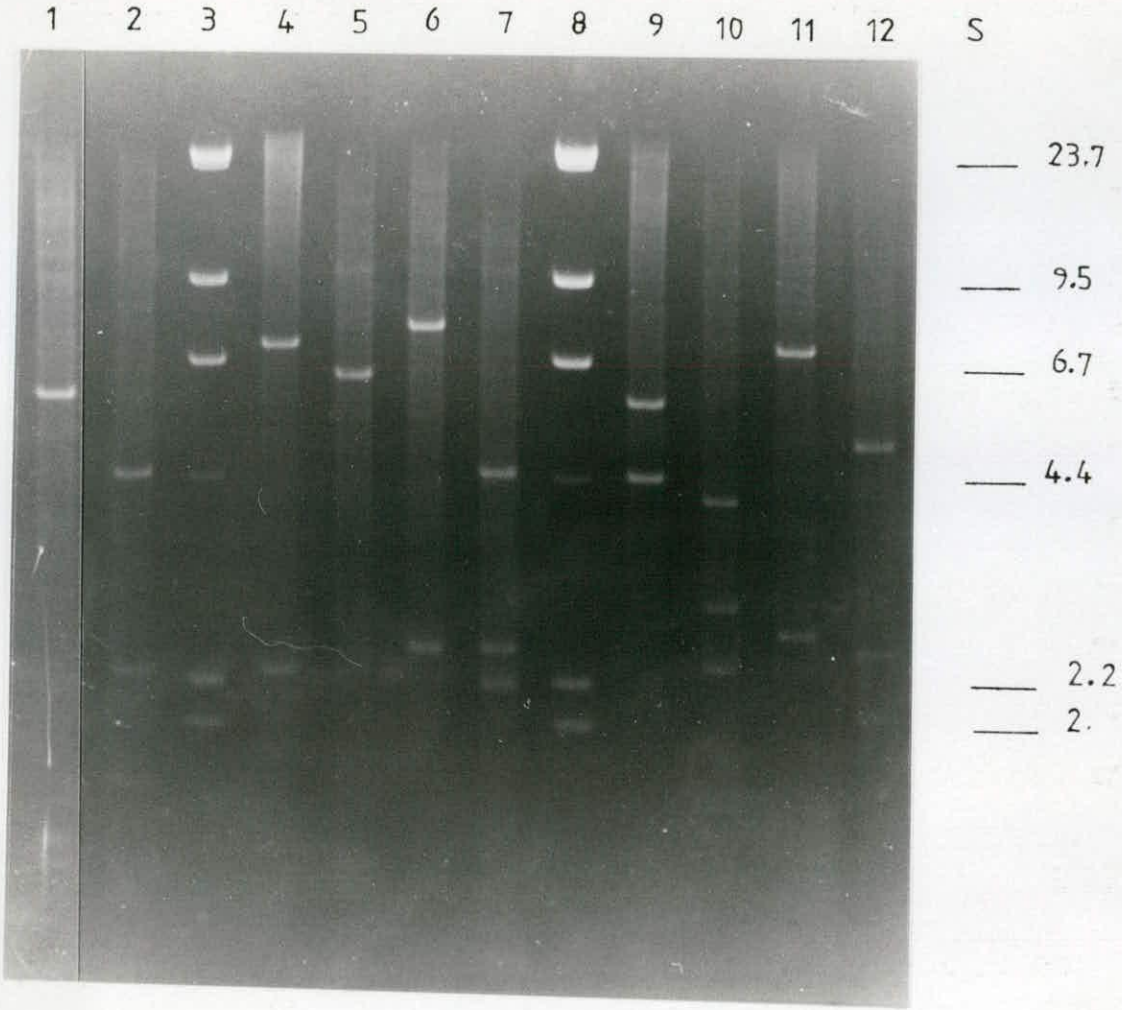


Fig. 5.1

resistant plaques could not be obtained with Bam H1/Bgl II generated fragments.

In the course of the continuing investigation of the primary structure of the β-lactamase, the recombinant plasmid pSS20 (pBR322 carrying *Rps. capsulata* sp 108 derived 5.8 kb Bam H1 fragment) was physically mapped. Several fragments spanning the whole length of the 5.8 kb DNA were cloned into a λ vector and the gene was further located to a 2.3 kb Sal 1 insert. The nucleic acid sequence of the rhodopseudomonad DNA was determined and the amino acid sequence of the *Rps. capsulata* sp 108 β-lactamase was deduced. The nature of the β-lactamase gene transcript was also examined to determine the transcriptional start and stop signals in photosynthetic bacteria.

5.2. pSS20 plasmid

The plasmid pSS20 is derived from an insertion of the *Rps. capsulata* 5.8 kb fragment into the Bam H1 site of pBR322 (Scahill, 1981). The plasmid was mapped by single and double restriction enzyme digestion and agarose electrophoretic resolution as outlined in section 2.2.3 (Fig. 5.1). The restriction enzyme sites and approximate fragment sizes deduced for the different fragments in the Bam H1 insert are shown in Fig. 5.2. The Sal 1 fragment was generated because of a flanking Sal 1 site within the pBR322 DNA only 275 bps from the Bam H1 junction site (Sutcliffe, 1978).

5.3. Subcloning the derivatives into plasmid vector

In order to locate the rhodopseudomonad β-lactamase gene more closely, the 5.8 kb Bam H1, 2.5 kb Bgl II, 2.6 kb Sal 1 and 2.2 kb Xho 1 generated fragments (Fig. 5.2) were purified from the rest of the pBR322 sequences as explained in section 2.2.3. The purified fragments were then separately cloned into the Bam H1 cut pLG401 vector DNA (Fig. 5.3) for the Bam H1 and the Bgl II fragments or into the Sal 1 site of the same vector for the Sal 1 and Xho 1 fragments.

pLG 401 is an expression vector constructed in the course of this work. It

Fig. 5.3: Plasmid pLG 401 construction

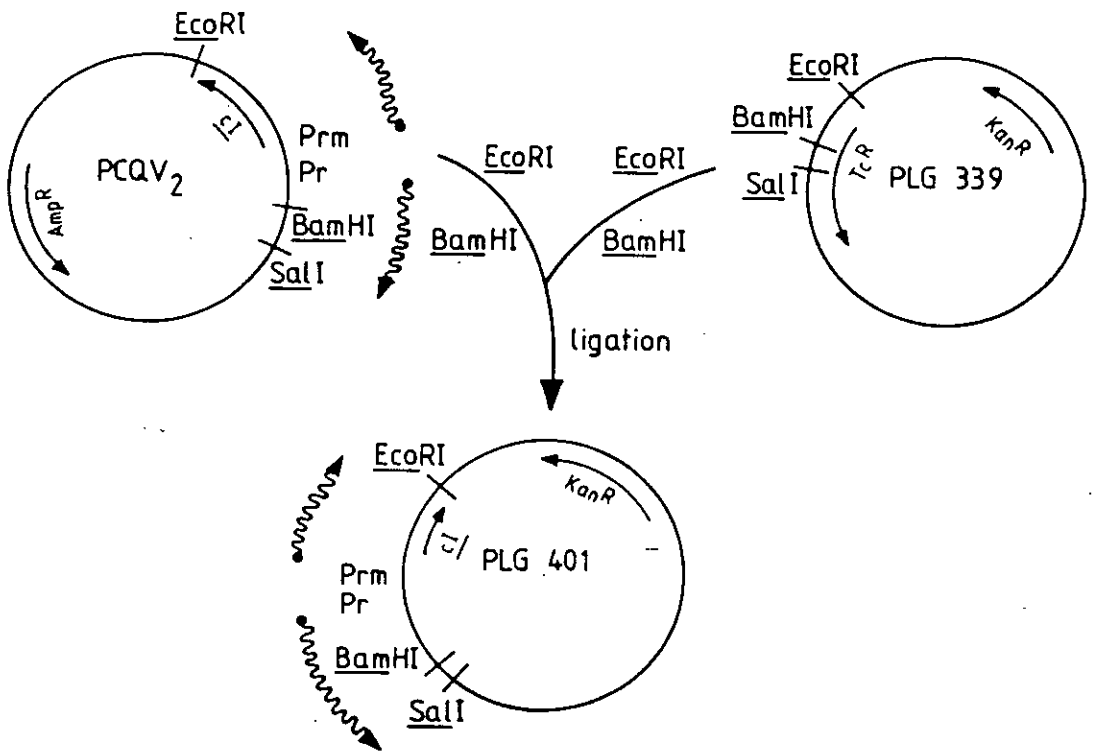


Fig. 5.3: Diagramatic representation of the scheme used in generating pLG 401. The Eco R1/Bam H1 fragment of pLG 339 is replaced by that from pCQV2 (Queen, 1983).

was derived by replacing the Eco R1/Bam H1 fragment in the plasmid pLG 339 vector DNA with an Eco R1/Bam H1 fragment from an expression vector PCQV2 (Queen, 1983). DNA sequences upstream and spanning the beginning of the pLG 339 Tc^R gene are thus replaced by that encoding a ts cI gene and the λPr promoter(Fig. 5.3). Subsequent DNA cloned into the Bam H1 site or the nearby Sal 1 site, can then be transcribed from the λPr promoter at the non-permissive temperature of 42 ^{0}C.

Ampicillin resistant colonies which produce β-lactamase were observed for *E. coli* transformed with recombinant plasmid DNA carrying either the Bam H1 fragment (positive control) or the Sal 1 piece. The recombinant plasmids resulting from the Sal 1 cloning were recovered by the minipreparation method (Birnboim & Doly, 1979) and subjected to restriction enzyme analysis. Though the 2.6 kb Sal 1 was not resolved (Fig. 5.9a), the plasmid was later observed to hybridize to a radioactive probe made from the *Rps. capsulata* β-lactamase gene (see section 5.9 below). This plasmid- pSS200 was not used in further investigations as the inability to excise the cloned Sal 1 fragment, suggests there might have been some gene rearrangement.

5.4. Subcloning the 5.8 kb Bam H1 fragment into λcI857

Xho 1 and Sal 1 fragments prepared as outlined in section 5.3 above were ligated into the Sal 1 site of λcI857 Sam in separate experiments. The two Sal 1 sites on λ DNA occur within 0.6 kb of each other within the gam gene (Sanger *et al*, 1982; Daniels *et al*, 1983). This is part of the replaceable inessential region. Since only 0.6 kb is lost in a Sal 1 digest, the insert on ligation must be less than 3 kb for efficient phage packaging (Murray, 1983). This is because λ will only package DNA that is 80 - 105% of its full genomic size (Weil *et al*, 1973). Thus, both target fragments were within the desired range.

β-lactamase producing phages were obtained only with recombinant phage

Table 5.1: Codon usage table in *Rps. capsulata* β-lactamase gene

F	TTT	0.	S	TCT	3.	Y	TAT	0.	C	TGT	1.	
F	TTC	7.	S	TCC	6.	Y	TAC	1.	C	TGC	2.	
L	TTA	0.	S	TCA	0.	*	TAA	0.	*	TGA	0.	
L	TTG	1.	S	TCG	5.	*	TAG	0.	W	TGG	6.	
L	CTT	1.	P	CCT	3.	H	CAT	1.	R	CGT	3.	
L	CTC	19.	P	CCC	5.	H	CAC	3.	R	CGC	12.	
L	CTA	0.	P	CCA	1.	Q	CAA	0.	R	CGA	1.	
L	CTG	15.	P	CCG	6.	Q	CAG	3.	R	CGG	4.	
I	ATT	0.	T	ACT	1.	N	AAT	2.	S	AGT	0.	
I	ATC	11.	T	ACC	8.	N	AAC	6.	S	AGC	7.	
I	ATA	0.	T	ACA	1.	K	AAA	1.	R	AGA	0.	
M	ATG	9.	T	ACG	12.	K	AAG	5.	R	AGG	3.	
V	GTT	3.	A	GCT	3.	D	GAT	4.	G	GGT	3.	
V	GTC	12.	A	GCC	14.	D	GAC	15.	G	GGC	17.	
V	GTA	0.	A	GCA	2.	E	GAA	5.	G	GGA	3.	
V	GTG	9.	A	GCG	20.	E	GAG	18.	G	GGG	5.	

TOTAL CODONS= 308.

	T	C	A	G	
1	21.77	24.68	45.83	41.43	
2	59.18	28.85	44.44	20.87	
3	19.05	46.47	9.72	37.69	
1	10.39	25.00	21.43	43.18	
2	28.25	29.22	20.78	21.75	
3	9.09	47.08	4.55	39.29	
%	15.91	33.77	15.58	34.74	OBSERVED, OVERALL TOTALS
%	22.02	24.33	22.67	30.97	EXPECTED, EVEN CODONS PER ACID

Fig. 5.5: Nucleotide sequence of *Rps. capsulata* β-lactamase gene and flanking sequences

```
GGATCCAGAACTTCTACAACGCCCATCGGCTATCTCCCGCGACGGGCGCCAGGAGCCACT
        10                  30                  50
CCTAGGTCTTGAAGATGTTGCGGGTAGCCGATAGAGGGCGCTGCCCGCGGTCCTCGGTGA


CTCCCAGCCCTCAACCCGTCGCCGGGAAATCGACACTCCCTTACTCTGGCTGATGTGAAG
        70                  90                  110
GAGGGTCGGGAGTTGGGCAGCGGCCCTTTAGCTGTGAGGGAATGAGACCGACTACACTTC


GGCGGGAAGCGGACCTTCGCCGCGCTTGGGTCGAACGTCGCGGAAGGGCCGTCTCTGCTG
        130                 150                 170
CCGCCCTTCGCCTGGAAGCGGCGCGAACCCAGCTTGCAGCGCCTTCCCGGCAGAGACGAC


GGCAGACCCCACTCGCTTGGGGGGCAGCAAGCTGTGCCGCCGCAACTCTGCTTCGAGCCC
        190                 210                 230
CCGTCTGGGGTGAGCGAACCCCCCGTCGTTCGACACGGCGGCGTTGAGACGAAGCTCGGG


TTTCGAGCCCTTCGTGTCGGTTTTGGGCTTCATGTTGCATGACGGGCATCCAAGTGGGTG
        250                 270                 290
AAAGCTCGGGAAGCACAGCCAAAACCCGAAGTACAACGTACTGCCCGTAGGTTCACCCAC


CCTGCCAATGTCCGCGGCGCTTAGCGCGCACGGCAGGCTTGGGGGGGGGGTGACACGAAT
        310                 330                 350
GGACGGTTACAGGCGCCGCGAATCGCGCGTGCCGTCCGAACCCCCCCCCCACTGTGCTTA


TGGATGCTTACTCAGCGCTCTGGCCGGTCAGCCAACGCGAGAACGTGCTCATCGCTGATG
        370                 390                 410
ACCTACGAATGAGTCGCGAGACCGGCCAGTCGGTTGCGCTCTTGCACGAGTAGCGACTAC
          *  E  A  S  Q  G  T  L  W  R  S  F  T  S  M  A  S


TGGCGGGCCGGTCGGACGGCCAGGCGAGATAGTAACGTCCGACAGAAACTGTGACGCCGA
        430                 450                 470
ACCGCCCGGCCAGCCTGCCGGTCCGCTCTATCATTGCAGGCTGTCTTTGACACTGCGGCT
T  A  P  R  D  S  P  W  A  L  Y  Y  R  G  V  S  V  T  V  G


AAGGCTGTGCCAAGCGGCCCTGCGCGATGTAGCTCTCGAACATGGAGATCGGCAGCAGGG
        490                 510                 530
TTCCGACACGGTTCGCCGGGACGCGCTACATCGAGAGCTTGTACCTCTAGCCGTCGTCCC
F  P  Q  A  L  R  G  Q  A  I  Y  S  E  F  M  S  I  P  L  L


CGACGCCGGCGCCTGAGGTTGCCAGCTCTGCCAAAGCCACCGAGCTGTCAAAGACCGGCC
        550                 570                 590
GCTGCGGCCGCGGACTCCAACGGTCGAGACGGTTTCGGTGGCTCGACAGTTTCTGGCCGG
A  V  G  A  G  S  T  A  L  E  A  L  A  V  S  S  D  F  V  P


CGGTGACAGGCGGGCAAGGCACGCCTGCGGCCTCAAACCAGCCCGGCCATTCGGCGCTTC
        610                 630                 650
GCCACTGTCCGCCCGTTCCGTGCGGACGCCGGAGTTTGGTCGGGCCGGTAAGCCGCGAAG
G  T  V  P  P  C  P  V  G  A  A  E  F  W  G  P  W  E  A  S
```

```
GGTAGCTGCGCAGAAGCGTTACTTGCCCGAGGTCCGAGGGATGCAGCAGCCGTGACGCGA
      670                 690                 710
CCATCGACGCGTCTTCGCAATGAACGGGCTCCAGGCTCCCTACGTCGTCGGCACTGCGCT
R  Y  S  R  L  L  T  V  Q  G  L  D  S  P  H  L  L  R  S  A


GGCCCGGCGCGCAGAGCGGCGCCATCGGCGCCTCTGCCAGTGGGATCGCATCGTGTCCTG
      730                 750                 770
CCGGGCCGCGCGTCTCGCCGCGGTAGCCGCGGAGACGGTCACCCTAGCGTAGCACAGGAC
L  G  P  A  C  L  P  A  M  P  A  E  A  L  P  I  A  D  H  G

                                    PSSURF-1 ⟶
                                    M  S  S  P  S  R  R  I  S  T  R
TCCAGCCTCCAGTGCCGAAACGGATCGCCATGTCGAGCCCCTCGCGCAGGATCTCCACCC
      790                 810                 830
AGGTCGGAGGTCACGGCTTTGCCTAGCGGTACAGCTCGGGGAGCGCGTCCTAGAGGTGGG
T  W  G  G  T  G  F  R  I  A  M  D  L  G  E  R  L  I  E  V


  L  L  L  V  E  I  L  R  S  I  W  G  W  A  C  R  N  A  S  S
GGTTGTTGTTGGTCGAGATCCTCAGGTCGATCTGGGGATGGGCCTGTCGGAACGCCTCGA
      850                 870                 890
CCAACAACAACCAGCTCTAGGAGTCCAGCTAGACCCCTACCCGGACAGCCTTGCGGAGCT
R  N  N  N  T  S  I  R  L  D  I  Q  P  H  A  Q  R  F  A  E


  R  G  I  S  Q  H  I  A  K  V  V  L  T  P  T  F  S  T  S  M
GCCGCGGCATCAGCCAGCACATCGCGAAGGTCGTGTTGACGCCAACCTTTAGCACCTCGA
      910                 930                 950
CGGCGCCGTAGTCGGTCGTGTAGCGCTTCCAGCACAACTGCGGTTGGAAATCGTGGAGCT
L  R  P  M  L  W  C  M  A  F  T  T  N  V  G  V  K  L  V  E


  S  R  R  P  P  R  R  S  S  T  R  D  I  A  S  N  P  C  S  S
TGTCGCGCCGCCCGCCGAGACGGTCGAGCACCCGCGACATCGCGTCGAACCCATGCTCAA
      970                 990                1010
ACAGCGCGGCGGGCGGCTCTGCCAGCTCGTGGGCGCTGTAGCGCAGCTTGGGTACGAGTT
I  D  R  R  G  G  L  R  D  L  V  R  S  M  A  D  F  G  H  E


  T  G  K  S  R  R  P  S  S  V  G  I  S  P  W  E  V  L  R  N
GCACGGGAAAGAGCAGGCGTCCCTCGTCGGTCGGGATCAGCCCCTGGGAGGTCCTCAGGA
     1030                1050                1070
CGTGCCCTTTCTCGTCCGCAGGGAGCAGCCAGCCCTAGTCGGGGACCCTCCAGGAGTCCT
L  V  P  F  L  L  R  G  E  D  T  P  I  L  G  Q  S  T  R  L


  S  A  V  P  S  R  S  S  S  R  A  T  W  W  L  T  A  A  W  V
ACAGCGCTGTGCCCAGCAGGTCCTCGAGCCGCGCGACCTGGTGGCTGACGGCAGCCTGGG
     1090                1110                1130
TGTCGCGACACGGGTCGTCCAGGAGCTCGGCGCGCTGGACCACCGACTGCCGTCGGACCC
F  L  A  T  G  L  L  D  E  L  R  A  V  Q  H  S  V  A  A  Q


  T  R  S  S  I  A  A  L  V  K  L  P  W  R  M  A  T  S  N  T
TGACGCGCAGCTCGATCGCGGCCTTGGTAAAGCTGCCCTGGCGCATGGCGACCTCGAATA
     1150                1170                1190
ACTGCGCGTCGAGCTAGCGCCGGAACCATTTCGACGGGACCGCGTACCGCTGGAGCTTAT
T  V  R  L  E  I  A  A  K  T  F  S  G  Q  R  M  A  V  E  F
```

Fig. 5.5 (continued)

```
  R  R  A  L  S  G  R  S  G  R  S  M  A  V  P  Q  L  M  A  G
CGCGCAGGGCATTGAGCGGCAGGTCGGGGCGATCCATGGCAGTACCTCAGCTAATGGCAG
      1210                 1230                 1250
GCGCGTCCCGTAACTCGCCGTCCAGCCCCGCTAGGTACCGTCATGGAGTCGATTACCGTC
V  R  L  A  N  L  P  L  D  P  R  D  M
                  ←—— PSSURF-2

  H  K  I  C  R  L  M  A  A  I  K  P  K  Y  T  V  R  M  T  S
GCCATAAGATTTGTCGTTTGATGGCTGCGATCAAGCCGAAATATACAGTCCGCATGACAT
      1270                 1290                 1310
CGGTATTCTAAACAGCAAACTACCGACGCTAGTTCGGCTTTATATGTCAGGCGTACTGTA

                                                  β-lactamase →
  C  D  V  A  S  Q  A  P  L  A  Q  I  L  K  G  M  V  P  C  G
                                                        M  R
CATGTGATGTCGCATCCCAGGCGCCACTGGCGCAGATCCTGAAAGGAATGGTCCCATGCG
      1330                 1350                 1370
GTACACTACAGCGTAGGGTCCGCGGTGACCGCGTCTAGGACTTTCCTTACCAGGGTACGC

  S  P  L  P  S  C  R  V  S  R  Q  G  S  L  S  A  C  P  W  P
  F  T  A  T  V  L  S  R  V  A  T  G  L  A  L  G  L  S  M  A
GTTCACCGCTACCGTCCTGTCGCGTGTCGCGACAGGGCTCGCTCTCGGCCTGTCCATGGC
      1390                 1410                 1430
CAAGTGGCGATGGCAGGACAGCGCACAGCGCTGTCCCGAGCGAGAGCCGGACAGGTACCG

  R  P  P  S  P  K  R  L  S  R  R  S  P  K  P  S  P  G  S  R
  T  A  S  L  A  E  T  P  V  E  A  L  S  E  T  V  A  R  I  E
CACGGCCTCCCTCGCCGAAACGCCTGTCGAGGCGCTCTCCGAAACCGTCGCCCGGATCGA
      1450                 1470                 1490
GTGCCGGAGGGAGCGGCTTTGCGGACAGCTCCGCGAGAGGCTTTGGCAGCGGGCCTAGCT

  N  S  S  A  P  A  S  A  S  R  S  W  R  P  A  R  V  G  P  G
  E  Q  L  G  A  R  V  G  L  S  L  M  E  T  G  T  G  W  S  W
GGAACAGCTCGGCGCCCGCGTCGGCCTCTCGCTCATGGAGACCGGCACGGGTTGGTCCTG
      1510                 1530                 1550
CCTTGTCGAGCCGCGGGCGCAGCCGGAGAGCGAGTACCTCTGGCCGTGCCCAACCAGGAC

  L  T  A  R  T  S  F  S  S  *
  S  H  R  E  D  E  L  F  L  M  N  S  T  V  K  V  P  V  C  G
GTCTCACCGCGAGGACGAGCTTTTCCTCATGAACAGCACGGTCAAGGTGCCGGTCTGCGG
      1570                 1590                 1610
CAGAGTGGCGCTCCTGCTCGAAAAGGAGTACTTGTCGTGCCAGTTCCACGGCCAGACGCC

  A  I  L  A  R  W  D  A  G  R  L  S  L  S  D  A  L  P  V  R
CGCCATCCTCGCGCGTTGGGACGCGGGCAGGCTGTCGCTCTCCGATGCGCTGCCGGTGCG
      1630                 1650                 1670
GCGGTAGGAGCGCGCAACCCTGCGCCCGTCCGACAGCGAGAGGCTACGCGACGGCCACGC
```

Fig. 5.5 (continued)

```
 K  A  D  L  V  P  Y  A  P  V  T  E  T  R  V  G  G  N  M  T
CAAGGCCGACCTCGTGCCCTACGCGCCCGTCACGGAGACGCGGGTCGGCGGCAACATGAC
      1690                 1710                 1730
GTTCCGGCTGGAGCACGGGATGCGCGGGCAGTGCCTCTGCGCCCAGCCGCCGTTGTACTG

 L  D  E  L  C  L  A  A  I  D  M  S  D  N  V  A  A  N  I  L
CCTCGACGAGCTCTGCCTCGCGGCGATCGACATGAGCGACAATGTGGCGGCGAACATCCT
      1750                 1770                 1790
GGAGCTGCTCGAGACGGAGCGCCGCTAGCTGTACTCGCTGTTACACCGCCGCTTGTAGGA

 I  G  H  L  G  G  P  E  A  V  T  Q  F  F  R  S  V  G  D  P
GATCGGGCATCTCGGGGGGCCGGAGGCGGTGACGCAGTTCTTCCGCAGCGTCGGCGACCC
      1810                 1830                 1850
CTAGCCCGTAGAGCCCCCCGGCCTCCGCCACTGCGTCAAGAAGGCGTCGCAGCCGCTGGG

 T  S  R  L  D  R  I  E  P  K  L  N  D  F  A  S  G  D  E  R
GACGAGCCGTCTCGACCGCATCGAGCCCAAGCTGAACGACTTCGCTTCTGGAGACGAGCG
      1870                 1890                 1910
CTGCTCGGCAGAGCTGGCGTAGCTCGGGTTCGACTTGCTGAAGCGAAGACCTCTGCTCGC

 D  T  T  S  P  A  A  M  S  E  T  L  R  A  L  L  L  G  D  V
GGACACCACGAGCCCGGCCGCCATGTCCGAGACGCTGCGAGCGCTGCTGCTGGGCGACGT
      1930                 1950                 1970
CCTGTGGTGCTCGGGCCGGCGGTACAGGCTCTGCGACGCTCGCGACGACGACCCGCTGCA

 L  S  P  E  A  R  G  K  L  A  E  W  M  R  H  G  G  V  T  G
GCTGTCTCCGGAGGCCCGCGGGAAGCTGGCGGAGTGGATGCGCCACGGCGGCGTGACCGG
      1990                 2010                 2030
CGACAGAGGCCTCCGGGCGCCCTTCGACCGCCTCACCTACGCGGTGCCGCCGCACTGGCC

 A  L  L  R  A  E  A  E  D  A  W  L  I  L  D  K  S  G  S  G
CGCATTGCTGCGCGCCGAGGCCGAGGACGCCTGGCTGATCCTCGACAAGTCGGGCAGCGG
      2050                 2070                 2090
GCGTAACGACGCGCGGCTCCGGCTCCTGCGGACCGACTAGGAGCTGTTCAGCCCGTCGCC

 S  H  T  R  N  L  V  A  V  I  Q  P  E  G  G  A  P  W  I  A
AAGCCACACGCGCAACCTCGTCGCGGTGATCCAGCCTGAAGGCGGAGCGCCCTGGATCGC
      2110                 2130                 2150
TTCGGTGTGCGCGTTGGAGCAGCGCCACTAGGTCGGACTTCCGCCTCGCGGGACCTAGCG

 T  M  F  I  S  D  T  D  A  E  F  E  V  R  N  E  A  L  K  D
GACCATGTTCATCTCGGATACGGACGCGGAGTTCGAGGTTCGCAACGAGGCGCTCAAAGA
      2170                 2190                 2210
CTGGTACAAGTAGAGCCTATGCCTGCGCCTCAAGCTCCAAGCGTTGCTCCGCGAGTTTCT

                                                  terminator
 L  G  R  A  V  V  A  V  V  R  E  *                ------>  <------
TCTGGGTAGGGCGGTGGTCGCGGTTGTTCGCGAATAGCCTATCCCAGGCGCGGCTTCCGC
      2230                 2250                 2270
AGACCCATCCCGCCACCAGCGCCAACAAGCGCTTATCGGATAGGGTCCGCGCCGAAGGCG
```

Fig. 5.5 (continued)

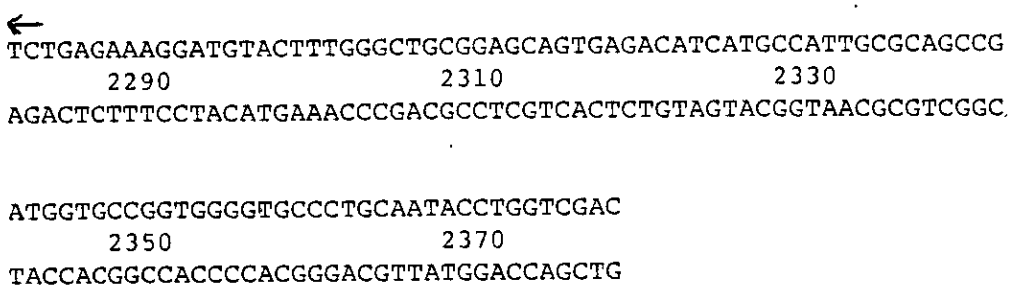


Fig. 5.5: Nucleotide sequence of the rhodopseudomonad derived DNA fragment. The boxed sequences indicate *E. coli*-like ribosome binding sites (Shine/Dalgarno sequences), while, the arrows show the mapped β-lactamase terminator sequences. Amino acid translations of large ORFs have been presented in their one-letter notation.

Fig. 5.5 (continued)

Fig. 5.4: Amp^R plaques

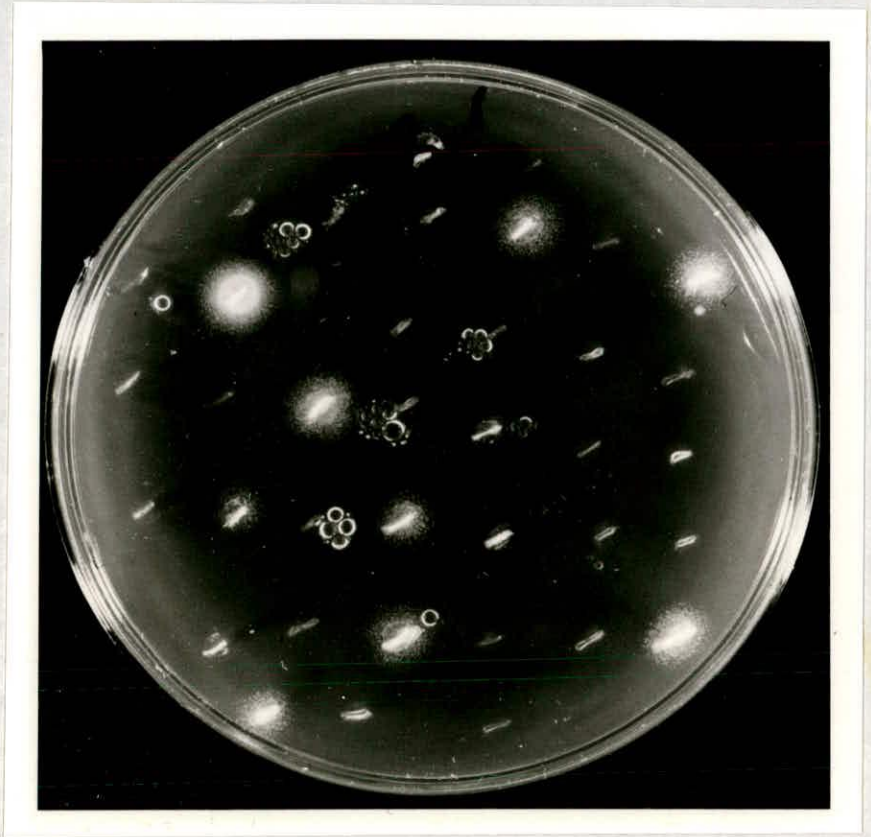


Fig. 5.4: Ampicillin resistant plaques. β-lactamase producing plaques conspicuously surrounded by a halo of bacterial growth.

DNA arising from cloning the Sal 1 fragment. These were identified as plaques surrounded by a halo of cells when the plaques were transferred onto an ampicillin sensitive *E. coli* FS 1585 lawn on L-broth ampicillin plates (Fig. 5.4). The amp^R plaques were later confirmed for β-lactamase production by staining with nitrocefin (section 3.4).

5.5. DNA sequence of the *Rps. capsulata* Bam H1/Sal 1 fragment

The DNA sequencing method and compilation of data used were the same as those outlined in section 4.2 for the pRLG 304 fragment. The DNA sequence of the *Rps. capsulata* 2.3 kb Sal 1/Bam H1 fragment is shown in Fig. 5.5. The internal Bgl II site was located and served as a confirmatory site during the DNA sequencing. There were two Xho 1 sites at positions 898 and 1105 respectively. Their proximity might account for why only one site was detected in that region during restriction enzyme mapping.

Each nucleotide was sequenced on average, four times. After sequencing 140 clones, the region from nucleotide 392 to 673 had only been sequenced on one strand, although in seven different experiments. To solve the problem, two clones (72 and 113) covering this region were made double stranded, excised with Bam H1 and Eco R1 and cloned into M13 mp19 double stranded DNA (clone turn around; section 2.2.4.7). Templates from cells transformed with the recombinant phage DNA were then sequenced to provide the complementary strand of the region.

The G + C content calculated in the sequenced region was 68.5% (T, 15.9%; C, 33.8%; G, 34.7%; A, 15.6%; ref Table 5.1). This falls within the range reported for *Rhodopseudomonas* spp (Jan de Bont *et al*, 1981) and from direct calculations of the recently sequenced cytochrome *c* gene (Daldal *et al*, 1986).

Fig. 5.8: Comparison of the rhodopseudomonad β-lactamase & Tem 1 sequences

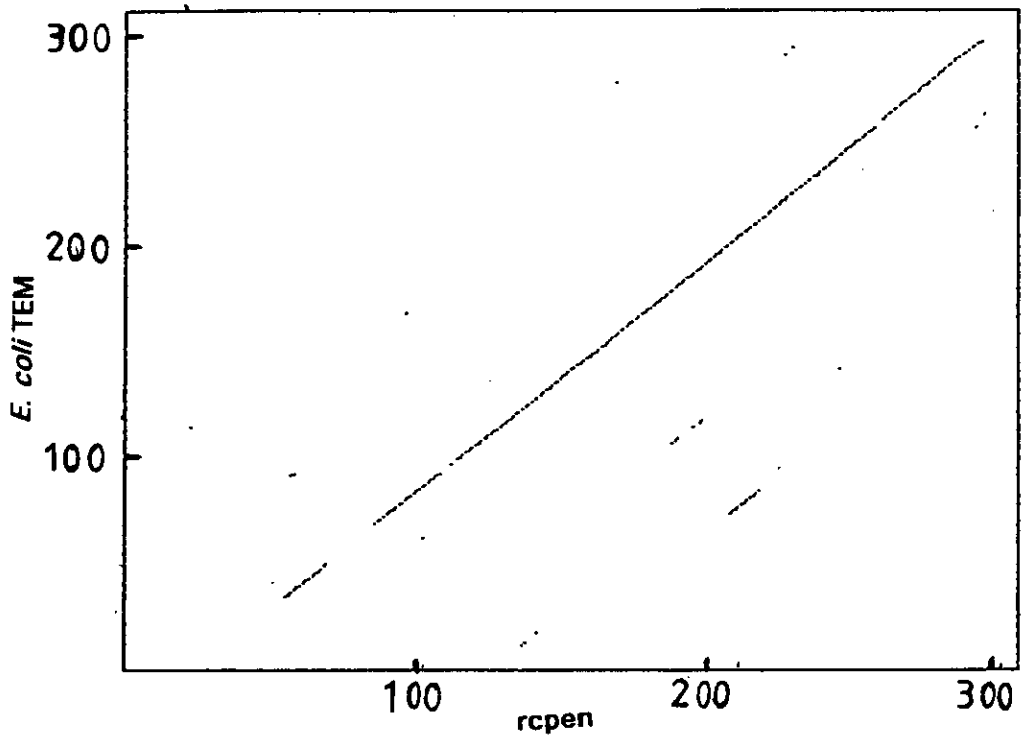


Fig. 5.8: Comparison of the protein sequences of rcpen and *E. coli* TEM β-lactamase using the program DIAGON. The window length is 21 amino acids.

Fig. 5.7: Framescan of the rhodopseudomonad DNA fragment

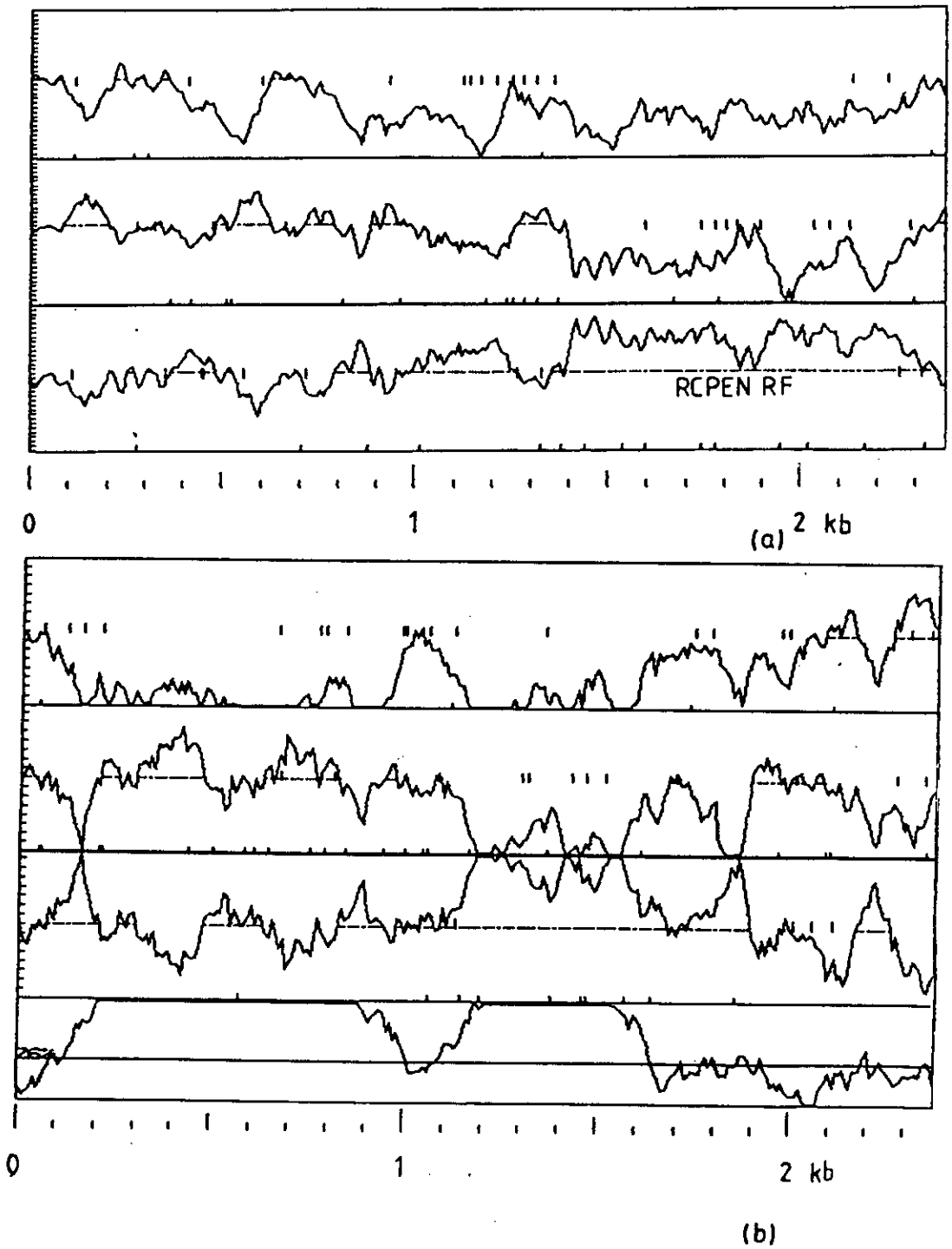


Fig. 5.7: Gene predictions for the *Rps. capsulata* DNA fragment using the codon preference option of Staden & McLachlan (1982). This is one of the functions of ANALYSEQ (explained in the text). The abscissa represents the DNA in (a), the same orientation as Fig. 5.5 and (b), in reverse orientation. The probability is plotted on the ordinate. The method assumes that the codon usage of genes from related organisms is similar and in this case the codon usage of *Rps. blastica* atp operon (Tybulewicz *et al*, 1984) was used.

The probabilities for each of the six reading frames have been plotted one above the other, every 3 codons. a continuous line at the mid-point of a reading frame (this is at the 50% level of probability) indicates which of the three frames is most likely to be coding. The initiation codons, ATG or GTG, are marked as vertical bars at the base of each plot and the termination codons as vertical bars along the 50% level.

Fig. 5.6: ORFs & stop codons in the rhodopseudomonad DNA fragment

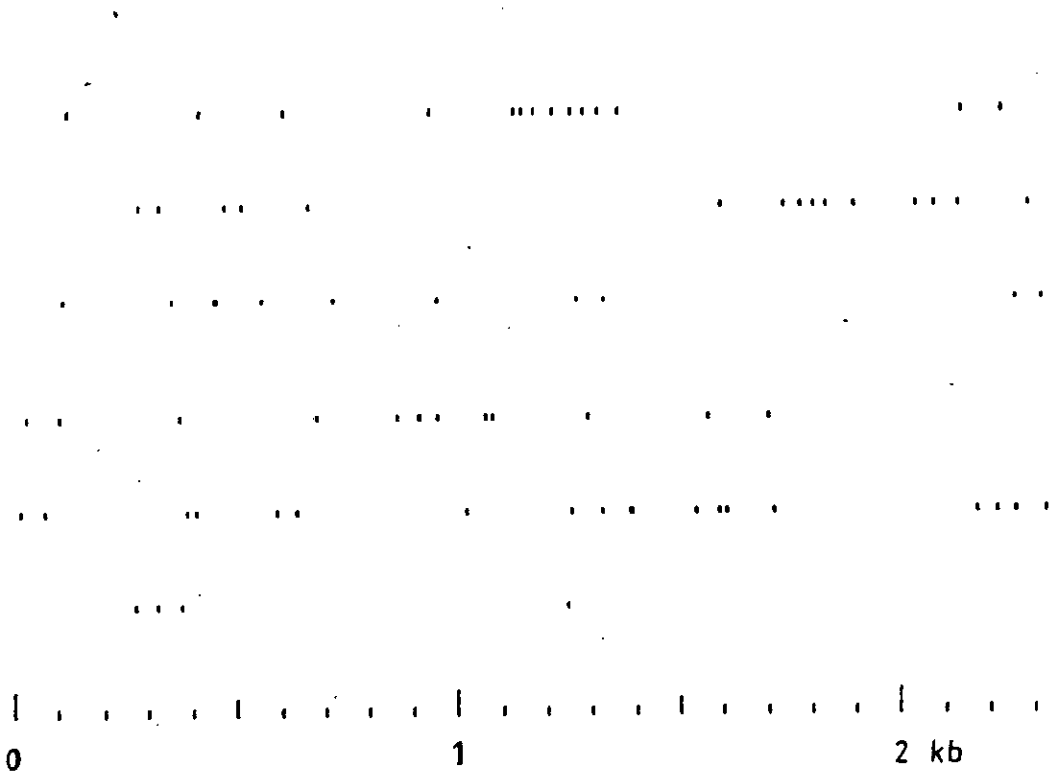


Fig. 5.6: ORFs predicted by ANALYSEQ in the sequenced rhodopseudomonad DNA. The positions of the stop codons are represented as vertical bars at the base of each of the six reading frames. The absence of stop codons in long stretches of DNA is suggestive of a potential coding region.

5.6. Analysis of the sequence data

The analysis of the *Rps. capsulata* nucleotide sequence was similar to that carried out for the *Ps. aeruginosa* sequence (see section 4.3). The open reading frames were identified using the program ANALYSEQ (Fig. 5.6).

The coding sequences could not, however, be confirmed by the use of positional base preference methods as very poor graphs were obtained. The content of the sequence library used for generating the average frequencies is strongly biased towards certain organisms (Staden, 1984a) and so it is possible that the very high G + C content (68.5%) of the *Rps. capsulata* DNA affected the analysis. The codon usage option was thus employed.

This option uses the method of Staden and McLachlan (1982) to calculate the probability that a given gene codes for a protein by having a codon usage resembling that of a "standard" gene. This method assumes that codon preferences in related organisms are similar since, according to the Wobble hypothesis (Crick, 1966), the first and second bases of codons are usually specific with the third base being more variable. As a result, there is a bias in this third base for those which are more common in the DNA.

To confirm the coding regions in the *Rps. capsulata* DNA, a codon usage table compiled from studies on the nucleotide sequence of *Rps. blastica* atp operon (Tybulewicz *et al*, 1984) was used as the standard gene. Two coding regions were identified in the forward orientation (Fig. 5.7). Another ORF was also observed in the opposite orientation. The β-lactamase gene was identified by comparing the translations of these regions to the TEM 1 β-lactamase protein sequence held in the EMBL library using the program DIAGON (Fig. 5.8). A short sequence resembling the Shine/Dalgarno sequence (Shine/Dalgarno, 1974) predicted for rhodopseudomonad genes (Gibson *et al*, 1979) was found 7 bps upstream from the start Met of the β-lactamase gene. Its corresponding protein

FIG. 5.9: LOCATION OF *RPS. CAPSULATA* SP 108 B-LACTAMASE GENE

(A) 0.75% agarose gel electrophoresis of total genomic and plasmid DNA digests of *Rhodopseudomonas* spp.

(B) Hybridization analysis of Southern transfers of DNA from gel (A) using a hybridization probe generated from template carrying DNA within the rhodopseudomonad β-lactamase gene.

Track	DNA	Restriction endonuclease
1	pSS20	Sal 1
2	pSS200	Sal 1
3	λLS20	Bam H1
4	*Rps. capsulata* sp 108 plasmids	Bam H1
5	*Rps. capsulata* sp 108 plasmids	undigested
6	*Rps. capsulata* sp 108 total DNA	Bam H1
7	λcI857	Hind III
8	*Rps. capsulata* St. Louis plasmid	Bam H1
9	*Rps. capsulata* St. Louis plasmid	undigested
10	*Rps. capsulata* St. Louis total DNA	Bam H1
11	*Rps. capsulata* sp 109 plasmids	Bam H1
12	*Rps. capsulata* sp 109 plasmids	undigested
13	*Rps. capsulata* sp 109 total DNA	Bam H1
14	*Rps. sphaeroides* plasmid DNA	Bam H1
15	*Rps. sphaeroides* plasmid DNA	undigested
16	*Rps. sphaeroides* total DNA prep.	Bam H1

Notes:

1. The hybridization of the probe in tracks 1, 2, 3, 6 & 13 suggest that the β-lactamase gene may be chromosomally mediated in the photosynthetic bacteria.

2. The size markers (S) are derived from a Hind III digest of λcI857.

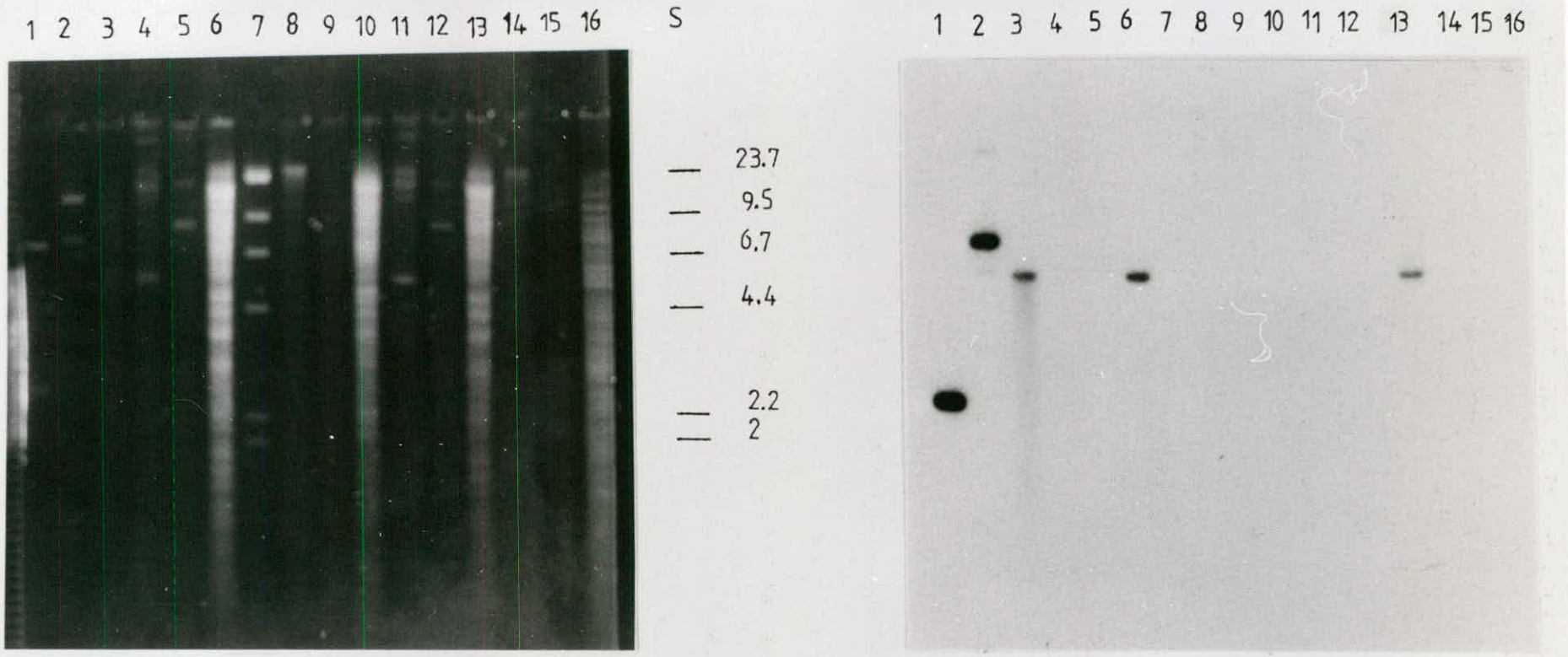

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
S
23.7
9.5
6.7
4.4
2.2
2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
A
B

Fig. 5.9

sequence is named rcpen.

5.7. Chromosomal or plasmid location of the β-lactamase gene

The original clone of the *Rps. capsulata* β-lactamase gene was derived from a total genomic digest. Scahill (1981) had also observed and studied several native plasmids in strains sp 108 and sp 109. Both had four plasmids (127, 92, 14.9 and 8.4 kb). No differences in plasmid sizes or in restriction fragment patterns generated could be observed when the strain sp 108 was grown in the presence or absence of penicillin and growth in EtBr did not abolish the penicillin resistant phenotype. Hybridization of λLS20 (original 12 kb Eco R1 clone) to plasmid preparations and the total genomic digests had suggested a plasmid location for the gene. Interestingly, there was no hybridization to total genomic DNA.

In the investigations reported in this thesis, a M13 hybridization probe was made from a template carrying an insert extending from nucleotide 1523 to 1757 (a region within the β-lactamase coding sequence), and used to probe genomic and plasmid digests of *Rps. capsulata* strains sp 108 and sp 109. *Rps. capsulata* (St. Louis) and *Rps. sphaeroides* which are both penicillin sensitive, were included as controls.

The probe hybridized to bands corresponding to the 5.8 kb Bam H1 fragments in *Rps. capsulata* sp 108 and sp 109 total genomic DNA tracks only (Fig. 5.9). This result suggests that the β-lactamase gene is probably chromosomally mediated or that, if it is on a plasmid, the plasmid is not recovered by the method of plasmid preparation used here. The latter is unlikely since the alkaline denaturation method (section 2.2.2.1.iii) used had been successful in isolating the *Ps. aeruginosa* RMS 149 plasmid which was very difficult to recover by more conventional methods.

No difference was observed between the hybridization of this probe to DNA isolated from ampicillin induced and non-induced strain sp 108 growths. This is

Fig. 5.10: Graphical representation of *Rps. capsulata* gene & surrounding sequences

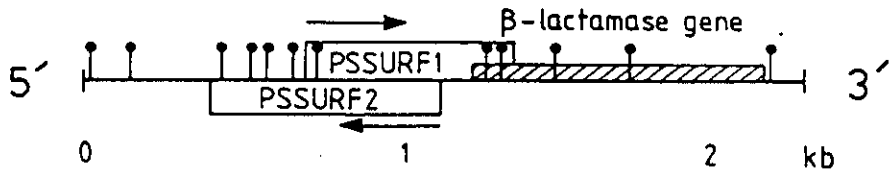
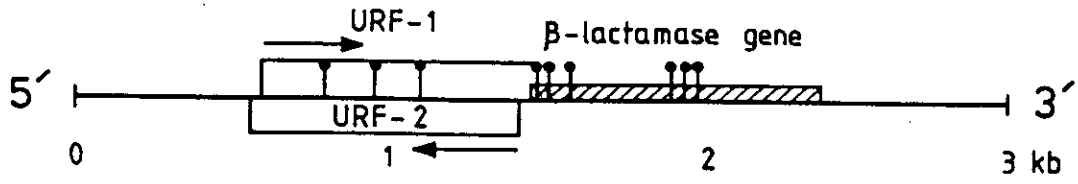


Fig. 5.10: Graphical representation of the rhodopseudomonad β-lactamase gene and surrounding DNA sequences. The arrows indicate the direction of transcription of the ORFs. The symbol ⚲ indicates possible stem-loop forming sequences.



For reference: Fig. 4.6: Graphical representation of the RMS 149 β-lactamase gene and flanking sequences.

in agreement with Scahill's (1981) earlier finding. The failure of λLS20 DNA to hybridize to Eco R1 digested sp 108 total DNA digest was postulated by Scahill to be because plasmid DNA constitutes only a fraction of total genomic DNA (Scahill, 1981). Judging by the hybridization of a similarly generated probe to the RMS 149 plasmid and to total DNA preparation (section 4.5.2.1), the lack of hybridization to plasmid DNA in this instance seems more indicative of a chromosomal location of the β-lactamase gene. The clearly distinguishable plasmid types in the *Rps. capsulata* undigested plasmid track (Fig. 5.9a) confirm that the plasmid preparation was satisfactory. Thus, in contrast to Scahill's (1981) finding, the *Rps. capsulata* sp 108 β-lactamase gene may be chromosomal.

5.8. Other ORFs

PSSURF-1

This ORF starts upstream of the β-lactamase gene from nucleotide 1592. PSSURF-1 has a putative Shine/Dalgarno sequence AACGGA (Fig. 5.5) that is not as good as the β-lactamase perfect consensus one. It codes for a protein of molecular weight, 28,169.

PSSURF-2

In the complementary strand, PSSURF-2 starts from position 1239 and ends at nucleotide 368. This second ORF is not in the same phase as PSSURF-1 and as such it might be a protein coding region itself (see section 4.3.4.2). It codes for a protein of molecular weight, 31,453.

An extensive search, using the UWGCG program WORDSEARCH to compare these putative proteins against entries in the NBRF protein library, failed to show up any significant homology. Interestingly, the gene organization around the β-lactamase gene in *Rps. capsulata* sp 108 (Fig 5.10) is very similar to that observed with the *Ps. aeruginosa* RMS 149 plasmid encoded enzyme (ref. section

TABLE 5.2: PUTATIVE SECONDARY STRUCTURES IN THE RHODOPSEUDOMONAD DNA

Stem sequences	loop size	Location	$\triangle G$ (kcal mol^{-1})
20 CGCCCAUCG 48 GCGGGCAGC	11	within external URF	−14.6
283 CGGGCAUCCA 304 GUCCGUGGGU	2	terminator of external URF?	−17.4
431 GUCGGACG 463 CAGCCUGC	16	at terminus of PSSURF2	−11.2
578 CACCGAGCUG 605 GUGGCCCGGC	8	within PSSURF2	−10.6
607 CAGGCGGGC 627 GUCCGCACG	3	before PSSURF1	−12.6
735 AGCGGCGCC 755 UCUCCGCGG	3	before PSSURF1	−15.4
828 AGGAUCUC 862 UCCUAGAG	19	before PSSURF1	−8.2
1394 GUCCUGUCGCG 1418 CGGGACAGCGC	3	beginning of β-lactamase gene	−15.2
1538 GAGACCGG 1565 CUCUGGUC	12	within β-lactamase gene	−4.8
1719 CGCGGGUCG 1745 GCUCCCAGU	9	within β-lactamase gene	−7.6
2017 GAUGCGCC 2046 UUACGCGG	13	within β-lactamase gene	−12.6
2302 GGCUGCGGAG 2339 CCGACGCGUU	18	after end of β-lactamase gene	−11.6

Notes:

- The possible secondary structures of RNA from the rhodopseudomonad β-lactamase gene and surrounding DNA were predicted using the program STEMLOOP.
- The free energies of pairing were calculated using the rules of Tinoco *et al.*, (1973). 1 kcal = 4.184 kJ.

4.6). In both cases, an upstream gene terminates within the coding region of the β-lactamase gene and, in the complementary strand, another gene starts off just before the putative start of the β-lactamase gene.

The two PSSURFs were compared with the pseudomonad URF-1 and URF-2 using the programs BESTFIT and GAP. No significant homologies were detected and it is thus possible that the similarities in the gene arrangement are coincidental.

5.9. Possible secondary structures

The program STEMLOOP (UWGCG) indicated twelve potential stem-loop forming sequences in the 2.3 kb rhodopseudomonad DNA fragment, with eight or more base pairs in the stem (Table 5.2). Two of these occur in the 5' flanking sequences with one well placed to act as a terminator sequence to an ORF which starts outside the 2.3 kb fragment (Fig. 5.10). Four potential stem-loop structures are clustered in the region between nucleotides 283 and 756 upstream of the PSSURF-1 start. Apart from two other loop-forming sequences which are within the β-lactamase gene so that they occur in the putative carboxyl terminus of PSSURF-1 protein, there is only one such structure within PSSURF-1 itself and this is situated 17 bps after its start codon.

Four possible stem-loops are interspersed within the β-lactamase gene, the first one being 18 bps from the putative initiation codon. The β-lactamase gene ends in a poorer stem-loop structure with four G-C bonds but leads to an A-T rich region.

While the latter may function in rho-independent termination of transcription (Platt, 1981), the other loops may act to regulate gene expression during translation. As pointed out in section 4.6, some of them may have no special roles.

5.10. Repeat sequences

No repeat lengths of any major significance were observed in the DNA sequence using the program REPEAT (UWGCG). An interesting observation, however, was the dyad symmetrical sequence occurring before the PSSURF-1 Shine/Dalgarno sequence at nucleotide 738 to 752 which has a similar sequence to another between nucleotides 1340 and 1353, just before the β-lactamase Shine/Dalgarno sequence as shown below:-

738 GGCGCCATCGGCGCC 752

1340 GGCGCCACTGGCGC 1353

Their similarity in sequence and in positioning, i.e. before a putative ribosomal binding site, would seem indicative of a regulatory role. Regulatory proteins e.g. cro, cI and lac repressor bind to DNA with dyad symmetry (Wharton *et al.*, 1984; Anderson *et al.*, 1984; e.t.c.).

There was, also, an 11 bp direct repeat at either end of the β-lactamase gene at nucleotide 1334 to 1344 and 2261 to 2271. As three other 11 bp direct repeats were also observed to occur within the β-lactamase gene and other parts of the 2.3 kb fragment, it was not clear if the flanking direct repeats had any significance or if they were simply coincidental.

Genetically unstable ampC mutants which carry tandem copies of the ampC gene have been reported (Edlund *et al.*, 1979). The ampC containing repeats were 10 - 20 kb in size, and in one case, it was demonstrated that the initial duplication had occurred through an unequal recombination event between two 12 bp repeats, 9.8 kb apart on either side of the ampC gene (Edlund & Normark, 1981). It is thus possible that these *Rps. capsulata* flanking 11 bps repeats may have been responsible for the highly unstable penicillin resistant isolates earlier described by Wall *et al.* (1975).

FIG. 5.11: S1 NUCLEASE MAPPING OF *RPS. CAPSULATA* B-LACTAMASE MRNA

1. Tracks TCGA represent dideoxynucleotide sequences of a marker M13 template.

2. Probe 2 refers to the probe spanning the β-lactamase gene Shine/Dalgarno sequence while probe 45 spans the putative termination loop.

3. Tracks a, probe hybridized to total RNA preparation and S1 digested; tracks b, no hybridization and S1 digestion; tracks c, untreated probe.

Note:

Probe 2 was completely protected by RNA during S1 mapping though, probe 45 was reduced by 48 bps (as determined by counting the number of sequencing bands between original probe and the protected fragment).

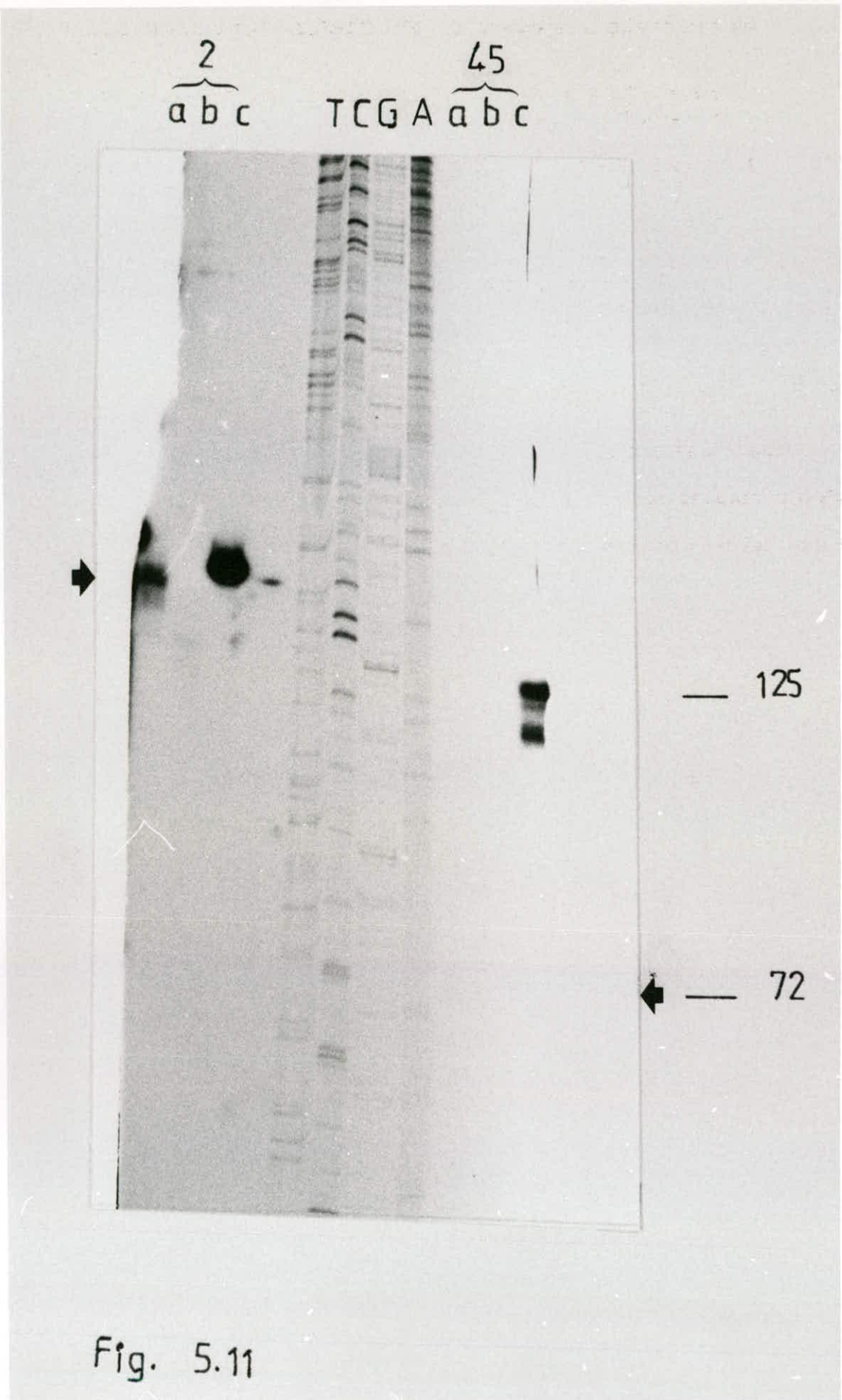


Fig. 5.11

5.11. Transcription of *Rps. capsulata* sp 108 β-lactamase gene

RNA mapping studies aimed at determining transcriptional start and stop sites in the rhodopseudomonad β-lactamase gene were carried out as described for the *Ps. aeruginosa* plasmid encoded enzyme (section 2.2.5.3). Two probes were used:-

Probe 2:	starts at the Dde 1 site at nucleotide 1246 upstream from the β-lactamase Shine/Dalgarno sequence and extends to position 1371, three bases downstream of the putative Shine/Dalgarno sequences.
Probe 45:	extends from the Bgl II site in the putative carboxyl-terminus of the β-lactamase (at nucleotide 2218), past the stop codon to nucleotide 2336.

After S1 digestion of probe/RNA hybrids and examination of the protected fragments on denaturing polyacrylamide gels, the 125 bases of probe 2 were found to be completely protected while probe 45 was reduced by about 48 bases (ref. Fig. 5.11).

These results suggest that the β-lactamase gene is transcribed from an upstream promoter, probably that of PSSURF-1. Work is in progress to map the putative promoter of PSSURF-1 by using a probe spanning DNA sequences upstream and downstream, of its Shine/Dalgarno sequence. The reduction in size of the termination-site probe by 48 bases can be accounted for by S1 digestion of the 17mer hybridization primer as well as the 18-20 bases beyond the termination point of the β-lactamase gene. Thus the results suggest that termination of the β-lactamase transcripts occur around its own terminator sequence mapped in Fig. 5.5.

Generally, in the S1 mapping experiments carried out in these investigations, the probes that were bound to longer transcripts tended to be fully protected (i.e. 17mer primer and all), while the end probes readily lost the primer. It is also possible that in the latter cases, some bases are lost either due to "breathing" at

the ends of the RNA/DNA hybrids, which allows S1 nuclease to digest into the double-stranded hybrid or due to exonucleolytic digestion of the RNA molecules during the extraction procedure (Tybulewicz *et al.*, 1984). The RNA preparations were made using the rapid hot phenol extraction procedure (section 2.2.2.4), in order to reduce to a minimum any exonucleolytic digestion. The breathing at the ends of the hybrid molecules cannot be prevented. More precise transcriptional start and stop sites can however, be determined by carrying out primer extension experiments. Here, prime-cut probes are hybridized to RNA and extended with reverse transcriptase. The probes can then be examined on sequencing-type gels for any increase in length.

5.12. Summary

The *Rps. capsulata* sp 108 β-lactamase gene was located to a 2.3 kb Sal 1/Bam H1 DNA fragment of the 5.8 kb Bam H1 insert in pSS20. Its nucleic acid sequence was determined and the amino acid sequence of a β-lactamase, clearly recognizable as a class A enzyme (Ambler classification), was identified.

This class of β-lactamase is commonly found in Gram-negative bacterial plasmids (Matthew & Hedges, 1976; Richmond *et al.*, 1980), but hybridization studies (section 5.7) suggest a possible chromosomal location for the *Rps. capsulata* sp 108 β-lactamase gene. β-lactamase genes are often organized as part of a transposable element (Richmond *et al.*, 1980), which can move between bacterial replicons independently of the cell's generalized recombination systems. Thus, they can exist as part of a chromosome, plasmid or phage DNA. In TnA, the β-lactamase gene ends 110 bps upstream from the start of the inverted repeats required for transposition (Heffron *et al.*, 1979). The DNA sequence here only extends 119 bps downstream of the β-lactamase stop codon. This means that there was not enough sequence information to check for TnA inverted repeat sequences.

Its 68.5% G + C content, which is more characteristic of rhodopseudomonad species (compare *E. coli*, 48% - 52% G + C), seems to indicate that this is an authentic rhodopseudomonad gene. Thus, the *Rps. capsulata* sp 108 β-lactamase might represent the first characterised example of a Gram-negative chromosomal, inducible class A β-lactamase. This raises the possibility that such a homologous gene in Enterobacteriaceae might be the source of the TEM gene commonly found on transposons (see section 7.8 below).

CHAPTER 6

CLASSIFICATION AND EVOLUTION OF CLASS A β-LACTAMASES

β-lactamases have been used as tools for a variety of molecular biological studies (section 1.2). Particular interest in them as detoxifying agents against the chemotherapeutically important β-lactam antibiotics means that a wealth of data has been accumulated about the various types found in diverse bacteria. Reference has already been made to the usefulness of a classification of enzymes, based on primary sequence relationships (section 1.2.3). It is hoped that these studies will provide information about the origin and genetic relationships of the β-lactamases, and at the same time, extend our knowledge of bacterial protein evolutionary processes.

Reviews concerning the use of macromolecular sequences for taxonomic and phylogenetic analyses are numerous (Sneath, 1974; Peacock, 1981; Falsenstein, 1981; Ambler, 1984). Thus, this chapter will only give a brief outline of the methods used and how they have been applied to the *Ps. aeruginosa* RMS 149 plasmid and *Rps. capsulata* sp 108 β-lactamases.

6.1. Molecular evolution of enzymes

6.1.1. Introduction

The study of the evolution of new enzymes (metabolic activities) is only a part of a larger field which deals with the evolution of the organism.

Geological data show that conditions on earth are not static and that, indeed, at the beginning, they were different from what they are now. The planet is believed to have emerged by the condensation of a cloud of gases and dust. Thus, the then abundant simple chemicals like N_2, NH_3, CH_4 and H_2O were transformed into the more complex macromolecules now characteristic of life, in

reactions driven by the energy of the sun (Oparin, 1965). Miller (1973) have demonstrated the soundness of these postulates in experiments in which they generated large quantities of amino acids and other organic molecules by subjecting a mixture of H_2, NH_3, CH_4 and H_2O to an electric discharge in vacuo.

With the evidence from the basic biochemical unity of all present day organisms, the most economical hypothesis is that all life forms have arisen from one source. Living cells probably arose on earth by the spontaneous aggregation of molecules about 3.5 thousand million years ago. From our knowledge of present day organisms, at least three steps must have occurred before the first cell emerged:

1. formation of RNA polymers capable of self replication;
2. development of protein synthesis directed by RNA molecules;
3. compartmentalization of the above by a lipid membrane (Alberts *et al.*, 1983).

Thus, the first life forms - most probably anaerobic heterotrophs - came into existence in an aqueous oceanic environment in which were dissolved large masses of prebiotic compounds (the "primeval soup"). Organisms are postulated to have evolved by developing more enzymes to produce types more suited to the new ecological niches.

Natural selection involves the interaction of the phenotype with the environment. Ultimately, the nature of the phenotype, depends directly or indirectly, on genetically determined structures and hence, the function of enzymes and other proteins. The flow of information within the cells was postulated by Crick (1958): DNA is transcribed into RNA which is then translated into proteins. All the information for the fate of the polypeptide is determined by the nature of its constituent amino acids. Therefore, the folding of individual proteins and association with other macromolecules occurs without further direct

intervention of the genetic material.

Translation of genetic material into a polypeptide chain involves triplet codons for 20 possible amino acids. All genomic modifications, therefore, affect proteins to a greater or lesser extent, depending on the type and position of the amino acid changed. It is, hence, possible to visualize the situation whereby organisms evolve via such changes. The observations that proteins which serve the same functions in all organisms frequently show strong similarity in structure (e.g. heme-bearing proteins - globins have striking similarities amongst eukaryotic species (Dayhoff, 1972), as well as the disappearance in higher organisms of some enzymes present in lower ones and the appearance of many new enzymes in higher organisms, attest to the general validity of these postulates (Smith, 1970).

6.1.2. Evolutionary studies

Two main approaches are currently used in studying the processes involved in the evolution of enzyme systems (Hegeman and Rosenberg, 1974). They involve:-

1. Selection of a group of different organisms and studying relationships of a character (or group of characters), in an attempt to define the pressures which act upon them (descriptive approach).

2. Selection of an organism and studying the changes of a character (or group of characters) produced by controlled applications of selection pressures (experimental approach).

Both methods have been applied to a variety of systems and have revealed:

- the genetic factors involved in the changing structure of homologous enzymes (e.g. sequence similarities of different serine proteases suggest that the enzymes evolved from a common precursor by gene duplication and subsequent independent evolution; Markland and Smith, 1971);

- the specific roles of amino acid residues in the mechanism of action and specificity as opposed to protein conformation and

- the selective factors operative in the evolution of enzymes, leading to the development of new metabolic pathways.

Micro-organisms have been used to investigate the last finding. Their high populations and rapid growth rate make bacteria ideal systems for studying evolution in the laboratory. Such experiments consist of placing the organism under conditions where they must develop new capabilities to grow. Thus, pathways for the acquisition of catabolic enzymes for novel substrates have been studied for *Klebsiella aerogenes* mutants that can metabolize 5-carbon sugars (Mortlock, 1982) and for the ebg operon of *E. coli* k12 - a new operon for lactose utilization (Campbell *et al*,1973; Hall & Hartl, 1974), etc.

These experiments in microbial evolution suggest that new metabolic capacities emerge by a plasmid acquisition, which introduces a new gene conferring on the host some selective advantage, or by the recruitment of other normal cell enzymes so that they can be utilized for a new function.

Plasmids are extrachromosomal DNA elements which are capable of replicating autonomously within the cell. The elements generally include no genes that are unconditionally needed for reproduction of the host. Characters encoded by plasmids are conventionally distinguished from chromosomal ones by the ability to readily transfer them to other bacteria or to lose them by curing. There are several chemical and physical methods of curing bacteria of their plasmids. Chemical agents include EtBr, SDS, rifampicin, and novobiocin. The two main physical methods used are elevated temperature and UV irradiation. Most of the chemical agents are capable of intercalating between DNA bases, so that the net result of treatments with curing agents is to disturb general DNA metabolism. In the process, the dispensible extrachromosomal replicons may be lost. Plasmids can however integrate into the chromosome, so that it can sometimes be difficult to tell what character is plasmid-borne.

The mutations that lead to the recruitment of normal cellular enzymes for a new role are either regulatory ones (a previously regulated enzyme becomes constitutively expressed) or structural ones. In either case, new enzymes are produced which interact with novel substrates better (Campbell *et al*, 1973; Hall & Hartl, 1974; Mortlock, 1982).

In summary, by combining the findings of laboratory simulated enzyme evolution and observations from descriptive studies, there is now more understanding of the mechanism of enzyme evolution.

6.1.3. Methods used in determining macromolecular relationships

Studies of the history of evolution have been based on analysis of the fossil record. If metabolic pathways evolved by the sequential addition of new enzymatic reactions to existing ones, then genetic sequence is also a historical record (Bryson & Vogel, 1965; Zukerkandl & Pauling, 1965). Therefore, a comparative analysis of genetic sequence can be used to establish genealogical relationships amongst organisms. The many approaches now available are summarized in Table 6.1.

Table 6.1: Parameters used in measuring molecular evolution

Macromolecule	method
Proteins	Gel electrophoresis
Proteins	Microcomplement fixation
Proteins	Amino acid sequence comparison
DNA/RNA	Nucleotide sequence comparison
DNA/RNA	Restriction mapping
DNA/RNA	Hybridization experiments

All the different methods listed above reflect the organism's genetic information directly or indirectly. Gel electrophoresis (separating proteins on the basis of their overall charge) is a measure of differences in amino acid composition. It has been used extensively to distinguish different but related proteins e.g. Matthew (1979) showed that Gram-negative bacterial β-lactamases could be grouped

according to their characteristic isoelectric points. Complement fixation (which measures interaction of test proteins with a reference antibody) and hybridization experiments have limited use as only very close relationships can be recognized. These methods are, however, indispensible for initial screening of large samples. The use of direct sequence comparisons is dealt with below.

6.1.4. Aligning macromolecular sequences

The best data for studying the relationship between two organisms or group of organisms is an aligned protein or nucleic acid sequence (Dayhoff, 1969; Woese, Stackebrandt, 1981; Ambler,1984). The unique advantage is due to the fact that sequence information is "digital" rather than "analogue" (Ambler, 1976; 1984). Thus, an aligned sequence data provides a better quantitative handle for these studies in contrast to the imprecise data obtained from comparing the gross physical properties.

Once aligned, the relationship can be considered as being homologous (i.e. both sequences come from the same ancestral gene) if their primary sequences show sufficient similarities in the number of invariant residues (i.e. residues occurring in every sequence at certain positions in the alignment) and conservative substitutions (a change in the residue at a particular alignment position such that the chemical properties at that point are maintained). They may be analogous (i.e. come from different ancestral genes) if no significant sequence similarities can be detected. Satisfactory significance in this context is difficult to formulate, and different workers are likely to disagree on any particular case. It follows, too, that proteins with similar functions need not be homologous (a situation defined as convergent evolution).

6.1.5. Quantifying sequence similarities and differences

Ordinarily, the number of invariant residues can be counted and similarity expressed as a percentage of total number of residues in one sequence. For

more complex analysis, e.g. building phylogenetic trees, each pair is compared position by position and the number of differences is converted to percentage difference between the sequences.

The number of differences between aligned sequences can be given equal value (Dayhoff & Eck, 1968) or, more frequently, may be based on the minimum nucleotide distances (i.e. the nature of the genetic code and rates of mutation at nucleotide level are taken into consideration - e.g., a valine to leucine mutation with a GTC to CTC codon change would have a value of 1 while a methionine to tryptophan change (ATG to TGG) would have a value of 2; Fitch, 1966). Unfortunately, it is not clear what value to assign an insertion or deletion event. It is conventional to score a value of 1 for every base involved but this is arbitrary as the size of a deletion is probably not directly related to its frequency of occurrence.

The values derived from comparing the homologous sequences from different organisms are then used to construct a matrix which gives a rough estimate of the evolutionary distance between any pair or group of organisms. Evolutionary trees which show more details in the nature of ancestral and present day organisms can be constructed by using the values from the matrix table. Phylogenetic trees are also drawn using the most probable ancestral sequences deduced from studying amino acid changes in aligned sequences.

6.1.6. Problems with deducing phylogeny from single genes

Several factors impair the use of sequence data for classification and evolutionary studies. Firstly, the same mutational event might occur in different lines to produce the same character at the same position (parallelism), so that the apparent degree of similarity is increased. The extent of parallelism in any data set is difficult to measure, but estimates indicate it may be as common as divergent events (Peacock, 1981).

FIG. 6.1: AMINO ACID SEQUENCE ALIGNMENT OF CLASS A B-LACTAMASES

```
                1                                                    50
pnsap.seq       .......... .......... .......... ......MKKL IFLIVIALVL
bclaci.seq      ........MI LKNKRMLKIG ICVGILGLSI TSLEAFTGES LQVEAKEKTG
blpenc.seq      .......... ...MKLWFST LKLKKAAAVL LFSCVALAGC ANNQTNASQP
pnecp.seq       .......... .......... .......... ..MSIQHFRV ALIPFFAAFC
ppapen.seq      .......... .......... .....MRKFQ CHFLSVPVAI LGCVGLICTS
rcpen.seq       .......... .......... .....MRFTA TVLSRVATGL ALGLSMATAS
Consensus       ........-- ---------- ---------- ---------- ----------
               51                                                   100
pnsap.seq       SACNSNSSHA KELNDLEKKY NAHIGVYALD TKSGKE.VKF NSDKRFAYAS
bclaci.seq      QVKHKNQATH KEFSQLEKKF DARLGVYAID TGTNQT.ISY RPNERFAFAS
blpenc.seq      AEKNEKTEMK DDFAKLEEQF DAKLGIFALD TGTNRT.VAY RPDERFAFAS
pnecp.seq       LPVFAHPETL VKVKDAEDQL GARVGYIELD LNSGKILESF RPEERFPMMS
ppapen.seq      AYAMDTGILD LAVTQEETTL QARVGVAVID TDSGLTWQ.H RGDERFPLNS
rcpen.seq       LAETPVEALS ETVARIEEQL GARVGLSLME TGTGWSWS.H REDELFLMNS
Consensus       ---------- ------e--- -a--g----- ---------- -----f---s
              101                                                  150
pnsap.seq       TSKAINSAIL LEQVPYNKLN KKVHI..NKD DIVAYSPILE KYVGKD.ITL
bclaci.seq      TYKALAAGVL LQQNSIDSLN EV..ITYTKE DLVDYSPVTE KHVDTG.MKL
blpenc.seq      TIKALTVGVL LQQKSIEDLN QR..ITYTRD DLVNYNPITE KHVDTG.MTL
pnecp.seq       TFKVLLCGAV LSRVDAGQEQ LGRRIHYSQN DLVEYSPVTE KHLTDG.MTV
ppapen.seq      THKAFSCAAV LAQADRHKLN LEQAIPIERT ALVTYSPVTE RVPPGGTLTL
rcpen.seq       TVKVPVCGAI LARWDAGRLS LSDALPVRKA DLVPYAPVTE TRV.GGNMTL
Consensus       t-k------- l--------- ---------- --v-y-p--e ----------
              151                                                  200
pnsap.seq       KALIEASMTY SDNTANNKII KEIGGIKKVK QRLKELGDKV TNPVRYEIEL
bclaci.seq      GEIAEAAVRS SDNTAGNILF NKIGGPKGYE KALRHMGDRI TMSNRFETEL
blpenc.seq      KELADASLRY SDNAAQNLIL KQIGGPESLK KELRKIGDEV TNPERFEPEL
pnecp.seq       RELCSAAITM SDNTAANLLL TTIGGPKELT AFLHNMGDHV TRLDRWEPEL
ppapen.seq      RELCRAAVSI SDNTAANLAL DAIGGARTFT AFMRSIGDDK TRLDRREPEL
rcpen.seq       DELCLAAIDM SDNVAANILI GHLGGPEAVT QFFRSVGDPT SRLDRIEPKL
Consensus       -----a---- sdn-a-n--- ---gg----- ------gd-- ----r-e--l
              201                                                  250
pnsap.seq       NYYSPKSKKD TSTPAAFGKT LNKLIANGKL SKENKKFLLD LMLNNKSGDT
bclaci.seq      NEAIPGDIRD TSTAKAIATN LKAFTVGNAL PAEKRKILTE WMLGNATGDK
blpenc.seq      NEVNPGETQD TSTARALVTS LRAFALEDKL PSEKRELLID WMKRNTTGDA
pnecp.seq       NEAIPNDERD TTMPAAMATT LRKLLTGELL TLASRQQLID WMEADKVAGP
ppapen.seq      NEATPGDARD TTTPIAAARS LQTLLLDGVL SAPARNELTQ WMLGDQVADA
rcpen.seq       NDFASGDERD TTSPAAMSET LRALLLGDVL SPEARGKLAE WMRHGGVTGA
Consensus       n--------d t----a---- l--------l -------l-- -m--------
              251                                                  300
pnsap.seq       LIKDGVPKDY KVADKSGQAI TYASRNDVAF VYPKGQSEPI VLVIF.TNKD
bclaci.seq      LIRAGIPTDW VVGDKSGAG. SYGTRNDIAV VWP.PNSAPI I..VLISSKD
blpenc.seq      LIRAGVPDGW EVADKTGAA. SYGTRNDIAI IWP.PKGDPV VLAVL.SSRD
pnecp.seq       LLRSALPAGW FIADKSGAG. ERGSRGIIAA LGP.DGKPSR IVVIY.TTGS
ppapen.seq      LLRAGLPRDW QIADKSGAG. GHGSRSIIAV VWP.PKRSAV IVAIYITQTA
rcpen.seq       LLRAEAEDAW LILDKSGSGS HTRNLVAVIQ PEGGAPWIAT MFISDTDAEF
Consensus       l--------- ---dk-g--- ---------- ---------- ----------
```

```
          301                                327
pnsap.seq     NKSDKPNDKL ISETAKSVMK EF.....
bclaci.seq    EKEAIYNDQL IAEATKVIVK GS.....
blpenc.seq    KKDAKYDDKL IAEATKVVMK ALNMNGK
pnecp.seq     QATMDERNRQ IAEIGASLIK HW.....
ppapen.seq    ASMSASNQAV SRIGSALAKA LQ.....
rcpen.seq     EVRNEALKDL GRAVVAVVRE .......
Consensus     ---------- ---------- -------
```

Notes:

1. pnsap = *Staph aureus* PC1; Ambler, 1975.
2. bclaci = *B. cereus* 569/H; Thatcher, 1975.
3. blpenc = *B. licheniformis* 749/C; Ambler & Meadway, 1969.
4. pncep = *E. coli* TEM 1; Sutcliffe , 1978, Ambler & Scott, 1978.
5. ppapen = *Ps. aeruginosa* RMS 149 specified enzyme (this thesis).
6. rcpen = *Rps. capsulata* sp 108 β-lactamase (this thesis).
7. The residues are aligned on the system of the Ambler classification.
8. The residues that are identical in all the sequences are represented in the consensus string.

Secondly, the concept of a molecular clock, on which evolutionary deduction rests, is not unequivocal as mutation rates differ in different proteins e.g. in histone IV, a highly conserved eukaryotic DNA binding protein, the rate of evolution is very low (about 600 million years for a 1% change in sequence). On the other hand, RNAse has evolved at different rates in different lines, so that, for example, rat ribonuclease has evolved much faster than RNAses from other mammalian orders (Berg & Bientema, 1975).

In addition, if a gene occurs in an organism because of a process of lateral gene transfer, sequence comparisons may produce anomalous results. Lastly, the correctness in sequence cannot be overemphasised, as incorrectly determined sequences can, and have, led to spurious trees (e.g. the classical rattlesnake cytochrome *c* case; Ambler, 1984)

Confusions due to parallelism can be avoided by considering more than one type of macromolecule. The problem with a representative molecular clock is more serious with prokaryotes as some eukaryotic fossil records exist and have been used to provide a reference framework. Nevertheless, it is hoped that by studying a large number of widely different proteins and nucleic acids in bacteria, it will be possible to find those which constitute the genetic core (i.e. they are not likely to be easily transferred) and are more suitable for producing a natural classification (Ambler, 1979a).

6.2. Results of analyses

6.2.1. Class A β-lactamase alignment

As both the β-lactamase genes from *Ps. aeruginosa* and *Rps. capsulata* were identified using the *B. licheniformis* 749/C β-lactamase protein sequence, it was evident that the two enzymes belong to the class A β-lactamase group (Ambler classification, 1980). The deduced protein sequences are aligned in Fig. 6.1 with other class A enzymes using the Ambler scheme (1980).

FIG. 6.1A: SEQUENCE CONSERVATION AMONGST GRAM-NEGATIVE CLASS A ENZYMES

```
            1                                                        50
papen.seq    .......... .......... .....MRKFQ CHFLSVPVAI LGCVGLICTS
rcpen.seq    .......... .......... .....MRFTA TVLSRVATGL ALGLSMATAS
pnecp.seq    .......... .......... .......... ..MSIQHFRV ALIPFFAAFC
Consensus    .......... .......... .....----- ---------- ----------
           51                                                       100
papen.seq    AYAMDTGILD LAVTQEETTL QARVGVAVID TDSGLTWQ.H RGDERFPLNS
rcpen.seq    LAETPVEALS ETVARIEEQL GARVGLSLME TGTGWSWS.H REDELFLMNS
pnecp.seq    LPVFAHPETL VKVKDAEDQL GARVGYIELD LNSGKILESF RPEERFPMMS
Consensus    ---------- --v---e--l -arvg----- ---g------ r--e-f---s
           101                                                      150
papen.seq    THKAFSCAAV LAQADRHKLN LEQAIPIERT ALVTYSPVTE RVPPGGTLTL
rcpen.seq    TVKVPVCGAI LARWDAGRLS LSDALPVRKA DLVPYAPVTE TRV.GGNMTL
pnecp.seq    TFKVLLCGAV LSRVDAGQEQ LGRRIHYSQN DLVEYSPVTE KHLTDG.MTV
Consensus    t-k---c-a- l---d----- l--------- -lv-y-pvte -----g--t-
           151                                                      200
papen.seq    RELCRAAVSI SDNTAANLAL DAIGGARTFT AFMRSIGDDK TRLDRREPEL
rcpen.seq    DELCLAAIDM SDNVAANILI GHLGGPEAVT QFFRSVGDPT SRLDRIEPKL
pnecp.seq    RELCSAAITM SDNTAANLLL TTIGGPKELT AFLHNMGDHV TRLDRWEPEL
Consensus    -elc-aa--- sdn-aan--- ---gg----t -f----gd-- -rldr-ep-l
           201                                                      250
papen.seq    NEATPGDARD TTTPIAAARS LQTLLLDGVL SAPARNELTQ WMLGDQVADA
rcpen.seq    NDFASGDERD TTSPAAMSET LRALLLGDVL SPEARGKLAE WMRHGGVTGA
pnecp.seq    NEAIPNDERD TTMPAAMATT LRKLLTGELL TLASRQQLID WMEADKVAGP
Consensus    n-----d-rd tt-p-a---- l--ll----l ----r--l-- wm----v---
           251                                                      300
papen.seq    LLRAGLPRDW QIADKSGAG. GHGSRSIIAV VWP.PKRSAV IVAIYITQTA
rcpen.seq    LLRAEAEDAW LILDKSGSGS HTRNLVAVIQ PEGGAPWIAT MFISDTDAEF
pnecp.seq    LLRSALPAGW FIADKSGAG. ERGSRGIIAA LGP.DGKPSR IVVIY.TTGS
Consensus    llr------w -i-dksg-g- ---------- ---------- ----------
           301                   322
papen.seq    ASMSASNQAV SRIGSALAKA LQ
rcpen.seq    EVRNEALKDL GRAVVAVVRE ..
pnecp.seq    QATMDERNRQ IAEIGASLIK HW
Consensus    ---------- -----a---- --
```

Notes:

1. papen = *Ps. aeruginosa* RMS 149 encoded β-lactamase (this thesis).

2. rcpen = *Rps. capsulata* sp 108 β-lactamase (this thesis).

3. pncep = *E. coli* TEM enzyme (Sutcliffe, 1978; Ambler & Scott, 1978).

4. While there are some sequence conservation amongst the Gram-negative bacterial enzymes, there were no sequence homologies in the carboxyl-termini.

FIG. 6.1B: SEQUENCE CONSERVATION AMONGST GRAM-POSITIVE CLASS A B-LACTAMASES

```
                1                                                    50
pnsap.seq       .......... .......... .......... ........MK KLIFLIVIAL
bclaci.seq      .......... MILKNKRMLK IGICVGILGL SITSLEAFTG ESLQVEAKEK
blpenc.seq      .......... .....MKLWF STLKLKKAAA VLLFSCVALA GCANNQTNAS
Consensus       .......... ---------- ---------- ---------- ----------
               51                                                   100
pnsap.seq       VLSACNSNSS HAKELNDLEK KYNAHIGVYA LDTKSGKEVK FNSDKRFAYA
bclaci.seq      TGQVKHKNQA THKEFSQLEK KFDARLGVYA IDTGTNQTIS YRPNERFAFA
blpenc.seq      QPAEKNEKTE MKDDFAKLEE QFDAKLGIFA LDTGTNRTVA YRPDERFAFA
Consensus       ---------- -------le- ---a--g--a -dt------- -----rfa-a
              101                                                  150
pnsap.seq       STSKAINSAI LLEQVPYNKL NKKVHI..NK DDIVAYSPIL EKYVGKDITL
bclaci.seq      STYKALAAGV LLQQNSIDSL NEV..ITYTK EDLVDYSPVT EKHVDTGMKL
blpenc.seq      STIKALTVGV LLQQKSIEDL NQR..ITYTR DDLVNYNPIT EKHVDTGMTL
Consensus       st-ka----- ll-q-----l n----i---- -d-v-y-p-- ek-v-----l
              151                                                  200
pnsap.seq       KALIEASMTY SDNTANNKII KEIGGIKKVK QRLKELGDKV TNPVRYEIEL
bclaci.seq      GEIAEAAVRS SDNTAGNILF NKIGGPKGYE KALRHMGDRI TMSNRFETEL
blpenc.seq      KELADASLRY SDNAAQNLIL KQIGGPESLK KELRKIGDEV TNPERFEPEL
Consensus       -----a---- sdn-a-n--- --igg----- --l---gd-- t---r-e-el
              201                                                  250
pnsap.seq       NYYSPKSKKD TSTPAAFGKT LNKLIANGKL SKENKKFLLD LMLNNKSGDT
bclaci.seq      NEAIPGDIRD TSTAKAIATN LKAFTVGNAL PAEKRKILTE WMLGNATGDK
blpenc.seq      NEVNPGETQD TSTARALVTS LRAFALEDKL PSEKRELLID WMKRNTTGDA
Consensus       n---p----d tst--a---- l--------l --e----l-- -m--n--gd-
              251                                                  300
pnsap.seq       LIKDGVPKDY KVADKSGQAI TYASRNDVAF VYPKGQSEPI VLVIF.TNKD
bclaci.seq      LIRAGIPTDW VVGDKSGAG. SYGTRNDIAV VWP.PNSAPI I..VLISSKD
blpenc.seq      LIRAGVPDGW EVADKTGAA. SYGTRNDIAI IWP.PKGDPV VLAVL.SSRD
Consensus       li--g-p--- -v-dk-g--- -y--rnd-a- --p-----p- ---------d
              301                          327
pnsap.seq       NKSDKPNDKL ISETAKSVMK EF.....
bclaci.seq      EKEAIYNDQL IAEATKVIVK GS.....
blpenc.seq      KKDAKYDDKL IAEATKVVMK ALNMNGK
Consensus       -k-----d-l i-e--k---k -------
```

Notes:

1. pnsap = *Staph aureus* PC1 β-lactamase.
2. bclaci = *B. cereus* 569/H β-lactamase.
3. blpenc = *B. licheniformis* 749/C enzyme.
4. The sequence conservation amongst Gram-positive β-lactamases, extends into the carboxyl termini of the protein sequences.

TABLE 6.2: SIMILARITY MATRIX FOR CLASS A B-LACTAMASES

1. *Staph. aureus* PC1	-					
2. *B. cereus* 569/H	34	-				
3. *B. licheniformis* 749/C	37	48.4	-			
4. *E. coli* TEM	29	30.2	30	-		
5. *Ps. aeruginosa* RMS 149	27	34	30	40	-	
6. *Rps. capsulata* sp 108	21	26	26	36	35.5	-
	1.	2.	3.	4.	5.	6.

Generally, the alignment is very good, suggesting that these sequences are all derived from the same ancestral gene. Many of the invariant residues were retained and all the enzymes are approximately the same size. The invariant active site serine (Cartwright& Coulson, 1980) has been numbered 100 according to the recommendation of S.G Waley (Pers. comm.) so as to give a stable numbering system without numbers less than one. Although the proteins coded for by the genes were not studied for their amino acid content in these investigations, the alignment observed is good evidence that the gene sequences, and hence, the deduced amino acid sequences are correct.

As was earlier observed, with the class A enzymes (Coulson, 1985), the overall similarity is best in the amino-termini of the enzymes, the carboxy-termini showing much more variation. When the sequences are considered in groups of Gram-positive and Gram-negative derived enzymes, however, some more sequence conservation within the groups is observed and the carboxy-termini variability disappears in the Gram-positive enzymes (Fig. 6.1a and 6.1b).

Consequently, there seems to be a widespread distribution of class A β-lactamases in bacteria which have undergone independent evolution in the two bacterial lineages.

6.2.2. Similarity matrix of class A β-lactamases

A pairwise comparison of the characterised class A enzyme sequences was carried out and a similarity matrix table was drawn (Table 6.2). All amino acid residue matches were scored as 1 and mismatches were scored as 0. The percentage similarity was then calculated from the total nummber of matches divided by the number of residues in the longer of the pair.

As was expected from the observations in section 6.2.1 above, the *Bacillus* enzymes were more related to each other and at the same time closer to the

staphylococcal enzyme than to the Gram-negative ones.

6.2.3. Relationship to other β-lactamases and PBPs

No significant similarities were evident when the pseudomonad and rhodopseudomonad β-lactamases were compared with classes B, C and D enzymes (Jaurin *et al*, 1982; Dale *et al*, 1985). The only similarity they share with classes C and D as well as other penicillin binding proteins is the active site serine and its surrounding residues (Broome-Smith *et al*, 1985).

An attempt was made to take into consideration the extensive base substitution amongst the different class A enzymes in making the sequence comparison with other β-lactamases. Thus, a consensus string made from the alignment in Fig. 6.1 (by representing the invariant residue as their amino acid symbols and silent mutations according to the nature of the residue e.g. N for neutral, C for charged, H for hydrophobic and X for nonspecific amino acids), was compared to the classes C and D enzyme sequences. The similarity did not extend beyond the four amino acids on either side of the active site serine, which supports the hypothesis that the different classes result from a polyphyletic origin of bacterial β-lactamases.

6.2.4. Secondary structure predictions and tertiary structure

Of all the amino acids making up proteins, only a limited number will participate actively in the catalytic mechanisms of the enzymes, namely, those with appropriate side chains that act in proton transfer, as group acceptors, as nucleophiles, or electrophiles etc. Some amino acids, particularly those with hydrophobic side chains, may not be directly involved in catalysis. These, in conjunction with others, serve to determine the precise folding of the polypeptide chain. The numerous observations that after denaturation or dissociation there is spontaneous reconstitution of simple and complex proteins as well as organelles (e.g. ribosomes) support these suggestions (Nomura, 1973).

Natural selection involves the interaction of the organism with the environment. Therefore, evolution is driven partly by selection of the biological conformations best suited to specific functions. In homologous enzymes, it follows that changes permitted in one lineage should correspond to those in other lines if they still perform similar functions. Thus, we can compare conformations of homologous proteins for the purpose of confirming their primary sequence relationships as well as providing some explanations for the roles of the various amino acids.

The tertiary structure of proteins (the spatial arrangement of all atoms of a polypeptide chain) is the most complete information that could be obtained to make the basis of this comparison. Progress in determining the tertiary structure of β-lactamases has been slow despite the extensive sequence information now available, and the ease of purifying the protein. This has been attributed to the rather flexible nature of β-lactamase molecules (L. Sawyer, pers. comm.). Extensive secondary structural predictions derived by the hydrophobicity profiles and the modified Chou Fassman's methods, have been reported for the characterised class A enzymes (Bunster & Cid, 1984). There was a general consensus in the folding pattern observed in all the enzymes. The enzymes were predicted to consist of short helical and extended regions interspaced with random coils, as well as a large number of reverse turns around the active site serine. Based on these findings, three-dimensional models of the spatial distribution of the class A β-lactamase atoms were proposed. The general picture was that the class A enzymes consisted of two domains and that a charged loop surrounded the active site serine located in the amino-terminal domain I.

Secondary structure predictions were carried out independently for the *Ps. aeruginosa* RMS 149 plasmid and *Rps. capsulata* sp 108 β-lactamases using the program CHOUFAS (Garnier *et al*, 1978) based on basic Chou and Fassman (1977)

principles. The predictions for the two new class A β-lactamase sequences were essentially similar to one another and to those predicted for other class A enzymes. Interestingly, this conservation of secondary structure extends into the variable carboxyl-terminal regions, confirming that the sequences are homologous. Towards the end of this project, the tertiary structure of two class A enzymes were finally determined (Kelly *et al.*, 1986, for *B. licheniformis*, and Samraoui *et al.*, 1986, for the *B. cereus* enzyme).

The tertiary structures reported for both the *B. cereus* and *B. licheniformis* enzymes present an altogether different picture. There are no seperate domains and the molecules consist of five stranded β-pleated sheets flanked on one side by three α-helices and five helices on the other. It is of interest to note, that this arrangement is similar to the three-dimensional model proposed for domain II of the class A enzymes by Bunster and Cid (1984) from secondary structure predictions (six β-pleated sheets flanked by helices). The similarity may be purely coincidental and as such it is not clear what implications (if any) this may have.

The tertiary structures were also observed to share extensive regions of homology with the three dimensional structure of the penicillin-sensitive D-alanyl D-alanine carboxypeptidase-transpeptidase from *Streptomyces* R61. This finding has been interpreted to mean that the two groups of enzymes have diverged from a common ancestor.

6.2.5. Nucleic acid sequence relationships amongst class A β-lactamases

It was observed in the gene hybridization experiments (sections 4.5.2 and 5.7), that neither the pseudomonad nor the rhodopseudomonad β-lactamase genes hybridized to the other or to the TEM gene on pBR322. This is expected from the wobble nature of the genetic code which allows flexibility in the third base of a triplet so that it is often dependent on the G + C content of the organism (Dayhoff, 1972). The nucleic acid sequences of the two genes in these

FIG. 6.2: NUCLEOTIDE SEQUENCE ALIGNMENT FOR CLASS A B-LACTAMASE GENES

```
                1                                                    50
ppapen.DNA      GAGACCACGC TTCAGGCTAG GGTCGGAGTG GCGGTGATCG ATACGGATTC
rrcpen.DNA      GAGGAACAGC TCGGCGCCCG CGTCGGCCTC TCGCTCATGG AGACCGGCAC
rtem.DNA        GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG
blpenc.DNA      GAGGAACAAT TTGATGCAAA ACTCGGGATC TTTGCATTGG ATACAGGTAC
Consensus       ga-------- t----gc--- --t-gg---- -------t-g a---------
               51                                                   100
ppapen.DNA      CGGCCTGACG TGGCAG...C ATCGTGGCGA CGAACGCTTC CCGCTGAACA
rrcpen.DNA      GGGTTGGTCC TGGTCT...C ACCGCGAGGA CGAGCTTTTC CTCATGAACA
rtem.DNA        CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA
blpenc.DNA      AAACCGGACG ...GTAGCGT ATCGGCCGGA TGAGCGTTTT GCTTTTGCTT
Consensus       ------g--- ---------- --cg----ga -ga-c--tt- ----t-----
              101                                                  150
ppapen.DNA      GCACGCATAA AGCCTTTTCC TGCGCGGCCG TTCTGGCCCA GGCCGACCGC
rrcpen.DNA      GCACGGTCAA GGTGCCGGTC TGCGGCGCCA TCCTCGCGCG TTGGGACGCG
rtem.DNA        GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TGTTGACGCC
blpenc.DNA      CGACGATTAA GGCTTTAACT GTAGGCGTGC TTTTGCAACA GAAATCAATA
Consensus       --ac----aa -g-------- ---g--g--- t--t----c- ----------
              151                                                  200
ppapen.DNA      CACAAGCTGA ACCTGGAGCA GGCGATACCG ATCGAGCGCA CAGCGCTGGT
rrcpen.DNA      GGCAGGCTGT CGCTCTCCGA TGCGCTGCCG GTGCGCAAGG CCGACCTCGT
rtem.DNA        GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA ATGACTTGGT
blpenc.DNA      GAAGATCTGA ACCAGAGA.. ....ATAACA TATACACGTG ATGATCTTGT
Consensus       --------g- --c------- -----t---- ---------- --g---t-gt
              201                                                  250
ppapen.DNA      CACATACTCA CCCGTGACGG AAAGGGTGCC ACCTGGCGGC ACGCTGACCC
rrcpen.DNA      GCCCTACGCG CCCGTCACGG AGACGCGGGT C...GGCGGC AACATGACCC
rtem.DNA        TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ...ATGACAG
blpenc.DNA      AAACTACAAC CCGATTACGG AAAAGCACGT TGATACGGGA ...ATGACGC
Consensus       ----tac--- cc--t-ac-g a-a-g----- -------gg- ----tgac--
              251                                                  300
ppapen.DNA      TGCGTGAGCT GTGCAGGGCC GCCGTCAGTA TCAGTGACAA CACAGCGGCC
rrcpen.DNA      TCGACGAGCT CTGCCTCGCG GCGATCGACA TGAGCGACAA TGTGGCGGCG
rtem.DNA        TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC
blpenc.DNA      TCAAAGAGCT TGCGGATGCT TCGCTTCGAT ATAGTGACAA TGCGGCACAG
Consensus       t----ga--t -------gc- -c--t----- --ag-ga-aa ----gc----
              301                                                  350
ppapen.DNA      AATTTGGCGT TGGATGCAAT CGGCGGGGCA CGGACATTCA CCGCGTTCAT
rrcpen.DNA      AACATCCTGA TCGGGCATCT CGGGGGGCCG GAGGCGGTGA CGCAGTTCTT
rtem.DNA        AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT
blpenc.DNA      AATCTCATTC TTAAACAAAT TGGCGGACCT GAAAGTTTGA AAAAGGAACT
Consensus       aa--t----- t--------t -gg-gg--c- -------t-a ---------t
              351                                                  400
ppapen.DNA      GCGGTCTATC GGTGACGATA AGACACGCCT GGATCGGCGA GAACCCGAAC
rrcpen.DNA      CCGCAGCGTC GGCGACCCGA CGAGCCGTCT CGACCGCATC GAGCCCAAGC
rtem.DNA        GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
blpenc.DNA      GAGGAAGATT GGTGATGAGG TTACAAATCC CGAACGATTC GAACCAGAGT
Consensus       --------t- gg-ga----- --a-----c- -ga-cg---- ga-cc--a--
              401                                                  450
ppapen.DNA      TCAACGAGGC CACGCCGGGG GATGCACGCG ACACGACAAC GCCAATTGCG
rrcpen.DNA      TGAACGACTT CGCTTCTGGA GACGAGCGGG ACACCACGAG CCCGGCCGCC
rtem.DNA        TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA
blpenc.DNA      TAAATGAAGT GAATCCGGGT GAAACTCAGG ATACCAGTAC AGCAAGAGCA
Consensus       t-aa-ga--- -----c---- ga----c--g a-ac-a--a- --c----gc-
```

```
            451                                                    500
ppapen.DNA    GCAGCGCGGA GCCTGCAAAC ACTGTTGCTC GACGGTGTCC TCTCCGCTCC
rrcpen.DNA    ATGTCCGAGA CGCTGCGAGC GCTGCTGCTG GGCGACGTGC TGTCTCCGGA
rtem.DNA      ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC
blpenc.DNA    CTTGTCACAA GCCTTCGAGC CTTTGCTCTT GAAGATAAAC TTCCAAGTGA
Consensus     ---------a ---t-c---- --t------- g--g-----c t--c------
            501                                                    550
ppapen.DNA    GGCTCGGAAC GAACTGACAC AATGGATGCT CGGGGATCAA GTTGCCGATG
rrcpen.DNA    GGCCCGCGGG AAGCTGGCGG AGTGGATGCG CCACGGCGGC GTGACCGGCG
rtem.DNA      TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC
blpenc.DNA    AAAACGCGAG CTTTTAATCG ATTGGATGAA ACGAAATACC ACTGGAGACG
Consensus     ----cg---- ----t----- a-tggatg-- ---------- ------g---
            551                                                    600
ppapen.DNA    CCTTGCTACG CGCTGGCTTG CCGAGGGATT GGCAAATTGC GGACAAGTCG
rrcpen.DNA    CATTGCTGCG CGCCGAGGCC GAGGACGCCT GGCTGATCCT CGACAAGTCG
rtem.DNA      CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT
blpenc.DNA    CCTTAATCCG TGCCGGTGTG CCGGACGGTT GGGAAGTGGC TGATAAAACT
Consensus     c--t--t-cg --c-g----- --g---g--t gg----t--- -ga-aa--c-
            601                                                    650
ppapen.DNA    GGAGCAGGT. ..TGGTCACG GATCACGTTC CATAATCGCC GTTGTCTGGC
rrcpen.DNA    GGCAGCGGAA GCCACACGCG CAACCTCGTC GCGGTGATCC AGCCTGAAGG
rtem.DNA      GGAGCCGGT. ..GAGCGTGG GTCTCGCGGT ATCATTGCAG CACTGGGGCC
blpenc.DNA    GGAGCGGCA. ..TCATATGG AACCCGGAAT GACATTGCCA TCATTTGGCC
Consensus     gg----g--- ---------g ---------- ---------- ----------
            651                                                    700
ppapen.DNA    C...GCCCAA GCGTTCAGCC GTCATTGTGG CGATCTACAT CACCCAAACC
rrcpen.DNA    CGGAGCGCCC TGGATCGCGA CCATGTTCAT CTCGGATACG GACGCGGAGT
rtem.DNA      A...GATGGT AAGCCCTCCC GTATCGTAGT TATCTAC... ACGACGGGGA
blpenc.DNA    G...CCAAAA GGAGATCCTG TCGTTCTTGC AGTATTA... TCCAGCAGGG
Consensus     ---------- ---------- ---------- ---------- ----------
            701                                                    750
ppapen.DNA    GCAGCGTCTA TGTCGGCAAG CAACCAGGCG GTGTCCAGAA TCGGATCAGC
rrcpen.DNA    TCGAGGTTCG CAACGAGGCG CTCAAAGATC TGGGTAGGGC GGTGGTCGCG
rtem.DNA      GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC
blpenc.DNA    ATAAAAAGGA CGCCAAGTAT GATGATAAAC TTATTGCAGA GGCAACAAAG
Consensus     ---------- ---------- ---------- ---------- ----------
            751                      772
ppapen.DNA    CTTAGCAAAG GCGTTGCAAT GG
rrcpen.DNA    GTTGTTCGCG AATAG..... ..
rtem.DNA      TCACTGATTA AGCATTGGTA ..
blpenc.DNA    GTGGTAATGA AAGCCTTAAA CA
Consensus     ---------- ---------- --
```

Notes:

1. ppapen = *Ps. aeruginosa* RMS 149 β-lactamase gene.

2. rrcpen = *Rps. capsulata* sp 108 β-lactamase gene.

3. rtem = *E. coli* TEM 1 gene on pBR322 (Sutcliffe, 1978).

4. blpenc = *B. licheniformis* gene (Neugebauer *et al*, 1981).

5. The nucleotide sequences were aligned in the same way as were the protein sequences in Fig. 6.1. The variable leader peptide sequences were excluded.

6. The high nucleotide polymorphism observed for the conserved amino acid sequences suggest that the genes have diverged considerably.

investigations, as well as those of two other class A β-lactamases (the pBR322 amp and *B. licheniformis* β-lactamase genes), were aligned according to the protein sequence alignment in Fig. 6.1 and examined for nucleotide polymorphism. The principles used were the same as Kreitman's (1983) for alcohol dehydrogenase genes. The sequences before the first invariant glutamic acid were not included so as to avoid the effect of the different signal peptides.

As can be observed in Fig. 6.2, in 32 of the 38 invariant amino acids, all possible variations in the third position of the codons are present. Generally, there is a great variation in codon content. This would suggest one of two possibilities:-

- that the β-lactamase genes sequences being compared are ancient genes which diverged a long time ago, or
- that they are new but highly mutable, so that constraints on the structure of the protein molecule are the main force keeping them as they are.

Hirota (1984; β-lactamase conference, Holy Island) predicted an ancient origin for β-lactamases based on their primary sequence relationships to PBPs. However, the absence of a reliable molecular clock in bacteria makes it difficult to conclude whether they are ancient or modern genes on the basis of nucleotide polymorphism.

6.3. Summary

The two β-lactamases in this study belong to class A (Ambler classification, 1980). Amino acid and nucleotide sequences differ extensively in the class A enzyme from various bacterial sources. Secondary structure predictions and three-dimensional evidence, however, suggest that most of the mutations are conservative.

The occurence of a class A β-lactamase in a photosynthetic bacterium raises interesting possibilities for the functions and origins of β-lactamases. This will be

discussed further in chapter 7.

CHAPTER 7
DISCUSSION

7.1. Introduction

The investigations reported here concerned novel β-lactamases from two different bacteria. The DNA sequences were determined for the β-lactamases specified by *Ps. aeruginosa* RMS 149 plasmid and *Rps. capsulata* sp 108. The nucleotide sequences as well as the deduced amino acid sequences were compared with other characterised β-lactamases in order to investigate the relationship between these enzymes. This discussion will be focussed on the implications of these findings for the function and the evolution of β-lactamases.

7.2. *Ps. aeruginosa* RMS 149 β-lactamase

7.2.1. The structural gene

The amino acid sequence deduced from the pseudomonad plasmid gene sequence suggests a molecular weight of 31,283, much larger than that reported for the excreted enzyme (Sawada *et al.*, 1974). S1 mapping experiments suggest that the RNA transcripts sometimes stop in the middle of the gene. Protein expression studies have not yet been completed, so the possibility still exists that the final product is smaller due to the abortion of transcription or post-translational processing. Nevertheless, reports of incorrectly determined molecular weights are abundant in the literature (ref. section 1.3). The good alignment of the pseudomonad enzyme to other class A β-lactamases and the observation that the full transcripts are made, suggest that this may well be another example of the difficulty of determining molecular weights satisfactorily.

The similarities amongst the different class A β-lactamase sequences are consistent with divergence from a common ancestral gene (section 6.3). The differences are, however, so extensive that the β-lactamase gene on the RMS

149 plasmid is unlikely to be a recent transposition of the commonly occurring Gram-negative TEM enzyme gene onto a pseudomonad-specific plasmid. Sequences similar to the TnA repeat sequences were not observed in flanking DNA sequences.

Thus, the RMS 149 plasmid encoded β-lactamase might represent the first example of a Gram-negative penicillin-destroying enzyme which differs from the TEM enzyme in all its physiological characteristics but is genetically related.

7.2.2. Control of gene expression

The pseudomonad β-lactamase is constitutive at low levels (Sawada *et al*, 1974). However, its transcript has an attenuator sequence similar to those of inducible bacterial synthetic operons which have an upstream leader region capable of forming secondary structures (Yanofsky, 1981; Platt, 1981; etc.). The RMS 149 leader transcript does not code for a peptide. The situation is thus similar to the promoter region of the *E. coli* ampC β-lactamase (a class C enzyme), except that in the latter case, there is another Shine/Dalgarno sequence downstream of the leader transcript (Jaurin & Grundstrom, 1981).

In ampC gene expression, studies with a mutant in the leader transcript region that destabilizes the attenuator loop structure, suggest a growth rate-dependent regulation. This is postulated to be mediated via anti-termination due to increased ribosome binding at higher growth rates (Jaurin *et al*, 1982). This growth rate-dependent control is independent of the rel system (which mediates similar stringent control on the cell's translational machinery) as it is not relieved in rel$^-$ hosts (Jaurin & Normark, 1979). It seems likely that expression studies might reveal similar possibilities for the RMS 149 β-lactamase leader transcript.

Transcription studies (section 4.8), however, suggest that the attenuator is read through from an upstream gene which overlaps its promoter. Therefore, the

gene expression controls in the pseudomonad enzyme synthesis might be more complex, as a choice has to be made between reading through and stopping or initiating from the β-lactamase promoter. The low levels of expression and the presence of attenuator-like sequences upstream from the gene would seem to suggest that we are looking at a gene that has lost its inducibility, either in a control protein or in its sequences, such that it is now expressed constitutively. The gene is expressed in *E. coli* in sufficient quantities to protect the organism against 100 -500 µg/ml ampicillin (section 3.2). Hence, whatever the changes (if any), the effects have not been degenerative as far as penicillin resistance is concerned.

Nothing has been found in the literature about sequences upstream of the promoter regions of Gram-positive class A β-lactamases. So, no comparisons could be made. Genetic and biochemical analysis of β-lactamase expression in Gram-positive bacteria, however, suggest the involvement of two regulatory genes (Imsande, 1978). Sherrat and Collins (1973) analysed a set of *B. licheniformis* penicillinase mutants by transformation. They reported that a penicllinase repressor is specified by the pen I gene which is linked to the structural pen P gene. Another remote regulatory region which is not linked to pen P was also implicated. Mutations in the latter were observed to reduce the basal cell level of β-lactamase expression and completely blocked induction. A similar arrangement has been described for *Staph. aureus* β-lactamase by Cohen *et al*, (1970). It was suggested that whereas the penicillinase is plasmid encoded in this case, the unlinked regulatory gene was chromosomally encoded. The model proposed for regulation in the Gram-positive β-lactamases (Imsande, 1978) is that, the linked regulatory gene codes for a repressor protein which is inactivated by the β-lactam activated antirepressor protein, specified by the remote regulatory gene. Thus, in Gram-positive bacteria, there is evidence that class A β-lactamase expression is normally under very stringent controls.

7.3. *Rps. capsulata* sp 108 β-lactamase

Rps. capsulata sp 108 is the only isolate of a photosynthetic bacterium so far isolated that is resistant to penicillins through the possession of a β-lactamase. Early experiments showed it was difficult to isolate as a soluble protein (S. Scahill, Pers. comm.). The alternative strategy of isolating and characterising the gene was therefore adopted. The gene coding for the enzyme was located by recombinant DNA techniques and its nucleic acid sequence was determined using the Sanger dideoxy sequencing technique. The deduced amino acid sequence indicates that it is a class A enzyme (Ambler classification, 1980), similar to those commonly found in enterobacteria, staphylococci and *Bacillus* spp.

The sequence similarity extends right through the coding region, but a 36% exact similarity to the TEM enzyme (section 6.2.2) suggests that the two genes have diverged considerably. Unfortunately, there was not enough sequence information downstream of the β-lactamase gene to search for sequence relationship to the inverted repeats of TnA.

Southern hybridization experiments suggested a possible chromosomal location of the structural gene in both the inducible strain sp 108 and sp 109 (a penicillin sensitive derivative of the original isolate). The latter strain produces only basal levels of the enzyme (Scahill, 1981). The hybridization experiments did not reveal the presence of similar genes on *Rps. capsulata* (St Louis) or *Rps. sphaeroides*, which are both penicillin sensitive. In order to investigate the presence or absence of β-lactamase enzymes in these other photosynthetic bacteria, liquid cultures were grown up as described in section 2.2.1.2. The cells were lysed in a French Press (section 3.5) and the resulting cell lysates were tested for β-lactamase activity using the isoelectricfocusing method (Matthew *et al*, 1979). β-lactamase active bands were detected in both strains. If the penicillin sensitive sp 109 strain is indicative of the normal situation in photosynthetic bacteria, then the penicillin resistant phenotype of strain sp 108 might be due to

an overproduction of a class A β-lactamase which may be chromosomaly mediated.

The sequences involved in transcript initiation could not be unequivocally identified but initiation from an upstream gene was strongly suggested (section 5.11). Curiously, the organisation of genes around the rhodopseudomonad β-lactamase was very similar to that around the pseudomonad one (Fig. 5.9). The different URFs did not, however, match one another significantly, so it is not clear if the gene arrangement was merely coincidental. Since PSSURF-1 overlaps the β-lactamase gene, it will be interesting to see how many proteins are made in protein expression studies. It is possible that in sp 108, alterations have occurred so that the controls operate to favour β-lactamase expression.

7.4. Advantages in studying DNA sequence information directly

A satisfactory evaluation of evolutionary origins and relationships requires measurements of genetic variations in natural populations. In recent years, gross physical properties of macromolecules such as electrophoretic mobilities have allowed the screening of large samples, but protein sequences have proved more effective for estimating the extent of genetic variation at structural loci.

Recent advances in recombinant DNA technology mean that DNA sequences can be rapidly determined, thus making it possible for allelic differences between individuals to be directly identified at the nucleotide level. This complete resolution of genetic variation not only solves the problem of detecting all amino acid substitutions between the proteins, but also identifies nucleotide variation that is not translated into protein differences. The latter can be pointers to how long the genes have been diverging from one another.

Consequently, the methods used in these investigations have been highly satisfactory for obtaining sequence information. Direct protein sequence determination would otherwise have been difficult. Firstly, the respective

β-lactamases are produced in very small quantities in both organisms. Secondly, the rhodopseudomonad is difficult to grow, requiring a large surface area for exposure of small containers to light of the right wavelength and for periods up to one week. Finally, in order to increase yields, one would have had to generate magno-constitutive mutants. As plasmids and total DNA can be made from wet cell masses between 4 - 15 g (section 2.2.2), it was possible to obtain these quantities from manageable culture volumes.

β-lactamases afford an excellent selection marker and so the initial cloning was easy, once a vector was obtained which did not carry an amp gene. In my experience, locating the gene to as small a DNA fragment as possible proved to be the most difficult aspect. This was due to the poor distribution of restriction enzyme sites. Thus, during the DNA sequencing exercise, the Bankier/Barrell (1983) shotgun modification of the Sanger dideoxy DNA sequencing technique was very useful because randomly generated fragments could be cloned. The fragments were evenly distributed and as computer methods were used for sequence compilation, DNA sequencing was almost a joy.

A major advantage of studying DNA sequences directly, as opposed to amino acid sequence determination, was that the control sequences involved in gene expression were examined as well. An interesting revelation is the possible effect of the state of β-lactamase gene expression on the penicillin sensitivity or resistance phenotype of bacteria. Further reference will be made to this later in section 7.7.

7.5. Accuracy in DNA sequencing

The sequences presented in these investigations are believed to be correct, though there a number of ways in which errors can arise. The shotgun cloning of fragments to be sequenced allows the generation of random sequences without *a priori* knowledge. Thus, the very fact that they all match in a contiguous

fragment is an indication that the match is genuine. Each base is determined on the average three to four times. Whenever discrepancies arose between gel readings, the autoradiographs were re-examined and many differences were simple errors made during reading and were corrected.

The intensity of bands in sequencing autoradiographs varies, the C track being the most irregular. There are, however, rules to DNA sequencing band intensities which are well documented (Bankier & Barrell, 1983).

Compressions arising from the stacking of bases in G-C rich regions were another problem. Inspection of sequencing gels in the opposite orientation often allowed the establishment of the correct consensus as no region was found to be compressed on both strands.

Mutations arising in cloned DNA are a potential source of changes in the sequence. No differences ascribable to mutation were found between templates containing the same segments of DNA during the compilation of data. Therefore, neither single base changes nor DNA rearrangements occurred during the passaging and subcloning procedures. The possibility exists that some changes could have occurred in cloned DNA before the sequencing steps. It is unlikely that there were major/significant deletions or rearrangements, other than small ones, because no discrepancies were observed between the length of sequenced DNA and the sizes of recombinant DNA fragments deduced by electrophoretic methods.

The deduced amino acid sequences reported here should ideally be confirmed by some protein chemical analyses of the gene products. There are plans for doing this for both the pseudomonad and the rhodopseudomonad sequences in the near future. Nevertheless, the good alignment of the amino acid sequences of the newly sequenced β-lactamases with other characterised class A ones, suggest that the sequences are correct.

7.6. Evolution of β-lactamase activity

The β-lactamases produced by bacterial species are of several types. They are either metallo or serine enzymes. The more common types are the class A serine enzymes.

In the present study, the genes for two other class A enzymes from *Ps. aeruginosa* RMS 149 plasmid and *Rps. capsulata* sp 108 were investigated. The overall gene organization around both β-lactamase coding sequences is similar and RNA mapping studies indicate that they can be transcribed from upstream promoters. DNA sequence information suggests that the pseudomonad gene might be placed under some kind of transcriptional/translational control and that *Rps. capsulata* strains sp 108 and sp 109 represent the inducible and non-inducible versions of a similar chromosomal gene. Although it is not yet certain that *Rps. capsulata* sp 109 represents the original state of β-lactamase production in the photosynthetic bacterium, or if it is just a micro-constitutive mutant, these findings suggest that the current mode for screening for β-lactamase production may be inadequate.

Penicillin resistant colonies are initially picked from antibiotic agar plates. Organisms which grow are then analysed for β-lactamase activity. As was pointed out in section 1.1, the organisms which grow at the initial selection are those which have an effective permeability barrier or a penicillin hydrolysing/modifying enzyme. As a result, some "β-lactamase producing" isolates could represent some form of regulatory mutants which would be rare in natural populations. Reports that β-lactamase activity can be detected from most bacteria (Matthew & Harris, 1976; Hamilton-Miller & Smith, 1979) including very β-lactam sensitive species (Florey *et al*, 1947), seem to support this hypothesis.

Studies on the origin of new enzymatic activities in micro-organisms reveal that most occur in two stages. First, there are modifications in the regulation of

enzymes involved in other established pathways and later, in the structural genes themselves, to produce enzymes fitter for the new role(s) (Hall & Knowles, 1976; Hall & Zuzel, 1980; Mortlock, 1982). It has been postulated that this is as a result of performing these experiments with organisms that have evolved precise mechanisms for the regulation of synthesis and activity of their metabolic pathways. Thus, if an organism is capable of using an already present enzyme activity to establish growth under novel conditions, a regulatory mutation might be required to permit enzyme synthesis. Even if the enzyme activity can only be established by a mutation in the structural gene, another mutation in the regulation of the altered gene may also be required.

The initial selection of penicillin resistant bacteria from natural isolates is very similar to the laboratory conditions used in evolutionary studies (ref. section 6.2.2). For this purpose, bacteria are subjected to new compounds such that they must develop new capabilities to survive and grow. It is thus postulated here that, when bacteria are challenged with antibiotic, they may recruit some cellular proteins to destroy the toxins or block their entry. This hypothesis can be tested by searching for β-lactamase genes in penicillin sensitive bacteria. Chromosomal digests can be cloned into high expression vectors and successful recombinant plasmids/phages screened for β-lactamase activity. The presence of β-lactamase staining bands in cell lysates of photosynthetic bacteria observed in the course of these investigation (section 7.3), suggests that such experiments might prove fruitful.

7.7. Possible metabolic role for β-lactamases

As a corollary of the deduction in the last section, if β-lactamase activity is always present in bacteria, these proteins may perform some metabolic function. The only cellular metabolism so far associated with class A β-lactamase is spore coat formation in the Bacillaceae (Ozer and Saz, 1970). A β-lactamase negative mutant of *B. cereus* 569 H (*B. cereus* 569 H pen$^-$), which was impaired in

sporulation at the late steps V and VI (Murrel, 1967) so that matured spores were not released, could be rescued by exogenously added β-lactamase. There was then no delay at the late steps. The resulting spores appeared normal under electron microscopy and were as heat resistant as wild-type ones. The wide distribution of the class A enzymes amongst bacterial species, however, suggests that they might have some other universal function. Their sequence similarity with other PBPs around the active site serine (Broome-Smith *et al*, 1985) suggests the possibility that β-lactamases are evolutionarily derived from PBPs. This is seemingly supported by the observed tertiary structure similarity between Carboxypeptidase and *B. cereus* β-lactamase (Samraoui *et al*, 1986; Kelly *et al*, 1986). Thus, if they have any general physiological function, it is likely to be associated with cell wall metabolism.

Proteins with a similar molecular weight and activity profile have not, however, been identified in the β-lactam labelling experiments (ref. section 1.5.1). It is unlikely that such proteins would be discovered as the labelled β-lactam would be rapidly hydrolysed. This suggestion is further supported by similar labelling experiments where cefotaxime (a highly β-lactamase resistant β-lactam antibiotic; Richmond, 1980) was used instead of benzyl-penicillin and five PBPs different from those currently characterised were identified in *E. coli* (Labia *et al*, 1980).

Ambler's (1979) observation that the *Staph. aureus* class A β-lactamase copurifies with a small peptide similar to the cell wall crosslinking peptide has recently been repeated in this laboratory (I. Scragg, Pers. comm.). The association is not covalent as they can be separated by gel filtration in 50% formic acid (Ambler, 1979). Such observations have not been reported for other β-lactamases. The observations that classes A and C enzymes can hydrolyse and transacylate depsipeptide substrates (analogues of peptidoglycan crosslinking peptides), albeit slowly, (Pratt and Govandhan, 1984) would seem to suggest that

they are built for the slower events of transpeptidation, which according to Mirelman (1980), are important for maturation of the polymer. Such a function need not be crucial, especially in the stable conditions of laboratory growth.

7.8. Evolution of class A β-lactamase genes

The inference made in the preceding section is based on purely circumstantial evidence. Therefore, other β-lactamase regulatory sequences and gene arrangements need to be studied. Tentatively, it is suggested that there might be two gene systems for class A enzymes: one which is chromosomal and stringently controlled and another, (less controlled), which has been acquired by a transposon and transferred inter-species via plasmid conjugation. Fixation has been helped by medical and agricultural use of β-lactam antibiotics.

This prescribed system of events is not peculiar to class A β-lactamase genes. A recent study of the evolutionary relationships between class C chromosomal β-lactamases from *E. coli*, *Enterobacter cloacae*, *Citrobacter freundii* and *Ps. aeruginosa* (Normark & Lindberg, 1985) suggests a similar situation. In all but the *Citrobacter* sp, which is a soil isolate, the gene organization of the ampC gene is similar and its expression is repressed. Its promoter is buried within a fumarate reductase gene required for anaerobic respiration such that in the conditions of the gut, the ampC gene is rarely expressed. In *Citrobacter* sp, the gene control sequences are similar but there is another ampR gene preceding the β-lactamase structural gene. AmpR is a negative regulatory protein which represses β-lactamase synthesis in the absence of inducer. Though most ampC genes are chromosomally located, a plasmid-borne related gene has been reported (Jack & Richmond, 1970). Thus, the ampC gene of *Citrobacter* sp with its more complex organization has been suggested to be evolutionarily older than the *E. coli* one. Another evolutionary parallel is evident in the lac plasmids currently observed in a wide range of enterobacterial species (Guiso & Ullman, 1976). In one case, the lac region is

homologous to that of *E. coli* and it is borne on a transposable element (Cornelius *et al*, 1978).

7.9. Other possible lines of research

7.9.1. *Ps. aeruginosa* RMS 149 β-lactamase

Protein expression studies need to be done so as to provide enough protein for protein chemical studies. Thus, the deduced protein sequence could be confirmed, and the correct molecular weight of the exported protein can be determined. N-terminal amino acid analysis will provide information about the protein secretory mechanism used in this periplasmic enzyme. The attenuator sequences upstream of the structural gene might serve as an interesting system for a study. Mutants could be generated so as to allow the determination of the role of the stem-loop structure. Also, it will be interesting to find out how the expression of the β-lactamase gene is related to that of the upstream gene.

Thus, the pseudomonad β-lactamase gene might prove a useful system for examining the unique expression controls on genes with an unusual Shine/Dalgarno sequence to initiation codon space.

7.9.2. *Rps. capsulata* sp 108

More DNA sequence determination needs to be done for the rest of the 5.8 kb Bam H1 fragment, so that sequences flanking the β-lactamase genes can be examined for TnA repeats. Penicillin sensitive photosynthetic bacteria can be examined for genes homologous to class A β-lactamase genes as prescribed in section 7.6.

In the course of these investigations, gene hybridization was not detected between the *Rps. capsulata* sp 108 β-lactamase gene sequence and other penicillin sensitive *Rhodopseudomonas* spp. The lack of such reactivity with the related genes on the RMS 149 plasmid and pBR322 indicates that this is not

necessarily a detracting finding.

7.10. Conclusion

Studies in the function and evolution of enzymes are closely linked, because it is difficult to think of the use of an enzyme without considerations of the molecular mechanisms of the genetic changes which produced it (Pollock, 1971). Function is largely a subjective quality measurable only in relation to the limited criteria available in the laboratory. Yet in the neo-Darwinian understanding of evolution, the process is determined to a large extent by natural selection of organisms showing greater efficiency in new environments (in terms of greater chances of survival and more rapid growth).

Thus, all interpretations of function made here have been attempts to correlate the extensive distribution of β-lactamases and the known physiology of bacterial species. There is consensus in natural and experimental evidence that the ability to produce β-lactamase confers a selective advantage on bacteria. The studies reported here provide preliminary evidence that penicillin resistant or penicillin sensitive phenotypes may be partly dependent on the state of certain regulatory gene(s).

It has therefore been suggested that the different types of β-lactamases might be examples of cellular metabolic enzymes being recruited to tackle β-lactam toxins. It is hoped that comparative studies on other β-lactamase genes, particularly from non-pathogenic bacteria, will cast more light on the subject of β-lactamase evolution.

Bibliography

Abraham, E.P. & Chain, E.B. (1940) Nature (London) 146, 837

Adhya, S. & Gottesman, M. (1978) Ann. Rev. Biochem. 47, 967-996

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J.D. (1983) Molecular Biology of the Cell, pp 4-11 Garland Publishing Inc. New York & London

Ambler, R.P. (1975) Biochem. J. 151, 191-218

Ambler, R.P. (1976) in Structure-Function Relationships in Proteins (Markham, R. & Horne, R.W., eds.), pp 1-14, Elsevier-North Holland, Amsterdam

Ambler, R.P. (1979) in β-lactamases (Hamilton-Miller, J.M.T. & Smith, J.T., eds.), pp 99-125, Academic Press, London

Ambler, R.P. (1979a) in Evolution of Protein Molecules (Matsubara, H. & Yamanaka, T., eds.), pp 314-322, Japan Scientific Societies Press, Tokyo

Ambler, R.P. (1980) Phil. Trans. R. Soc. Lond. B 289, 321-331

Ambler, R.P. (1984) in Computer-Assisted Bacterial Systematics (Jones, D., Goodfellow, M. & Priest, F.G., eds.), pp 307-335, Cambridge University Press, Cambridge

Ambler, R.P. & Meadway, R. (1969) Nature (London) 222, 24-26

Ambler, R.P. & Scott, G.K. (1978) Proc. Natl. Acad. Sci. USA 75, 3732-3736

Ambler, R.P., Daniel, M., Hermoso, J., Meyer, T.E., Bartsh, R.G. & Kamen, M.D. (1979) Nature (London) 278, 659-660

Anderson, E.S. & Datta, N. (1965) Lancet 7382, 407-409

Anderson, S. (1981) Nucl. Acids Res. 9, 3015-3027

Anderson, J., Ptashne, M. & Harrison, S.C. (1984) Proc. Natl. Acad. Sci. USA 81, 1307-1311

Appleyard, R.K. (1954) Genetics 39, 440-452

Batchelor, F.R., Chain, E.B., Richards, M. & Rolinson, G.N. (1961) Proc. R. Soc. Lond. B 154, 5522-5531

Baldwin, G.S., Waley, S.G. & Abraham, E.P. (1979) Biochem. J. 179, 459-463

Bankier, A.T. & Barrell, B.G. (1983) in Techniques in Life Sciences, B5:Nucleic Acid Biochemistry (Flavell, R.A., ed.), B508, pp 1-34, Elsevier Scientific Publishers, Ireland

Bellard, G. (1973) Eur. J. Biochem. 36, 32-38

Benton, W.D. & Davies, R.W. (1977) Science 196, 180

Berg, A. & Bientema, J.J. (1975) Nature (London) 253, 207-210

Bergstrom, S., Lindberg, F.P., Olsson, O. & Normark, N. (1983) J. Bacteriol. 155, 1297-1305

Bergy's Manual of Systematic Bacteriology, (1984) Volume I (Editorial Board:R.G.E. Murray, chairman, Kreig, N.R. & Holt, J.G.), Williams & Wilkins, Baltimore & London

Bicknell, R., Emmanuel, E.L., Gagnon, J. & Waley, S.G. (1985) Biochem. J. 229, 791-797

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

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

Boyer, H.W. & Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472

Broda, P.M.A. (1979) Plasmids, W.H. Freeman, Oxford & San Francisco

Broome-Smith, J.K., Edelman, A., Yousif, S. & Spratt, B.G. (1985) Eur. J. Biochem. 147, 437-446

Brown, D. (1982) Gene 20, 139-141

Bryan, L.E., Shahrabadi, M.S. & Van den Elzen, H.M. (1974) Antimicrob. Agents Chemother. 6, 191-199

Bryson, V. & Vogel, H.J. (1965) Evolving Genes and Proteins, Academic Press, New York

Bunster, M. & Cid, H. (1984) FEBS Lett. 175, 267-274

Campbell, J.H., Lengyel, J.A. & Langridge, J. (1973) Proc. Natl. Acad. Sci. USA 70, 1841-1845

Carrey, E.A. & Pain, R.H. (1978) Biochim. Biophys. Acta 533, 12-22

Carrey, E.A., Mitchinson, C., Pain, R.H. & Virden, R. (1980) Phil. Trans. R. Soc. Lond. B 289, 372-373

Cartwright, S.J. & Coulson, A.F.W. (1980) Phil. Trans. R. Soc. Lond. B 289, 370-372

Charlier, P., Dideberg, O., Frere, J.-M, Moews, P.C. & Knox, J.R. (1983) J. Mol. Biol. 171, 237-238

Chou, P.Y. & Fassman, G.D. (1977) J. Mol. Biol. 115, 135-175

Citri, N. & Pollock, M.R. (1966) Adv. Enzymol. 28, 237-323

Clarke, P.H. & Richmond, M.H. eds., (1975) Genetics & Biochemistry of *Pseudomonas*, John Wiley & Sons, London and New York

Cohen, S., Vernon, E.G. & Sweeney, H.M. (1970) J. Bacteriol. 103, 616-621

Cohen, S.A. & Pratt, R.F. (1980) Biochemistry 19, 3996-4003

Compos, M., Ruiz, J., Marcela, M., Garces, E. & Ward, P. (1985) Comp. Biochem. Physiol. 80B, 45-49

Cornelius, G. Ghosal, D. & Saedler, H. (1978) Mol. Gen. Genet. 160, 215-224

Coulson, A.F.W. (1985) Biotechnol. Genet. Engng. Rev. 3, 219-253

Crick, F.H.C. (1958) Symp. Soc. Expt. Biol. 12, 138-163

Daldal, F., Cheng, S., Applebaum, J., Davidson, E. & Prince, R.C. (1986) Proc. Natl. Acad. Sci. USA 83, 2012-2016

Crick, F.H.C. (1966) Scientific American 215, 55-62

Dale, J.W. (1979) in β-lactamases (Hamilton-Miller, J.M.T. & Smith, J.T. eds.), pp 73-98, Academic Press, London & New York

Dale, J.W. & Smith, J.T. (1971) Biochem. J. 123, 507-512

Dale, J.W. & Smith, J.T. (1976) Biochem. Biophys. Res. Commun. 68, 1000-1005

Dale, J.W., Mossakowska, D., Stephenson, P. & Wall, S. (1985) FEBS Lett. 191, 39-44

Daniels, D., Schroeder, J., Szybalski, W., Sanger, F. & Blattner, F. (1983) in Lambda II (Hendrix, J.W., Roberts, F.W., Stahl, F. & Weisberg, R.A. eds.), pp 467-518, Cold Spring Harbour Laboratory

Datta, N. & Kontomichalou, P. (1965) Nature (London) 234, 222-241

Datta, N. & Richmond, M.H. (1966) Biochem. J. 98, 204-209

Dayhoff, M.O. & Eck, R.V. (1968) Atlas of Protein Sequence and Structure 1967-1968, National Biomedical Research Foundation, Silver Spring, Maryland

Dayhoff, M.O. (1969) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Silver Spring, Maryland

Dayhoff, M.O. (1972) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Silver Spring, Maryland

Davies, R.B., Abraham, E.P., Flemming, J. & Pollock, M.R. (1975) Biochem. J. 145, 409-411

DeLucia, M.L., Kelly, J.A., Mangeon, M.M., Moews, P.C. & Knox, J.R. (1980) Phil. Trans. R. Soc. Lond. B 289, 274-376

Dyke, K.G.H. (1979) J. Med. Microbiol. 2, 261-278

Edlund, T., Grundstrom, T. & Normark, S. (1979) Mol. Gen. Genet. 73, 115-125

Edlund, T. & Normark, S. (1981) Nature (London) 292, 269-271

Falsenstein, J. (1981) J. Mol. Evol. 17, 368-376

Farrell, P., Deininger, P.L., Bankier, A. & Barrell, B.G. (1983) Proc. Natl. Acad. Sci. USA 80, 1565-1569

Fisher, J., Belasco, J.G., Khosla, S. & Knowles, J.R. (1980) Biochemistry 19, 2895-2901

Fitch, W.M. (1966) J. Mol. Biol. 16, 9-16

Florey, H.W., Chain, E., Heatley, N.G., Jennings, M.A., Saunders, A.G., Abraham, E.P. & Florey, M.E. (1947) Antibiotics 2, 1090-1110, Oxford Univ. Press

Frere, J.-M. & Joris, B. (1985) Crit. Rev. Microbiol. 11, 299-396

Garnier, J., Osguthorpe, D.J. & Robson, B. (1978) J. Mol. Biol. 120, 97-120

Gest, H. (1974) J. Bacteriol. 88, 1201-1202

Ghysen, J.-M., Frere, J.-M, Leyh-Bouille, M., Coyette, J., Dusart, J. & Nguyen-Disteche, M. (1979) Ann. Rev. Biochem. 48, 73-101

Gibson, K.D. & Niedman, R.A. (1970) Arch. Biochem. Biophys. 141, 694-704

Gibson, J., Stackebrandt, E., Zahler, L., Gupta, R. & Woese, C.R. (1979) Curr. Microbiol. 3, 59-64

Gibson, T.J. (1984) Ph.D Thesis University of Cambridge

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

Gold, L., Stormo, G. & Saunders, R. (1984) Proc. Natl. Acad. Sci. USA 81, 7061-7065

Groove, D.C. & Randal, W.A. (1955) in Assay Methods of Antibiotics p 16, Medical Encyclopedia Inc. Chicago

Guiso, N. & Ullman, A. (1976) J. Bacteriol. 127, 691-697

Hall, A. & Knowles, J.R. (1976) Nature (London) 264, 803-804

Hall, B.G. & Hartl, D.L. (1974) Genetics 76, 391-400

Hall, B.G. & Zuzel, T. (1980) Proc. Natl. Acad. Sci. USA 77, 3529-3533

Hamilton-Miller, J.M.T. (1982) J. Antimicrob. Chemother. 9, 87-90

Hamilton-Miller, J.M.T. & Smith, J.T. eds., (1979) β-lactamases, Academic Press, London

Hanahan, D. (1983) J. Mol. Biol. 166, 557-580

Hansen, F.G. & Von Meyenberg, K. (1974) Mol. Gen. Genet. 175, 135-144

Hedges, R.W. & Jacob, A.E. (1974) Mol. Gen. Genet. 132, 31-40

Hedges, R.W. & Matthew, M. (1979) Plasmid 2, 269-278

Hegeman, G.D. & Rosenberg, S.L. (1970) Ann. Rev. Microbiol. 24, 429-462

Heffron, F., McCarthy, J.B., Ohtsubo, H. & Ohtsubo, E. (1979) Cell 18, 1153-1163

Hennessey, T.D. (1967) J. Gen. Microbiol. 49, 277-285

Hill, P. (1970) M.Sc. Thesis, Edinburgh University

Holloway, B.W. (1969) Bacteriol. Revs. 33, 419-443

Holloway, B.W. (1973) in First International Symposium on the Genetics of Industrial Micro-organisms (Vanek, Z., Hostalek, Z. & Cudlin, J., eds.), pp 45-53, Academia Prague

Hu, & Messing, (1982) Gene 17, 271-277

Huang. H.T., Seto, T.A. & Shull, G.M. (1963) Appl. Microbiol. 11, 1-6

Hussein, M., Carlino, A., Madonna, J.M. & Lampen. O. (1985) J. Bacteriol. 164, 223-229

~~Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C. & Rupley, J.A. (1972) Enzymes 7, 665-870~~

Imsande, J. (1978) Microbiol. Revs. 42, 67-83

Izaki, K., Matsuhashi, M. & Strominger, J.L. (1968) J. Biol. Chem. 243, 3180-3192

Jack, G.W. & Richmond, M.H. (1970) J. Gen. Microbiol. 61, 43-61

Jacob, F. & Monod, J. (1961) J. Mol. Biol. 3, 318-356

Jacoby, G.A. (1974) Antimicrob. Agents Chemother. 6, 239-252

Jacoby, G.A. (1977) in Microbiology - 1977 (Schlessinger, D., ed.), pp 119-126 American Society for Microbiology, Washington DC

Jacoby, G.A. (1979) in *Ps. aeruginosa*: Clinical Manifestations of Infection & Current Therapy (Doggert, D. ed.), pp 271-309, Academic Press, New York

Jan de Bont, A.M., Scholen, A., & Hansen, T. (1981) Arch. Microbiol. 128, 271-274

Jaurin, B. & Normark, S. (1979) J. Bacteriol. 138, 896-902

Jaurin, B. & Grundstrom, T., (1981) Proc. Natl. Acad. Sci. USA 78, 4897-4901

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

Kelly, J.A., Dideberg, O., Charlier, P., Wery. J.P., Libert, M., Moews, P.C., Knox, J.R., Dioz, C., Fraipont, C., Joris, B., Dusart, J., Frere, J.-M. & Ghysen, J.-M. (1986) Science 231, 1429-1430

~~Koster, W.R. & Matthews, B.W. (1977) J. Biol. Chem. 252, 7704-7710~~

Kirby, W.M. (1944) Science 99, 452-453

Kirby, K.S. (1956) Biochem. J. 64, 405-408

Klein, B. & Murray, K. (1979) J. Mol. Biol. 133, 289-294

Knott-Hunziker, V., Waley, S.G., Orlek, B.S. & Sammes, P.G. (1979) FEBS Lett. 99, 59-61

Knott-Hunziker, V., Petursson, S., Waley, S.G., Jaurin, B. & Grundstrom, T. (1982) Biochem. J. 207, 315-322

Kogut, M., Pollock, M.R. & Tridgell, E.J. (1956) Biochem J. 62, 391-401

Kreft, J., Burger, J. & Goebel, N. (1983) Mol. Gen. Genet. 190, 384-389

Kreitman, M. (1983) Nature (London) 304, 412-417

Labia, R., Kazmierczak, A., Guionie, M. & Massou, J.M. (1980) J. Antimicrob. Chemother. 6, Suppl. A, 19-23

Lai, J.S., Sarvas, M., Brammer, W.J., Neugebauer, K. & Wu, H.C. (1981) Proc. Natl. Acad. Sci. USA 78, 3506-3510

~~Lampen, J.O., Nielsen, J.B.K., Izui, K. & Caulfield, M.P. (1980) Phil. Trans. R. Soc. Lond. B 289, 345-348~~

Lederberg,J. (1956) Proc. Natl. Acad. Sci. USA 42, SM4

Lennox, E.S. (1955) Virology 1, 190-206

Lindberg, F., Lund, B. & Normark, S. (1986) Proc. Natl. Acad. Sci. USA 83, 1891-1895

Lowry, O.H., Rosebrough, N.J., Farr, A.L. & Randall, R.J (1951) J. Biol. Chem. 193, 265-275

Macleod, A.R., Karn, J. & Brenner, S. (1981) Nature (London) 291, 386-390

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor,

Markland, F.S. & Smith,E.L. (1971) Enzymes 3rd edn. 3, 561

Marmur, J. (1961) J. Mol. Biol. 3, 208-218

Marrs, B. (1974) Proc. Natl. Acad. Sci. USA 71, 971-973

Matthew, M. (1978) FEMS Microbiol. Letts. 4, 241-244

Matthew, M. (1979) J. Antimicrob. Chemother. 5, 349-358

Matthew, M., Harris, A.M., Marshall, M.J. & Ross, G.W. (1975) J. Gen. Microbiol. 88, 169-178

Matthew, M. & Harris, A.M. (1976) J. Gen. Microbiol. 94, 55-67

Matthew, M. & Hedges, R.W. (1976) J. Bacteriol. 125, 713-718

Matthew, M., Hedges, R.W. & Smith, J.T. (1979) J. Bacteriol. 138, 657-662

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

Maxam, A.M. & Gilbert, W. (1980) Meth. Enzymol. 65, 499-560

McPhail, M. & Furth, A.J. (1973) Biochem. Soc. Trans. 1, 1260-1263

Mehta, R.J. & Nash, C.H. (1978) J. Antibiot. 31, 239-240

Mendel, R.J. & Higa, A. (1970) J. Mol. Biol. 53, 159-162

Messing, J. & Viera, J. (1982) Gene 19, 269-276

Meynell, E., Meynell, C.C. & Datta, N. (1968) Bacteriol. Revs. 32, 55-83

Miller, S. & Orgel, L.E. (1973) The Origins of Life, Prentice-Hall, Englewood Cliffs, New Jersey

Mirelman, D. (1980) in Bacterial Outer Membranes (Inoyue, M. ed.), pp 115-166, John Wiley & Sons Inc. New York

Mizuno, T. & Kageyama, M. (1979) J. Biochem. 84, 179-191

Moore, R. (1980) Ph.D Thesis, Monash University, Australia

Mortlock, R.P. (1982) Ann. Rev. Microbiol. 36, 259-284

Murray, N.E. (1983) in Lambda II (Hendrix, J.W., Roberts, F.W., Stahl, F. & Weisberg, R.A. eds.), pp 395-432, Cold Spring Harbour Laboratory

Murray, N.E., Brammer, W.J. & Murray, K. (1977) Molec. Gen. Genet. 150, 53-61

Murray, N.E. & Kelly, W.S. (1979) Molec. Gen. Genet. 175, 77-87

Murrell, W.G. (1967) Adv. Microbial Physiol. 1, 133-251

Neu, H.C. & Chou, J. (1967) J. Bacteriol. 94, 1934-1945

Neugebauer, K., Sprengl, R. & Schaller, H. (1981) Nucl. Acids Res. 9, 2577-2588

Nielsen, J.B. (1983) Meth. Enzymol. 97, 153-158

Nielsen, J.B., Caulfield, M.P. & Lampen, J.O. (1981) Proc. Natl. Acad. Sci. USA 78, 3511-3515

Nielsen, J.B. & Lampen, J.O. (1983) Biochemistry 22, 4652-4656

Nomura, M. (1973) Science 179, 864-873

Normark, S. & Lindberg, F. (1985) Bio-Essays 3, 22-26

Novick, R.P. (1969) Bacteriol. Revs. 33, 210-257

O'Callaghan, C.H., Muggleton, P. & Ross, G. (1968) in Antimicrobial Agents and Chemotherapy (Hobby, G. ed.), p 57, American Society for Microbiology

O'Callaghan, C.H., Morris, A., Kirby, S. & Shingler, A.H. (1972) Antimicrob. Agent. Chemother. 1, 283-288

Oegema, T.R., Hascall, V.C. & Dziewiatkowski, D. (1975) J. Biol. Chem. 250, 6151-6159

Oparin, A.I. (1965) Life, Its Origin, Nature and Development, Academic Press Inc., New York

Ozer, J.H. & Saz, A.K. (1970) J. Bacteriol. 102, 65-71

Parkinson, J.S. (1968) Genetics 59, 311-325

Peacock, D. (1981) in Biochemical Evolution (Gutfruend, H. ed.), pp 88-115, Cambridge University Press, London

Perret, C.J. (1954) Nature (London) 174, 1012-1013

Pfenning, N. (1977) Ann. Rev. Microbiol. 31, 275-290

Platt, T. (1981) Cell 24, 10-23

Pollock, M.R. (1950) Br. J. Exp. Pathol. 31, 739-753

Pollock, M.R. (1956) J. Gen. Microbiol. 15, 154-169

Pollock, M.R. (1960) Enzymes 4, 269-278

Pollock, M.R. (1971) Proc. R. Soc. Lond. B 179, 385-401

Pratt, R.F. & Govandhan, C.P. (1984) Proc. Natl. Acad. Sci. USA 81, 1302-1306

Pribnow, D. (1979) in Biological Regulation and Development (Goldberger, R.F. ed.), 1, pp 219-277, Plenum Press, New York

Queen, C. (1983) J. Mol. & Appl. Genet. 2, 1-10

~~Richmond, M. H. (1965) J. Bacteriol. 90, 370-374~~

Richmond, M.H. (1980) J. Antimicrob. Chemother. 6, Suppl A, 33-35

Richmond, M.H. & Sykes, R.B. (1973) Adv. Microbial Physiol. 9, 255-261

Richmond, M.H., Bennett, P.M., Choi, C.L., Brown, N., Brunton, J., Grinsted, J. & Wallace, I. (1980) Phil. Trans. R. Soc. Lond. B 289, 349-359

Rigby, P.W.J., Dieckman, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. 113, 237-251

Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D.L. (1978) Nature (London) 272, 414-423

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

Rossman, M.G. & Argos, P. (1977) J. Mol. Biol. 109, 99-129

Sabath, L.D., Jago, M. & Abraham, E.P. (1965) Biochem. J. 96, 739-404

Saino, Y., Kobayashi, F., Inonye, M. & Mitsuhashi, S. (1982) Antimicrob. Agents Chemother. 22, 564-570

Salsier, W.L., Gesteland, R. & Bolle, A. (1967) Nature (London) 215, 588-591

Samraoui, B., Sulton, B.J., Todd, R.J., Artymuik, P.J., Waley, S.G. & Phillips, D.C. (1986) Nature (London) 320, 378-380

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

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. & Roe, B.A. (1980) J. Mol. Biol. 143, 161-178

Sanger, F., Coulson, A.F.R., Hong, G.F., Hill, D.F. & Petersen, G.B. (1982) J. Mol. Biol. 162, 729-773

Saunders, V.A., Saunders, J.R. & Bennett, P.M. (1976) J. Bacteriol. 125, 1180-1187

Sawada, Y., Yaginuma, S., Tai, M., Iyobe, S. & Mitsuhashi, S. (1974) in Microbial Drug Resistance (Mitsuhashi, S. & Hashimoto, H. eds.), pp 381-397 Univ. of

Tokyo Press, Tokyo

Saz, A.K. & Lowery, D.L. (1979) in β-lactamases (Hamilton-Miller, J.M.T. & Smith, J.T. eds.), pp 465-479, Academic Press, London

Scahill, S. (1981) Ph.D Thesis, University of Edinburgh

Scherer, G.F., Walkinshaw, M.D., Arnott, S. & Morre, D.J. (1980) Nucl. Acids Res. 8, 3895-3907

Scott, G.K. (1972) Ph.D Thesis, University of Edinburgh

Sharp, P.A., Sugden, B. & Sambrook, J. (1973) Biochemistry 12, 3055-3063

Shaw, D.J. & Guest, J.R. (1982) Nucl. Acids Res. 10, 6119-6130

Sherratt, D.J. & Collins, J.F. (1973) J. Gen. Microbiol. 76, 217-230

Shine, J. & Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1343-1346

Sistrom, W.R. (1977) J. Bacteriol. 131, 526-532

Smith, E.L. (1970) Enzymes 3rd edn. 1, 267-339

Smith, H.O. & Nathans, D. (1973) J. Mol. Biol. 81, 419-423

Sneath, P.H.A. (1974) Symp. Soc. Gen. Microbiol. 24, 1-39

Sojka, G.A. Din, G.A. & Gest, H. (1967) Nature (London) 261, 1021-1022

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

Southern. E.M. (1979) Meth. Enzymol. 68, 152-176

Spratt, B.G. (1975) Proc. Natl. Acad. Sci. USA 72, 2999-3003

Spratt, B.G. (1977) Eur. J. Biochem. 72, 341-352

Spratt, B.G. (1978) Nature (London) 274, 713-715

Spratt, B.G. & Pardee, A.B. (1975) Nature (London) 254, 516-517

Spratt, B.G., Boyd, A. & Stoker, N. (1980) J. Bacteriol. 143, 568-591

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

Staden, R. (1982a) Nucl. Acids Res. 10, 4731-4751

Staden, R. (1984) Nucl. Acids Res. 12, 499-503

Staden, R. (1984a) Nucl. Acids Res. 12, 551-567

Staden, R. & McLachlan, A.D. (1982) Nucl. Acids Res. 10, 141-156

Stoker, N.G., Fairweather, N.F. & Spratt, B.G. (1982) Gene 18, 335-341

Stormo, G.D., Schneider, T. & Gold, L. (1982) Nucl. Acids Res. 10, 2971-2996

Stuber, D. & Bunjard, H. (1981)Proc. Natl. Acad. Sci. USA 78, 167-171

Sutcliffe, J.G. (1978) Proc. Natl. Acad. Sci. USA 75, 3737-3741

Thatcher, D.R. (1975) Biochem. J. 149, 313-326

Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205

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

Tipper, D.J. & Strominger, J.L. (1965) Proc. Natl. Acad. Sci. USA 54, 1133-1141

Tomasz, A., Albino, A. & Zanati, E. (1970) Nature (London) 227, 138-140

Tomasz, A. (1980) Phil. Trans. R. Soc. Lond. B 289, 303-308

Tybulewicz, V.L.J., Falk, G. & Walker, J.E. (1984) J. Mol. Biol. 179, 185-214

Van Niel, C.B. (1944) Bacteriol. Revs. 8, 1-118

Waley, S.G. & Cartwright, S.J. (1983) Medicinal Res. Revs. 3, 341-382

Wall, J.D., Weaver, P.F. & Gest, H. (1975) Arch. Microbiol. 105, 217-224

Waxman, D.J. & Strominger, J.L. (1980) J. Biol. Chem. 255, 3964-3976

Waxman, D.J. & Strominger, J.L. (1983) Ann. Rev. Biochem. 52, 825-869

Weaver, P.F., Wall, J.D. & Howard, G. (1975) Arch. Microbiol. 105, 207-216

Weil, J., Cunningham, R.M., Mitchell, B. & Bolling, B. (1973) Virology 50, 373-380

Wharton, R.P., Brown, E.L. & Ptashne, M. (1984) Cell 38, 361-369

Woese, C.R. & Stackebrandt, E. (1981) Symp. Soc. Gen. Microbiol. 32, 1-31

Yanofsky, C. (1981) Nature (London) 289, 751-758

Zukerkandl, E. & Pauling, L. (1965) J. Theor. Biol. 8, 357-366