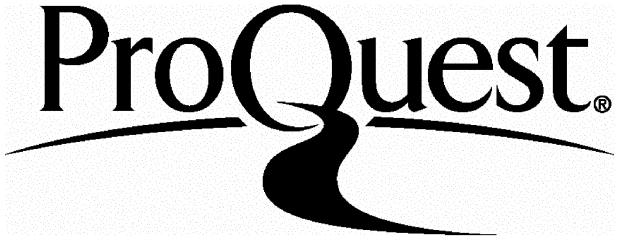
Regulation of the *Pseudomonas aeruginosa* Amidase Operon by Transcription Antitermination

by

Munuhe Simon James Wachira

A thesis submitted for the Degree of Doctor of Philosophy in the University of London

Department of Biochemistry and Molecular Biology
University College London
January 1994



Abstract

Previous studies have identified five genes, *amiE* (amidase structural gene), *amiB*, *amiC*(negative regulator gene), *amiR* (positive regulator gene) and *amiS*, in the inducible amidase operon of *Pseudomonas aeruginosa*. The amidase operon is regulated in part by transcription antitermination upstream of the operon at a Rho-independent terminator located between *amiE* and its promoter. In this study, the induction of amidase expression and the mechanism of transcription antitermination have been investigated.

An inducible amidase expression system has, for the first time, been reconstituted in *E. coli* by making a coordinate expression vector for *amiC* and *amiR*. The *amiC/R* expression vector was also used to restore an amidase inducible phenotype in a *Pseudomonas aeruginosa* amidase-constitutive mutant and in an amidase *amiC*$^-$, *amiR*$^-$ negative mutant. A deletion mutation within the *amiB* gene, previously thought to be involved in amidase induction, led to no change in amidase phenotype. It is now thought that *amiB* codes for the energy transducing component of an ABC type transporter system. The sequence of *amiS*, the potential membrane component of the putative ABC transporter has been confirmed.

In an attempt to purify AmiR for *in vitro* studies of the transcription antitermination mechanism, an AmiR overproducing plasmid has been constructed based on a T7 expression system. AmiR has been expressed using this system to about 20% of total cellular protein. Purification of active protein was, however, not achieved due to poor binding to ion exchange columns and aggregation. Identity of the protein was however confirmed by analysis on SDS-polyacrylamide gels, Western blotting using rabbit antisera raised against an MalE-AmiR fusion and by N-terminal sequencing.

The transcription antitermination mechanism has been investigated both *in vivo* and *in vitro*. Firstly, overexpression of the leader sequence *in trans* has been shown to reduce the direct AmiR interaction with the leader mRNA. Secondly, by cloning various lengths of the amidase leader region upstream of *amiR* followed by *in vivo* competition studies, it has been shown that AmiR mediates transcription antitermination by binding to the leader transcript within a region that runs up to the terminator upstream face.

In vitro RNA/protein interaction studies have been carried out. Bandshift studies showed complex formation between ^{32}P labelled leader RNA and a semipurified *E. coli* extract containing AmiR. The specific nature of the nucleoprotein complex was demonstrated by cold competition experiments. Bandshift experiments using transcripts encompassing different lengths of the leader region identified a 58b region running from +10 to +68 from the transcription start site as the binding region. The precise binding site was not determined.

Extensive site directed mutagenesis within the leader region has identified sequences that are critical for the antitermination reaction without any effect on the termination reaction, and other sequences that made the terminator 'leaky' without any apparent effect on antitermination.

Acknowledgments

I would like to thank the following people for their assistance: Ms Rachel Williams for helping with the thesis preparation; Mr Brian Coles for N-terminal sequencing of AmiR; Dr Roger Pennel for assistance with the immunoblotting; Dr John Ward for useful discussions; Dr Laurence Pearl for useful discussions and the use of his laboratory for carrying out some experiments, and Dr Chris Prodromou for introducing me to DNA sequencing. Special thanks go to Dr Stuart Wilson for the collaboration in the AmiR purification attempts, and invaluable assistance with the RNA bandshift assays.

Most of all I would like to acknowledge my supervisor Dr Robert Drew for guiding me through the project, expeditious reading of the manuscript and assistance with the administrative matters.

Last but not least I would like to thank my family, for unfailing support, and all the friends I have made in England.

This work was supported by a grant from Overseas Development Administration through the British Council.

Dedicated to Phyllis Wangari and Tim Wachira

Contents

List of Figures

List of Tables

Abbreviations

ATP	Adenosine triphosphate	kDa	KiloDalton
CTP	Cytosine triphosphate	Mr.	Relative molecular weight
GTP	Guanosine triphosphate	HTH	Helix turn helix
ITP	Inosine triphosphate	SDS	Sodium dodecyl sulphate
TTP	Thymidine triphosphate	TCA	Trichloroacetic acid
NTP	Nucleoside triphosphate	EDTA	Ethylenediaminetetraacetic acid
cAMP	Cyclic Adenosine triphosphate	PMSF	Phenylmethylsulphonyl-fluoride
DNA	Deoxyribonucleic acid	DEPC	Diethylpyrocarbonate
RNA	Ribonucleic acid	X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
ABC	ATP binding cassette-		
b	base	IPTG	Isopropyl-β-D-thiogalactopyranoside
bp	base pair	PEP	Phosphoenolpyruvate
kb	Kilo bases	PEG	polyethylene glycol
DS	Double stranded	SSC	Sodium saline citrate
SS	Single stranded	TE	Tris-EDTA
ORF	Open reading frame	PNK	Polynucleotide Kinase
SD	Shine-Dalgarno	CIAP	Calf intestinal alkaline
RBS	Ribosome binding site		phosphatase
MCS	Multiple cloning site	WT	Wild type
RAT	Ribonucleic antiterminator	Inc	incompatibility
TAS	Transactivation sequence	UV	Ultra violet
MPP	Miniplasmid preparation	2D	2-dimensional
PCR	Polymerase chain reaction	psi	Pounds per square inch
Ap/Amp	Ampicillin	HIV	Human
Cm	Chloramphenicol		immunodeficiency virus

CHAPTER 1

Introduction

1.1 Pseudomonas Biology

The pseudomonads have been widely studied and characterized because of their diverse properties. Some species are pathogenic to animals including humans and others to plants. Species with the ability to degrade environmentally harmful natural and synthetic compounds have been isolated and characterized. *Pseudomonas* species are capable of using a wide variety of organic substances as growth substrates and as such are useful in studying metabolic pathways, gene structure and function. Growth on many of these substances occurs via the induction of catabolic enzymes which generate substrates for central metabolic pathways, mainly the TCA cycle for energy metabolism. The TCA cycle is the central energy metabolism pathway and its intermediates exert a powerful catabolite repression on inducible pathways including glucose metabolism.

In cytological terms , *Pseudomonas* strains are defined as nonsporulating gram-negative rods with polar and sometimes subpolar flagella (Palleroni, 1984). Phylogenetic classification of *Pseudomonas* has been made possible by determination of homology between the highly conserved rRNAs (Palleroni, 1992; Palleroni, 1984). The *Pseudomonas* have been placed into 5 rRNA homology groups determined by rRNA oligoribonucleotide fragments/DNA hybridisation and competition experiments. *Ps. aeruginosa* is the type species of group I which includes the fluorescent *Pseudomonas* and some nonfluorescent species. This group has been confirmed retrospectively by the

immunological cross-reactivity of enzymes in the aromatic β–ketoadipate degradative pathway (Stanier *et al.*, 1970), cytochrome *c*-551 amino acid sequences (Ambler and Wynn, 1973) and DNA hybridisation data (Palleroni *et al.*, 1972). Some members of group I produce pigments some of which fluoresce under UV irradiation (fluorescent *Pseudomonas*). Many pigments are strong iron chelators which function as siderophores and have been found to be species specific. These siderophores are associated with reduced populations of pathogenic fungi and bacteria in strains of *Ps. putida* and *Ps. fluorescens* with plant growth enhancing effect (Loper *et al.*, 1984; Leong *et al.*, 1992). The opportunistic human pathogen *Ps. aeruginosa* is able to produce copious amounts of the polysaccharide alginate which causes respiratory distress and is a major factor in the morbidity and mortality cystic fibrosis patients (Goldberg, 1992).

RNA homology group II is entirely composed of pathogenic species (Palleroni, 1992); *Ps. mallei* and *Ps. pseudomallei* infect animals, and *Ps. cepacia, Ps. gladioli, Ps. caryophilla* and *Ps. solanacerium* are all plant pathogens. Ribosomal RNA group III includes *Ps. acidovoras* and *Ps. testosteroni.* Members of rRNA homology group IV *Ps. dimunita* and *Ps. vesicularis* although currently allocated to the genus *Pseudomonas* are very distantly related to the *Pseudomonas* Sensu Stricto (Group I) (Palleroni, 1992) and group V is the *Xanthomonas* genus. In addition classification has been carried out by investigating the cofactor specificity and allosteric control pattern for L-tyrosine (Byng *et al.*, 1980) and L-phenylalanine biosynthesis. These results have confirmed the rRNA homology groups and allowed lower level classification (Whitaker *et al.*, 1981).

1.2 Pseudomonas Genetics

1.2.1 Chromosome Mapping

Genetic mapping in *E. coli* and the demonstration of linkage map circularity was greatly facilitated by the sex factor, plasmid F. F integrates into the bacterial chromosome producing high frequency of recombination (Hfr) donors capable of polarized mobilization of markers from the point of integration (Hayes, 1968). Chromosome mapping in the *Pseudomonas* has been more difficult because the naturally occurring FP plasmids do not display marker transfer kinetics comparable to the F plasmid (Holloway and Morgan, 1986). A limited map was initially assembled for *Ps. aeruginosa* strain PAO using the naturally occurring plasmid FP2 which also confers resistance to mercuric ions but poor marker recovery made it unsuitable for mapping markers later than 25 minutes. A derivative of the IncP-I plasmid R68 carrying a duplication of the IS21 sequence (R68.45) has been found to function well in the construction of Hfr strains (Currier and Morgan, 1982). R68.45 has subsequently been used for chromosome mapping of many bacteria due to the wide host range specificity of the IncP-I replicon (Haas, 1983).

Linkage map circularity has been established for *Ps. aeruginosa* PAO and PAT, and the position of many loci has been determined (Holloway and Morgan, 1986). A variety of techniques were used in these investigations including transposon mutagenesis using Tn5 and Tn2521. Gene order was then determined by conjugation marker transfer in interrupted mating experiments combined with transduction analysis using bacteriophages F116 and G101 to determine the order of closely linked genes (Rothmel *et al.*, 1991). Chromosome linkage map circularity has been demonstrated, and several markers mapped, in *Ps. putida* PPN by time of entry kinetics using a mutant IncP-10 plasmid, R91-5,

loaded with transposon Tn501, to generate Hfr donors with a polarized chromosome transfer ability (Dean and Morgan, 1983). A number of markers have been mapped in *Ps. putida* PPG1 using the K factor recombinant plasmid with chromosome transfer ability XYL-K (Novick, *et al.*, 1976). Genetic analysis has also been carried out by conjugation in *Ps. fluorescens* and *Ps. glycinea*, by transduction in *Ps. acidovorans* and *Ps. syringae* pathovar *syringae* and by transposon mutagenesis in *Ps. solanacearum* and *Ps. syringae* (Holloway and Morgan, 1986)

Physical mapping of the *Pseudomonas aeruginosa* PAO genome has been undertaken using rare cutting restriction enzymes notably *Spe*I followed by fragment analysis with 2-D pulsed field gel electrophoresis (Tummler *et al.*, 1992). In these maps the origin of replication is used as the point of reference which is more desirable than the arbitrary origins used in chromosome maps generated by the traditional conjugation, transduction and transformation gene exchange experiments (Tummler *et al.*, 1992).

1.2.2 Genetic Organisation

Genetic studies have revealed a biased location of many biosynthetic genes and central metabolic pathway genes on only one half of the map (reviewed Krawiec and Riley, 1990). Biosynthetic genes which are commonly contiguous in *Enterobacteriaceae* are often noncontiguous in pseudomonads however, functionally related genes are clustered in noncontiguous groups termed "supraoperonic clustering" (Holloway and Morgan, 1986). Some exceptions have however been identified in *Ps. aeruginosa*; the flagellar genes, genes of the inducible arginine pathway (Holloway and Morgan, 1986), the amidase genes (Drew and Wilson, 1992) and possibly the alginate biosynthetic genes cluster (Chitnis and Ohman, 1993) are all organised into operons.

1.2.3 Transposable Genetic Elements

Mobile genetic elements provide the potential for 'add on' and the rearrangement of genetic material which are thought to be important events in bacteria evolution (Reznikoff, 1983). Transposons and insertion sequences have been reported in all *Pseudomonas* species (Holloway and Morgan, 1986): a 7.3 kb element, IS22, with a different restriction map from other reported insertion sequences has been identified in *Ps. aeruginosa.* A 6.8kb transposon, Tn2521, which codes for carbenicillin, streptomycin and sulphanilamide resistance has been identified in clinical isolates of *Ps. aeruginosa.* The toluene degradative pathway operons on the TOL plasmid pWWO of *Ps. putida* are contained in two transposons, a 70 kb element Tn4653 and the smaller 56 kb Tn4651 which can transpose into the host chromosome or into other plasmids. Thus horizontal evolution can occur by transfer of toluene degrading ability to new hosts (Tsuda and Lino, 1987; Haas 1983; Rothmel *et al.*, 1991).

Horizontal and vertical transfer of antibiotic resistance genes has been shown to be mostly due to R plasmids and the transposons they carry. Tn21-like transposons and some plasmids have been shown to have potential site specific mobile elements (integrons) made up of two conserved elements at the 5' and 3' ends between which discrete cassettes encoding multiple drug resistance are located (Stokes and Hall, 1989). The 5'-conserved segment encodes a putative site specific recombinase, Int, and a common promoter region, P_1 P_2, for expression of the integrated cassettes (Martinez and Cruz, 1990; Mercier *et al.*, 1990). The 3'-conserved segment carries a gene for sulphonamide resistance,*sulI*, and two open reading frames, ORF_4 and ORF_5 (Stokes and Hall, 1989). Many clinical isolates of *Ps. aeruginosa* and *Enterobacteriaceae* showing multiple resistance are known to harbor integrons on the R plasmids. The *Ps. aeruginosa* nonconjugative plasmid pVS1 carries an integron, InO, that appears to be an ancestor of the more complex integrons (Bissonnette

and Roy, 1992). The codon usage for the InO *int* and *sul1* genes resembles that of other integrons however, the integration site is unoccupied.

Many *Ps. aeruginosa* strains are lysogenic for one or more bacteriophages, however lysogeny is relatively rare in other *Pseudomonas* species (Holloway and Krishnapillai, 1975). Mutator phages that lysogenize by transposition are rare in *E. coli* the only examples being Mu and the closely related D108, however, many *Pseudomonas* phages which use the homology-independent recombination process of transposition for their replication have been identified (Rehmant and Shapiro, 1983). Bacteriophage D3112 has been used to develop *in vivo* cloning systems analogous to Mu (Rothmel *et al.*, 1991).

1.2.4 Plasmids

Plasmids are widely spread in all strains of *Pseudomonas* examined. In *Ps. aeruginosa* the narrow-host-range sex factors FP2 and FP5 have been identified. Large multidrug resistance broad host range plasmids such as the IncP-1 plasmid R68 have been found in many *Pseudomonas* isolates (Haas 1983). A large number of 'degradative' plasmids encoding catabolic pathways for a wide range of substrates have been isolated from soil bacteria mostly *Ps. putida* (Haas, 1983; Rothmel *et al.*, 1991). These include the Cma (chromosome mobilising ability) recombinant plasmid XYL-K transferrable by sex factor K which codes for the meta- cleavage pathway of xylene and toluene (Mylroie *et al.*, 1977), the broad host range TOL plasmid pWWO coding for meta-cleavage pathway of toluene and xylene (Franklin, *et al.*, 1981; Harayama, *et al.*, 1984) plasmids NAH7 and pWW60-1 which code for the meta- and ortho- cleavage pathways for naphthalene, respectively (Yen and Gunsalus, 1982; Cane and Williams, 1986). Complete degradation of recalcitrant compounds often involves pathways partly encoded by plasmids and partly by the chromosome (Haas, 1983). Plasmid CAM codes for the degradation of camphor to

isobutyrate which is further degraded to TCA cycle intermediates by chromosomally encoded enzymes. The distribution of plasmids in *Pseudomonas* possibly reflects the habitat of particular species with opportunistic pathogens carrying R (drug resistance) factors and the soil organism *Ps. putida* being endowed with degradative plasmids (Haas, 1983).

1.2.5 Nucleotide Composition and Codon Usage in *Pseudomonas*

Codon usage in genes of a particular organism is biased, and the choice between synonymous codons tends to be characteristic of the species (Grantham *et al.*, 1980 a; Grantham *et al.*, 1980 b). From a practical point of view, codon analysis is useful for identifying protein coding regions in a particular nucleotide sequence (Gribstov *et al.*, 1984;), design of oligonucleotide probes from amino acid sequences and may be useful in giving clues to resolving ambiguities in sequences caused by band compressions which are all too common in high G+C content DNA fragments (West and Iglewski, 1988).

The *Pseudomonas* have a high G+C content in their genomes (Rothmel, *et al.*, 1991): *Ps. aeruginosa* has 67%, *Ps. putida* 61-62% and *Ps. fluorescens* 59-63%. An analysis of the G+C content of *Ps. aeruginosa* coding sequences gave an average of 64.8%, 43.5% and 88.7% at positions 1,2 and 3, respectively, in the coding triplets with the exception of plasmid, phage and transposon encoded genes and the pilin chromosomal genes (West and Iglewski, 1988). Other salient points to emerge from this study were that C is preferred to G at the wobble position with the exception of leucine and proline where CTG and CCG are more common. Valine codons CTG and GTC and serine TCG and TCC appear with equivalent frequency.

1.2.6 Gene Cloning in *Pseudomonas*

Gene cloning techniques and accompanying *in vitro* manipulations are now universal. The choice of host-vector systems may be critical for *in vivo* expression of cloned genes because of host range limitation of vector replicons. Differences in transcription/translation signal recognition between species can result in poor heterologous expression of genes as seen with *Pseudomonas* genes cloned in *E. coli* (Nakazawa and Inouye, 1986) and overexpression of toxic proteins may lead to loss of host cell viability. Several plasmid vectors that can be used in *Pseudomonas* have been developed from gram negative broad-host range plasmids (Table 1.1) (Morales et al., 1990). The most versatile cosmid and plasmid vectors have been developed from IncP-4/IncQ plasmids RSF1010, R300B and R1162 by insertion of new antibiotic resistance genes, multiple cloning sites (MCS) and sometimes an additional replication origin to make them shuttle vectors. IPTG controlled expression vectors of the pMMB range have been constructed by introducing the hybrid *tac* promoter and *lac* repressor upstream of the MCS (Morales *et al.*, 1990).

In vivo cloning systems that use the transposable element Mu have been developed, because of its high transposition frequency, and used in a wide range of gram negative bacteria (Groisman and Casabadan 1987). The system however does not work well with *Pseudomonas* because of low phage titres, necessitating the development of an analogous system based on phage D3112 which can also be used for mutagenesis (Rothmel *et al.*, 1991). Phage D3112 uses a headful packaging mechanism like Mu and most of the viral properties have been deleted from the vector leaving just the integrative properties which increases the size of host DNA that can be packaged. Several mini-D3112 elements have been made by addition of markers and replicons making recovery of cloned genes on a plasmid possible (Rothmel *et al.*, 1991).

Table 1.1. Properties of IncQ-derived broad host range vectors

Vector	Size (kilobase pairs)	Markers	Cloning site(s)	Insertional change
General-type cloning				
pKT210	11.8	Cm, Sm	*Sst*I	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Eco*RI	$Sm^r \rightarrow Sm^s$
			*Hin*dIII	None
pKT231	13.0	Sm, Km	*Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Eco*RI, *Hpa*I	$Sm^r \rightarrow Sm^s$
			*Hin*dIII, *Xho*I	$Km^r \rightarrow Km^s$
			*Cla*I, *Xma*I, *Pvu*I	$Km^r \rightarrow Km^s$
			*Bam*HI, *Bgl*II	None
pKT248	12.4	Cm, Sm	*Eco*RI	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
			*Sal*I	$Cm^r \rightarrow Cm^s$
pDSK509	9.1	Km, Sm	*Pst*I, *Sal*I, *Xba*I	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Bam*HI, *Kpn*I	$Sm^r \rightarrow Sm^s$
			*Sst*I, *Eco*RI	$Sm^r \rightarrow Sm^s$
Joint replicon				
pKT230	11.9	Km, Sm	*Bam*HI	None
		mob$^+$	*Eco*RI, *Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
		pACYC177	*Hin*dIII, *Sma*I, *Xho*I	$Km^r \rightarrow Km^s$
pFG7	13.3	Ap, Sm	*Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
		Su, Tc	*Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Bam*HI, *Cla*I	$Tc^r \rightarrow Tc^s$
		pBR322	*Hin*dIII, *Sal*I	$Tc^r \rightarrow Tc^s$
pUI81	13.3	Sm, Su, Tc	*Sst*I, *Sst*II	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Bam*HI, *Cla*I	$Tc^r \rightarrow Tc^s$
		pBR322	*Hin*dIII, *Sal*I	$Tc^r \rightarrow Tc^s$
pMW79	13.3	Ap, Sm, Tc	*Sst*I	$Sm^r \rightarrow Sm^s$
		mob$^+$	*Sst*II	$Sm^r \rightarrow Sm^s$
		pBR322	*Bam*HI, *Cla*I	$Tc^r \rightarrow Tc^s$
			*Hin*dIII, *Sal*I	$Tc^r \rightarrow Tc^s$
pSUP104	9.5	Cm, Tc	*Eco*RI	$Cm^r \rightarrow Cm^s$
		mob$^+$	*Bam*HI, *Hin*dIII, *Sal*I	$Tc^r \rightarrow Tc^s$
		pACYC184	*Pst*I	None
Cosmid				
pMMB33, pMMB34	13.8	Km, *cos*	*Bam*HI, *Eco*RI	None
		mob$^+$	*Hpa*I, *Sst*I, *Sst*II	None
pFG6	15.3	Ap, Tc	*Bam*HI, *Cla*I	$Tc^r \rightarrow Tc^s$
		mob$^+$, *cos*	*Hin*dIII, *Sal*I	$Tc^r \rightarrow Tc^s$
pSUP106	9.9	Cm, Tc	*Eco*RI	$Cm^r \rightarrow Cm^s$
		mob$^+$, *cos*	*Bam*HI, *Hin*dIII	$Tc^r \rightarrow Tc^s$
		pACYC184	*Sal*I	$Tc^r \rightarrow Cm^s$
pJRD215	10.2	Km, Sm	*Bam*HI, *Kpn*I, *Mlu*I, *Hpa*I, *Nde*I, *Stu*I, *Xba*I, *Spe*I, *Eco*RI	None

Continued

Table 1.1. - *continued.*

Vector	Size (kilobase pairs)	Markers	Cloning site(s)	Insertional change
Promoter probe				
pKT240	12.9	Km, Ap *mob*$^+$	*Eco*RI, *Hpa*I	Sm^s → Sm^r
pCF32	15.1	Km, Ap *mob*$^+$	*Hin*dIII	*xylE* → *xylE*$^+$
pUI523A, pUI533A	9.7	Tc, *mob*$^+$, *lacZ*	*Kpn*I, *Dra*I, *Stu*I, *Sma*I	Fusion to *lacZ*
Controlled expression				
pMMB66EH	8.8	Ap, P_{tac}	*Eco*RI, *Sma*I, *Bam*HI	None
		mob$^+$, *lacI*^{q}	*Sal*I, *Pst*I, *Hin*dIII	None
pMMB67EH	8.8	Ap, P_{tac}, *mob*$^+$	*Eco*RI, *Sst*I, *Kpn*I, *Sma*I, *Bam*HI, *Xba*I	None
		lacI^q	*Sal*I, *Pst*I, *Sph*I, *Hin*dIII	None
pMMB207	9.0	Cm, P_{tac}, *mob*$^+$	Same as pMMB67EH	None
pMMB206	9.3	Cm, *mob*$^+$, *lacI*^{q}, P_{tac}-P_{tac}, *lacZα*	Same as pMMB67EH	*lacZα*$^+$ → *lacZα*
pNM185	14.0	Km, *mob*$^+$, Pm, *xylS*	*Eco*RI, *Sst*I, *Sst*II	None
pERD21	13.8	Km, *mob*$^+$	*Eco*RI, *Sst*I, *Sst*II	None
		Pm, *xylS*	*Kpn*I, *Sal*I, *Hpa*I, *Bcl*I	None
pPLGN1		Km, P_L *mob*$^+$, c1857	*Eco*RI	None

Taken from Morales *et al*, 1990

1.3 Mechanism of Gene Transcription in Bacteria

1.3.1 The DNA-Dependent RNA Polymerase

RNA polymerase catalyses the initiation, elongation, and termination of transcription using double stranded DNA as the template in addition to RNA primer synthesis in the replication of DNA. The synthetic reaction requires the presence of a divalent metal cation in addition to all four ribonucleoside triphosphates (rNTPs) and template DNA. The RNA chain is synthesized in the 5' to 3' direction complementary to the 3' to 5' coding strand of DNA. The *E.coli* enzyme is the best studied although RNA polymerases from several other bacterial species including *Pseudomonas* have been characterised (Burgess, 1976). All contain subunits homologous to the *E.coli* enzyme and heterologous subunits can be used to reconstitute active enzyme. Many bacterial RNA polymerases show immunological cross-reactivity of subunits indicating a high degree of conservation during evolution (Chamberlin, 1982).

The *E. coli* enzyme contains a catalytic core of two α-subunits (Mr. 36592) encoded by *rpoA*, one β-subunit (Mr. 150619) encoded by *rpoB* and one β'-subunit (Mr. 155162) encoded by *rpoC* (Chamberlin, 1982). In addition several sigma factors, the most abundant being the housekeeping σ^{70} (*rpoD*), confer promoter specificity to the holoenzyme and dissociate after transcription initiation (Gross *et al*, 1992). The β subunit is responsible for binding rNTPs and carries the catalytic site, the positively charged β' binds non-specifically to DNA and the α-subunits are required for the stability of the complex (Das, 1993).

Several other proteins are tightly bound to the RNA polymerase complex including the NusA protein which binds after transition of the initiation complex to the elongation complex and the ω subunit of Mr 8,000-10,000 which can be removed without any apparent effect on the catalytic mechanism of the enzyme.

All multi-subunit RNA polymerases examined to date have tightly bound zinc ions (Sentenac *et al.*, 1992). The *E. coli* enzyme has two Zn^{2+} ions one being on the β' subunit and the other possibly on β or interacting with these two subunits (Wu *et al.*, 1977). The purified core enzyme tends to aggregate to dimers and higher order multimers greatly favoured by low salt concentration, however the holoenzyme shows no tendency to aggregate (Chamberlin, 1982).

The three dimensional structure of the RNA polymerase core enzyme has been determined by electron crystallography (Darst *et al.*, 1989). It contains a structure similar to the DNA binding cleft of Klenow DNA polymerase (Ollis *et al.*, 1985) and in addition, a flexible domain which might surround the DNA and allow processivity in the polymerization of ribonucleotide triphosphates. The three dimensional structure is likely to hold true for all multi-subunit RNA polymerases because of the extensive homology seen between the β- (with ten conserved domains) and β'- (with eight conserved domains) subunits with the largest subunits of the yeast, Drosophila and murine RNA polymerases (Sentenac *et al.*, 1992).

1.3.2 Initiation of Transcription

Initiation of transcription depends at a minimum on i) the binding of RNA polymerase to

the promoter, a relatively weak and rapidly reversible step usually represented by the equilibrium constant $_B$ and ii) a slow isomerization step during which the DNA double helix is unwound near the site of initiation of transcription. This step is usually represented by the rate constant k_2 (Crothers and Steitz, 1992). RNA polymerases are unique among nucleic acid synthesizing enzymes in that they initiate RNA chains *de novo* by coupling two specific nucleoside triphosphate substrates to form the primer for subsequent chain extension (Chamberlin, 1982). The holoenzyme locates a promoter site specified by the sigma factor, the canonical hexanucleotides TTGACA-----TATAAT (located at -35 and -10 to the transcription start site respectively) for the *E. coli* σ^{70} holoenzyme. This leads to formation of a metastable closed promoter complex which in σ^{70} holoenzyme is followed by factor independent melting of the DNA to a single stranded bubble encompassing the -12 to +4 region (isomerization of the closed complex to the open promoter complex) that is resistant to dissociation by high salt or heparin (Das 1993). Depending on the promoter, rapid cycles of abortive initiation occur for each RNA that is initiated, releasing RNA transcripts of up to nine nucleotides without concomitant dissociation of the RNA polymerase from the promoter (Carpouis and Gralla, 1980; Munson and Reznikoff, 1981). Initiation involves not just the formation of the first bond but extension of the first 8-10 nucleotides. Transition into the stable elongation complex occurs at this stage and this transition is profoundly affected by the initially transcribed region (Brunner and Bujard, 1987; Kammerer *et al.*, 1986). Experimental evidence for these initiation phases has been obtained by Krummel and Chamberlin (1989). The initial transcription complex (ITC) protects the same upstream edge from DNase I digestion as the closed and open promoter complexes but the downstream edge of RNAP is increased. A concomitant DNA bending occurring as a result of changes in the enzyme conformation during isomerization is indicated by the appearance of a new DNase I hypersensitive site located upstream of the melted bubble.

The first RNA polymerase translocation is accompanied by loss of the sigma factor marking the transition to the initial elongation complex (IEC). The biochemical steps that constitute initiation have been summarised as follows (Krummel and Chamberlin, 1989):

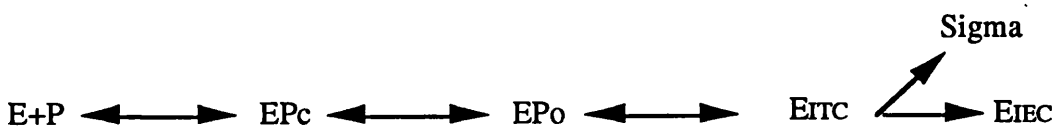


Fig 1.1a Steps in transcription initiation. E is the enzyme, P the promoter, EP_C and EP_O refer to closed and open promoter complexes, E_{ITC} refers to initial transition complex, E_{IEC} refers to initial elongation complex

1.3.3 Transcript Elongation

Two models for transcription elongation have been proposed. In one model, the ternary complex of RNA polymerase, DNA template and the growing RNA chain with an 18 (+/-1) bp transcription bubble, 12 bp RNA/DNA hybrid covering about 30 bp of DNA elongates in a continuous mode. This complex translocates 1 bp at time as 1 bp of DNA is opened at the leading edge of the transcription bubble and closed at the lagging edge per ribonucleotide added to the chain (Gamper and Hearst, 1982; von Hippel *et al.*, 1984; Yager and von Hippel, 1987). Based on *in vitro* analysis of experimentally halted ternary transcription complexes, a second discontinuous 'inchworm' model has been proposed whereby the leading edge translocates about 10 bp at the end of each elongation cycle (Fig 1.2) (Krummel and Chamberlin, 1992a and b; reviewed Das 1993). In this model the elongation cycle consists of synthesis of RNA at the catalytic site into a low and a high affinity RNA binding site, the latter being capable of holding up to a 9 nucleotide long transcript, and the translocation step involves

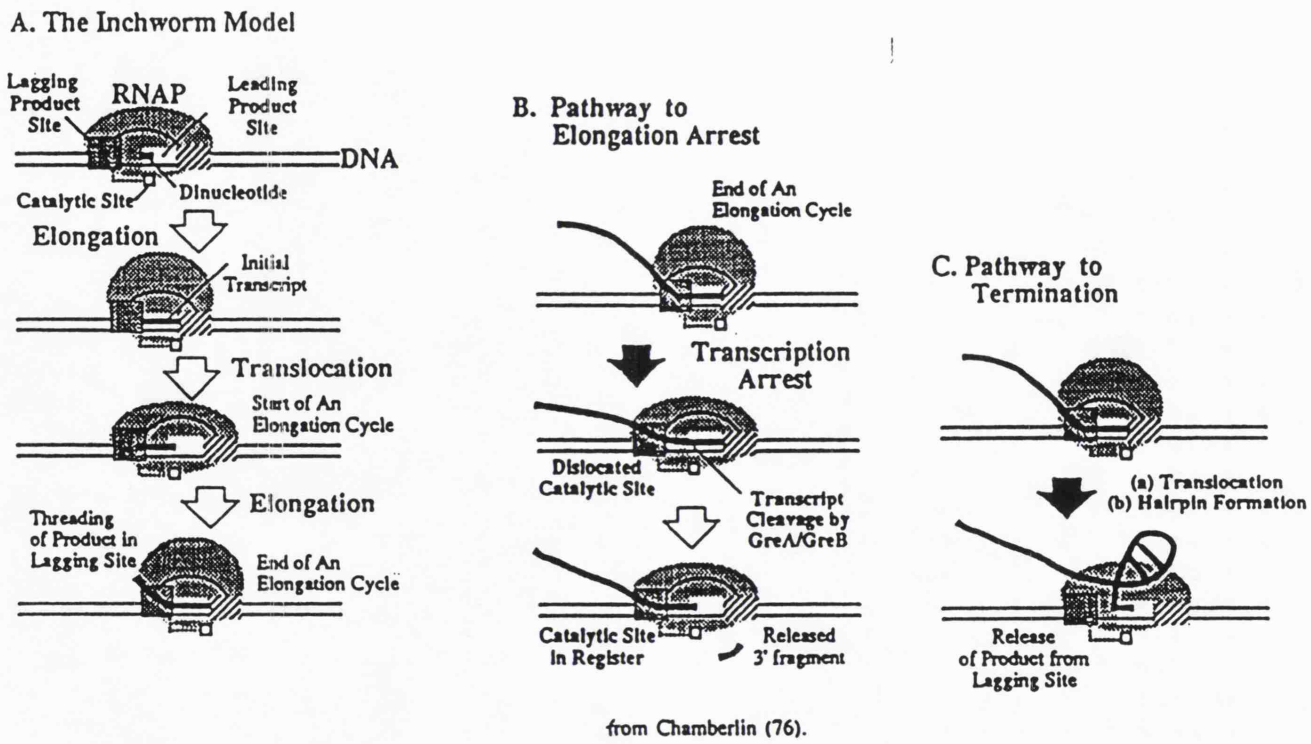


Fig 1.2 The inchworm model of transcript elongation. Adapted from Chamberlin (1976)

emptying of this site and rearrangement of the catalytic site in order to initiate another elongation cycle. This model is supported by the fact that the k_s values for addition of nucleotides during transcription and the stability of ternary complexes differ markedly between different sequences (Surrat *et al.*, 1991; Arndt and Chamberlin, 1990). DNA footprinting of transcription complexes halted artificially by limitation of one nucleotide and then restarted (transcription walking) resulted in an unchanged downstream boundary as 7 to 8 nucleotides are added while the upstream boundary moves with each successive addition to the growing chain (Krummel and Chamberlin, 1992 b). In this study, some complexes previously classified as "dead end" complexes could be switched to active elongation complexes by the use of ITP rather than GTP. DNâse footprinting of these complexes demonstrated that the downstream edge of the transcription complex only translocated after 9 nucleotides had been added to the growing RNA transcript.

Direct evidence for active centre rearrangements and the presence of low and high affinity transcript sites has been obtained from RNA polymerase initiation complexes (Mustaev *et al.*, 1993). Extension from ATP covalently linked to the evolutionary conserved His^{1237} within the active site on the β-subunit yields promoter specific families of transcripts of 7-8 nucleotides. GTP covalently linked to Lys^{1063} extends only by 2-3 bases whereas it grows to 8-9 bases when attached to His^{1237}. This led to the proposal of the active centre extending to accommodate a nanomer of RNA during the extension cycle, the filled site is then emptied by threading the RNA 5' terminus passed the His^{1237}.

1.3.4 Transcription Termination

The overall transcription cycle consists of initiation, elongation and termination. Chain elongation is not continuous and the RNA polymerase often encounters sequences that

temporarily halt transcription. Sometimes this is a decision point as to whether to continue elongation or to terminate RNA synthesis accompanied by transcript release and dissociation of the RNA polymerase from the DNA template. Pausing is in part regulated by NusA which is known to enter the transcription complex after dissociation of sigma (Friedman *et al.*, 1987) and is essential for cell growth (Nakamura and Uchida, 1983). Two types of terminator have been described in *E. coli*, simple and rho-dependent terminators (Friedman *et al.*, 1987; Yager and von Hippel, 1987; Das, 1993).

1.3.4.1 Simple Termination

Simple terminators have a canonical structure of a G+C rich axis of dyad symmetry followed by a run of Us in the RNA and terminate transcription with spontaneous release of the transcript without any apparent involvement of a protein factor (Yager and von Hippel, 1987). The termination event involves transcription of the G+C split dyad symmetry which is postulated to form a hairpin structure within the transcription bubble causing RNA polymerase to pause (Johnston and Roth, 1981; Lynn *et al.*, 1985; Yanofsky, 1981). Dissociation of the paused elongation complex is facilitated by the weak nature of the oligo U:oligo dA heteroduplex (Farnham and Platt, 1982). This however is not the full story. Sequences downstream of the termination site up to the 3' RNA polymerase footprint protection site significantly affect transcription termination efficiency (Reynolds and Chamberlin, 1992; Lee *et al.*, 1990) as does the DNA conformation (Reynolds *et al.*,1992). The simple termination event is itself complex involving many interactions, and some experimental data imply involvement of the hairpin loop sequence and sequences upstream of the terminator stem, that are not fully characterized (Yager and von Hippel, 1987). Template release at rho-independent terminators is then enhanced by entry of the sigma factor (Arndt and Chamberlain, 1988).

1.3.4.2 Rho-Mediated Transcription Termination

In contrast to the highly circumscribed sequences of simple terminators, rho-dependent terminators are highly heterogeneous and also occur within coding and regulatory regions of DNA. The mechanism involves rho-independent pausing of the RNA polymerase followed by rho-mediated release of transcript (Yager and von Hippel, 1987).

Rho functions as a hexamer as inferred from sedimentation coefficients in the presence of poly(C) and electron micrographs of uranyl acetate stained Rho reveal a ring structure composed of six globular units with a cleft on one side (Bear *et al.*, 1988; Oda and Takami, 1973 ; Richardson, 1990). Three structural domains of Rho have been revealed by limited proteolysis (Dolan *et al.*, 1990). The N-terminal domain which extends to a trypsin sensitive site, residue 130, contains a functional RNA binding site and shows strong homology to other RNA binding proteins (Query *et al.*, 1989). In addition residues 14 to 20 shows sequence similarity to the cytidine binding part of the regulatory subunit of aspartate transcarbamoylase. The second structural domain 167-359 has a functional ATP binding site and shows homology to nucleotide binding proteins (Dambroski and Platt, 1988). The carboxyl terminal domain does not appear to bind ATP or RNA but mutations within it impair Rho function.

Rho-dependent transcription termination is accompanied by ATP hydrolysis and the purified Rho protein is an RNA dependent nucleoside triphosphate phosphohydrolase (NTpase) (Lowery-Goldmer and Richardson, 1974; Howard and Crombrugghe, 1976). In the proposed mechanism , Rho interacts with unpaired cytidylate residues in the *rut* segments of the nascent transcript via clefts formed by each dimer in the presence of 3 ATP molecules per hexamer (Richardson, 1990; Yager and von Hippel 1987). This primary binding of Rho induces conformational changes coupled to ATP

hydrolysis that allows secondary RNA contacts which displace base pairing between various sections of the transcript and between the RNA/DNA hybrid without releasing the primary site. Contacts between Rho and the RNA polymerase have been observed however, the detailed interactions are unknown especially at the product site (Richardson, 1990). The sigma factor does not appear to play any role in this type of termination although Rho has been shown to bind to NusA-RNA polymerase complexes but not the core RNA polymerase (Darlin *et al.*, 1971).

Other factors have been shown to cause transcription termination at specific sites. Ribosomal protein L4 when produced in excess, causes termination within the 11-gene S10 operon at the potential rho-independent terminator located at the 3' end of the 140 base leader region by binding to a 9 base sequence (UAAUGGGUCC), which is identical to its binding site on 23S rRNA (Friedman *et al.*, 1987). Bacteriophages T3 and T7 early transcripts initiating at the A promoters are terminated *in vitro* at their early terminators with an efficiency of the order of 60-90% and downstream of the *in vivo* termination site (Dunn and Studier, 1980; Neff and Chamberlin, 1980; Kassavetis and Chamberlin, 1981). When the *E. coli* protein factor TAU is added, the termination efficiencies rise to nearly 100% and the termination sites are shifted to the *in vivo* locations (Yager and von Hippel; Briat and Chamberlin, 1984).

1.4 Control of Transcription in Bacteria cells

Bacteria have evolved a large number of regulatory mechanisms to ensure specific induction of catabolic operons in response to presence of growth substrates and to switch on biosynthetic genes only when there is a requirement for the end product. Mechanisms

for negatively and positively regulating gene expression at the transcriptional, translational and post translational levels have been seen in all organisms studied. All of the stages of transcription (initiation, elongation and termination) are potential sites for regulatory intervention.

1.4.1 Regulation of Transcription Initiation at σ^{70} Dependent Promoters

Transcription initiation regulatory factors function by repression, activation, and some factors activate some genes while repressing others.

1.4.1.1 Repression

Repressors of transcription initiation bind to specific sites (operators) near and often overlapping the promoter providing steric hindrance to the holoenzyme. In the *gal* and *ara* operons of *E. coli*, the repressors bind to two separate operators and cause DNA looping in the promoter region which blocks access of the RNA polymerase to the promoter. Binding to only one operator sequence provides partial control (Gralla, 1992). Other repressors have been shown to alter the initiation rate constants k_B and k_F in the following scheme (reviewed McClure, 1985) in which case the effect is mostly down modulation rather than complete shut off.

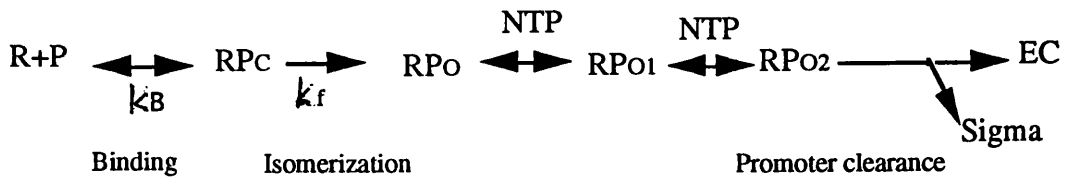


Fig 1.1b Reaction constants of transcription initiation

1.4.1.1.1 Lactose Operon

The lactose operon of *E. coli* remains the prototype of negative regulation of bacterial initiation of transcription (Miller, 1978; Gralla, 1992). The operon consists of four contiguous genes in two transcriptional units; *lacI* encodes the repressor protein, *lacZ* the structural gene for β-galactosidase, *lacY* the lactose permease, and *lacA* which codes for the thiogalactoside transacetylase. The *lac* repressor has two domains and functions as a tetramer although only one dimer binds the pseudosymmetric operator sequence (Gralla, 1992). The operator recognition specificity resides in the N-terminal domain. Mutations in this domain show reduced ability to bind to the operator without affecting tetramerization (Miller *et al.*, 1979; Kleina and Miller, 1990) and proteolytic fragments of this domain make identical contacts with the operator compared to the intact repressor (Gralla, 1992). However, the affinity is reduced perhaps due to loss of stabilization interactions (Gralla, 1990). The DNA binding domain has a helix-turn-helix (HTH) motif which is proposed to make contact with the DNA by penetrating the major groove of the β form double helix with minor contacts being made with the minor groove (Gralla, 1992). The amino acid sequence of the DNA binding domain shows strong homology to other repressors and only a few changes in the *lac* repressor make it capable of binding to the *gal* operator and *vice versa* (Haber and Adhya, 1988; Lehming *et al.*, 1990). Mutations within the carboxyl terminal domain lead to monomer dissociation and it is therefore thought to be important in subunit association (Gralla, 1992). Although the mechanism of inducer binding is not known, some important residues have been mapped soon after the N-terminal domain and they occur throughout the protein with a periodicity of 26 residues (Miller, 1978).

It was initially thought that repressor and RNA polymerase holoenzyme binding to their respective sites on the *lac* control region was mutually exclusive due to the overlap of

sequences. However, simultaneous binding of the RNA polymerase holoenzyme and the *lac* repressor has been demonstrated but the 3' contact of the polymerase on the template is changed; which seems to indicate that the repressor blocked isomerisation of the closed promoter complex to an open complex (Straney and Crothers, 1987; Gralla, 1990). Some experiments however, show that the RNA polymerase bound to the DNA together with the *lac* repressor synthesizes significant amounts of aborted transcripts but does not undergo the ITC to IEC transition (Krummel and Chamberlin, 1989). This situation has also been observed with the *gal* repressor (Dilauro *et al.*, 1979; Adhya and Miller, 1979).

Inducer (allolactose) binding to the LacI tetramer reduces its affinity for the operator significantly enough to cause dissociation under the low cellular concentrations of the repressor (10 molecules/cell).

In addition to operator O_1 (Fig 1.3), the *lac* operon has two other operators. O_3 is located 93 bp upstream of the promoter and is thought to have little function *in vivo*. O_2 is located 400 bp downstream within the *lacZ* gene and is of greater importance *in vivo*. The location of the three operators with respect to each other is in integral multiples of helical turns.

There is evidence that binding of a repressor tetramer to O_1 leads to DNA bending which brings it into contact with one or other of the other two operators through the free dimer by looping out the intervening sequence. This introduces cooperativity causing enhanced repression (Gralla, 1992).

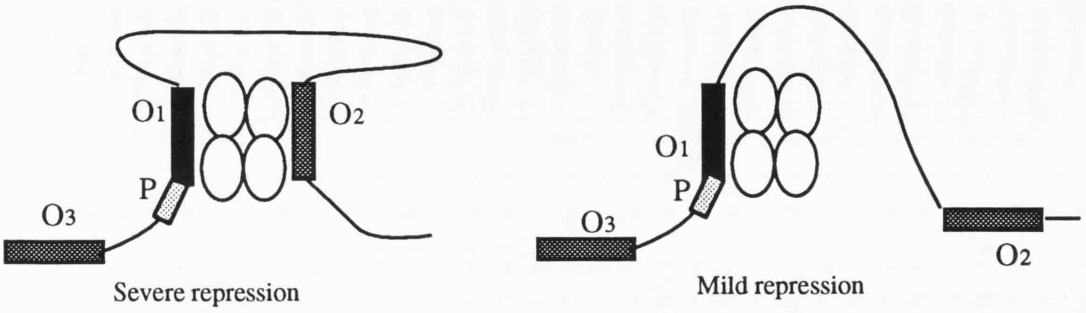


Fig 1.3 *lac* repressor binding to operator sequences

1.4.1.2 Activation of Transcription

The requirement for an additional protein to activate transcription implies that the either the RNA polymerase/promoter interaction is normally poor or that the closed complex to open complex isomerization reaction is difficult in the absence of the activator. Thus in the presence of activator/inducer k_B and/or k_F is enhanced and transcription initiation occurs. Regulator proteins which activate transcription fall into several families but the two main groups are factor-independent and factor-dependent.

1.4.1.2.1 *E. coli* Catabolite Activator Protein

The *E. coli* catabolite activator protein (CAP) is a factor-dependent global activator protein. When bound to its inducer, cAMP, CAP binds specifically to DNA sequences upstream of the promoter (at a variety of distances) and stimulates transcription of several sugar utilizing genes possibly by different mechanisms for different promoters (Crothers and Steitz, 1992). The enzyme adenyl cyclase synthesizes cAMP from ATP in response to complex signals linked to the sugar phosphotransferase system.

The structure of CAP has been determined and shown to have two domains: the larger N-

terminal domain binds cAMP within the interior of a 'β–roll structure; the carboxyl terminal domain has a helix-turn-helix DNA binding motif. CAP functions as a dimer with two deeply buried cAMP molecules which are thought to contribute to the dimer stability and may affect the relative orientation of the subunits (Crothers and Steitz, 1992). Mutations that enable CAP to stimulate transcription in the absence of cAMP have been found to fall within the cAMP binding domain and the D helix of the DNA binding domain (Aiba *et al.*, 1985).

Binding of the cAMP-CAP complex to the consensus pseudosymmetric DNA binding sequences is a two step process involving initial capture followed by formation of a stable complex which is associated with DNA bending. The *lac* operon is in addition to repression , positively controlled by cAMP-CAP which binds 41 bp upstream of the transcription start site and stimulates the closed complex formation (Crothers and Steitz, 1992). CAP stimulates transcription from *lac* promoter by increasing k_B (the equilibrium binding constant) 20-fold at the *lac* promoter (Reznikoff, 1992).

1.4.1.2.2 The Arabinose Operon

AraC is a bifunctional (activator/repressor) of the catabolic arabinose regulon of *E. coli.* AraC activates the *araBAD* operon, *araE* for low affinity sugar uptake, *araFGH* the high affinity arabinose uptake operon and autorepresses its own transcription (Schleif, 1993). The protein functions as a dimer made up of two identical subunits of molecular weight 30,500. Purification of AraC has been achieved with considerable difficulty because the protein does not bind to conventional columns and gives poor yields from precipitates (Schleif, 1992). The dimer contacts the major groove regions of the DNA substrate at a direct repeat binding site through a HTH motif within the carboxyl terminal domain which

is highly conserved amongst the AraC family of DNA binding proteins (Schleif, 1992).

Three AraC binding sites are located upstream of the *araBAD* promoter. In the absence of arabinose, the AraC dimer binds to the upstream $araO_2$ and a half site on the *araI*1 site by DNA looping possibly caused by cooperative dimerization of the monomers bound

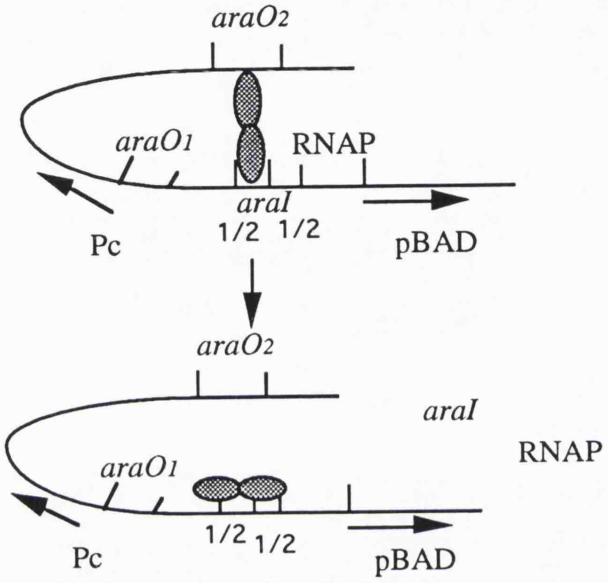


Fig. 1.4 AraC dimer activation of pBAD operon. The pc promoter controls expression of *araC* and is divergently transcribed relative to pBAD. *araI* site is divided into two half sites.

separately to the two sites (Shleif, 1992). Transcription activation is achieved by breakage of the loop when arabinose binds to the AraC dimer inducing allosteric changes that lead to the dimer occupying both half sites at *araI* and releasing *ara*O2 (Fig 1.4). Precisely how the occupancy of both half sites of *araI* leads to transcription activation is not known however, there is evidence that DNA bending upstream of the bound RNA polymerase may be an important factor in transcription (Schleif, 1992).

1.4.2 Transcription Activation by Alternate Sigma Factor

Bacteria contain many sigma factors that fall into two major families. The function of the sigma factor is to confer promoter specificity to the core RNA polymerase (Gross and Lonetto, 1992). Changes in promoter specificity thus provides the potential for discriminate transcription of genes and provides additional opportunities for regulation. The major *E. coli* sigma factor, σ^{70}, and related sigma factors like the heat shock response, σ^{32}, are competent to form 'open complexes' in the absence of auxiliary factors (Gross *et al.*, 1992).

σ^{54} (*rpoN*) is, however, structurally distinct from the σ^{70} family and allows the RNA polymerase to bind to promoters that contain conserved nucleotides at -24 and -12 from the transcription start point (CTGG_PyPuPyPu_ _ _ _ TTGCA). All σ^{54} related proteins have 2 strongly conserved regions (Gross and Lonetto, 1992); region I, located at the extreme N-terminus, contains a motif that resembles a glutamine rich leucine zipper, region III, located at the extreme C-terminus of the protein, contains a weak match to the leucine zipper and two potential HTH motifs. Promoter binding by the σ^{54} holoenzyme leads to formation of a heparin sensitive closed complex both *in vivo* and *in vitro* but does not have strand melting activity in the absence of an activator protein (Popham *et al.*, 1989; Wong *et al.*, 1987; Sasse-Dwight and Gralla, 1988). σ^{54} holoenzyme enzyme is the target for various enhancer binding proteins the best studied being NtrC (NRII) (Kustu *et al.*, 1989; Thony and Hennecke, 1989), NifA (nitrogen fixation protein A) (Dixon 1986) and XylR for xylene catabolism in *Ps. putida* (Gomada *et al.*, 1992). All these bind to upstream enhancer like elements and facilitate the closed to open complex change by RNAPσ^{54}

bound at the promoter.

1.4.2.1 Sensor-kinase/Response regulators

Many RpoN-dependent genes are coupled to two component signal transduction systems that respond to enviromental stimuli including cellular activities like chemotaxis (Lukat and Stock, 1993). The stimulus-response coupling often involves two families of proteins (Fig 1.5) (Stock *et al.*, 1990). One component of the regulatory pair is a sensor kinase that uses ATP and possibly other low molecular weight phosphodonors such as acetylphosphate, carbamoylphosphate and phosphoramidate (Lukat *et al.*, 1992) to autophosphorylate itself at a histidine located within a highly conserved domain within the family. The phosphate from the histidine protein kinase (HPK) is transferred to an aspartate residue within the conserved domain of the second component, the response regulator (RR) whose level of phosphorylation is further controlled by associated phosphatase activity. The response

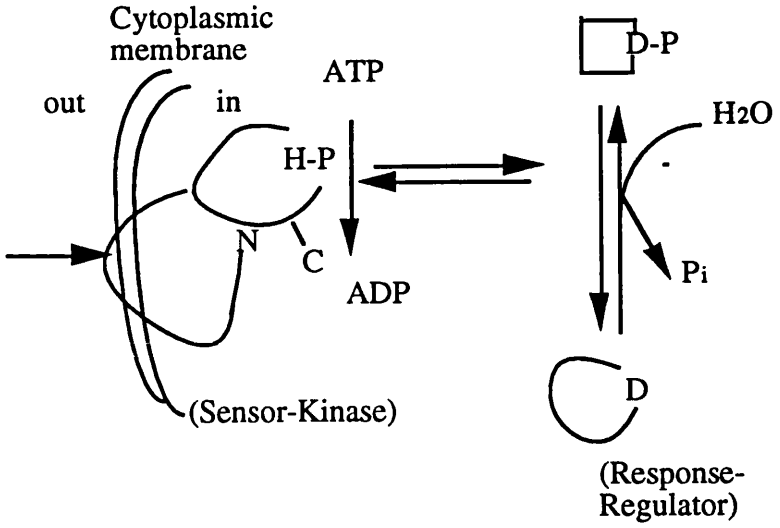


Fig 1.5 Signal transduction in bacteria Sensor-kinase/Response regulators

regulator then activates expression of the target gene by binding to enhancer elements and looping the DNA to form contacts with the closed promoter complex.

1.4.2.1.1 Control of nitrogen metabolism in bacteria

The central enzyme in bacteria for N-metabolism is glutamine synthetase which under N limiting conditions, mediates entry of ammonia into anabolic pathways. The amount of enzyme present is controlled by transcriptional regulation and the activity of the enzyme is controlled allosterically by adenylation (Maganasik, 1993). Enteric bacteria utilize ammonia preferentially as a nitrogen source by repressing enzymes required for utilisation of other nitrogen containing compounds at the transcriptional level (Magasanik and Neidhardt, 1987). When the cells are growing in a high ammonia environment, the NADP-linked glutamate dehydrogenase catalyses reductive amination of α-ketoglutarate, and by action of glutamine synthetase some of the glutamate formed is converted to glutamine in a reaction coupled to hydrolysis of ATP. In ammonia limiting situations, the high K_M of glutamate dehydrogenase becomes unsuitable for synthesis of glutamate and the cells rely more heavily on glutamine synthetase to assimilate nitrogen coupled with glutamate synthase which transfers the guanidino group of glutamine to a-ketoglutarate to form two molecules of glutamate (Magasanik, 1993). In *Pseudomonas aeruginosa*, glutamine synthetase is regulated by repression/derepression and by adenylation/deadenylation perhaps in a mechanism similar to the enterics (Janssen *et al.*, 1980)

The *glnALG* operon codes for glutamine synthetase (*glnA*), the response regulator NtrC (*glnG*) and the sensor-kinase NtrB (*glnL*). Transcription of the *gln ALG* can be initiated from three promoters (Magasanik, 1993): from *glnA*P1 and *glnA*P2 located upstream of *glnA* and *glnL*P located between *glnA* and *glnLG*. *glnA*P1 and *glnL*P are consensus σ^{70} promoters and, NtrC dimers bind to two sequences overlapping *glnA*P1 and one site overlapping *glnL*P causing partial repression when cells are grown in high ammonia concentrations (Magasanik, 1993). *glnA*P2 was identified as a promoter

(CTGG_PyPuPyPu_ _ _ _ TTGCA) recognised by the alternative sigma factor σ^{54} and activation occurs by means of NtrC.

The NtrC protein and other σ^{54} promoter activators are three domain polypeptides with a highly homologous central domain of approximately 220 amino acid in length, containing an ATP binding site, which is thought to be the RNA polymerase contact domain, and a carboxyl terminal domain containing a HTH DNA binding motif (Popham *et al.*,1989). NtrC and some other members of the family have a homologous N-terminal domain (which is otherwise very variable) which contains the phosphate acceptor aspartate residue and is most likely involved in cooperativity in situations of multiple enhancer elements

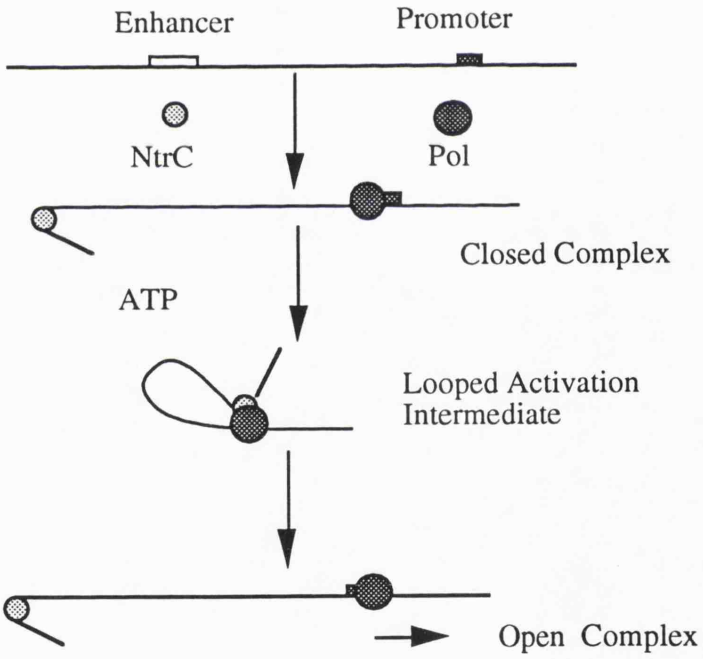


Fig 1.6 Stimulation of Open Complex formation at Sigma 54 Promoters

(Magasanik, 1993). NtrC binds to five sites in the *glnA* promoter region (Magasanik and Neidhardt, 1987), two of these sites, overlapping *glnA*P1, are high affinity and show full enhancer activity. Binding of phosphorylated NtrC to these sites causes looping out of the DNA to make protein-protein contact with the closed promoter complex at *glnA*P2 stimulating open complex formation in a reaction accompanied by hydrolysis of ATP (Fig 1.6) (Su *et al.*, 1990).

In some σ^{54} enhancer activated systems the DNA looping requires the integration host factor (IHF) binding of which causes DNA bending and the orientation of the enhancer element and the IHF binding site on the face of the helix becomes critical for activation (Weiss *et al.*, 1992).

1.4.2.1.2 Regulation of P_U promoter of TOL plasmid by XylR

The TOL plasmid pWWO of *Pseudomonas* strains (Worsey and Williams, 1975) encodes, on a 56 kb transposon the upper pathway operon *xylCMABN* for oxidative transformation of toluene/xylenes to the corresponding toluates/benzoates and the lower pathway, meta operon *xyl*DLEGF which cleaves the aromatic acids to TCA cycle intermediates (Tsuda and Lino, 1987; Harayama *et al.*, 1986; Nakazawa *et al.*,1990). Two regulatory genes *xylR* and *xylS* positively regulate both operons. In the presence of m-xylene, XylR activates transcription of both *xylCMABN* and *xylS* from consensus RpoN promoters and XylS then transactivates transcription of *xyl*DLEGF (Dixon, 1986; Inouye *et al.*, 1990).

A potential mechanism for XylR function has been proposed in which the protein binds to

palindromic sequences located 140 bp upstream of the transcription start sites, termed upstream regulator sequences (URS), and DNA looping occurs such that bound XylR contacts the closed promoter complex (de Lorenzo *et al.*, 1991). XylR binds the DNA through its carboxyl terminal domain in the presence of inducer to effect the activation. The DNA looping is facilitated by IHF whose recognition sequence 5'-ATCAANNNNTTR-3' (Friedman, 1988; Yang and Nash, 1989) has been identified at position -56 of *xylCMABN* such that XylR/URS is brought into contact with the closed promoter complex resulting in the isomerisation reaction and transcription initiation (de Lorenzo *et al.*, 1991). Deletion and insertion studies have shown that complete activation requires full turn or integral multiples of full turn separation between the URS and the promoter (Gomada *et al.*, 1992).

1.4.3 Regulation of transcription elongation

Early models of transcription elongation were based on a uniform RNA polymerase ternary complex which was unchanging during the elongation step (Yager and von Hippel, 1987). More recently, a discontinuous model which allows for interaction of the ternary complex with regulatory factors has been proposed (Krummel and Chamberlain, 1992; Das, 1993).

During the elongation step, the ternary complex is sometimes caught up in dead end complexes, and in eukaryotes continuation of elongation past the arrest sites requires the elongation factor TFIIS (Reines *et al.*, 1989; Sluder *et al.*, 1989; Kerpolla and Kane, 1990; SivaRaman *et al.*, 1990; Bengal *et al*; 1991). Recently discovered *E. coli* proteins GreA and GreB are functional homologs of TFIIS and a general biochemical pathway for their action has been proposed (Reines, 1992; Izban and Luse, 1992; Surrat *et al.*, 1991). GreB functions by causing transcript cleavage within the arrested transcription complex and allows the restart of elongation whereas GreA does not appear to function at paused

complexes but has to be present to prevent termination for continuation of elongation (Burokhov *et al.*,1993). These factors do not have any nuclease activity and are proposed to function by stimulating intrinsic RNA polymerase endonucleatic activity (Burokhov *et al.*, 1993).

1.4.4 Control of Transcription by Termination-Antitermination Systems

1.4.4.1 Regulation of Rho-Independent Terminators

1.4.4.1.1 Attenuation control in amino acid biosynthetic operons of *Enterobacteriaceae*

Amino acid biosynthetic operons in the enterics are commonly under attenuation control (Yanofsky, 1988). The common feature of the operons is the presence of a 150-300 base pair leader region between the promoter and the first structural gene of the operon which encodes a rho-independent terminator at its 3' end. Within the leader region are multiple codons for the specific amino acid followed by a stop codon. Delayed translation of these codons depending on the availability of the charged tRNA provides the signal to show amino acid limitation and causes operon expression.

The leader transcript can form three potential secondary structures that are important for regulation of these operons by transcription attenuation (Yanofsky, 1988). The first is the transcription pause signal that allows the ribosome to initiate translation and couples it to transcription (Landick *et al.*, 1985). The second and third secondary structures are mutually exclusive and function as transcription terminator and antiterminator, respectively.

In the *trp* operon of *E. coli* four sequences on the leader transcript can base pair as follows (Watson *et al.*, 1987): regions 1/2 and 3/4 in the absence of the ribosome, regions 2/3 the antiterminator with regions 1 and 4 unpaired in the antiterminator structure or region 3/4

with regions 1 and 2 being occupied by the ribosome in the termination mode. Under conditions of tryptophan limitation the ribosome stalls at the tryptophan codons thereby masking region 1 and the transcribing RNA polymerase exposes region three to base pair with region 2 before region 4 is transcribed. Under these conditions, transcription continues. Under high tryptophan conditions, the ribosome does not stall but proceeds to the stop codon masking regions 1 and 2 before region 3 is fully synthesized which results in 3/4 base pairing and termination of transcription.

Many amino acid biosynthetic operons are regulated by attenuation and in the *ilvGMEDA* operon for the biosynthesis of branched chain amino acids the leader region contains codons for all amino acids so that lack of any one of them leads to suppression of attenuation (Yanofsky, 1988). Charged tRNAs also appear to regulate expression of the transport systems for the branched chain amino acids where a rho-dependent terminator functions as the attenuator in an as yet uncharacterised mechanism (Landick, 1985).

1.4.4.1.2 Control of Attenuation by an RNA Binding Protein

Expression of the *trpEDCFBA* gene cluster of the gram positive bacteria *B. subtilis* and *B. pumilus* is controlled by transcription attenuation involving selection of alternate RNA secondary structures (Yanofsky, 1988). However, the signal this time is not charged tRNA but free tryptophan and the leader region does not encode an ORF unlike the enteric system (Kuroda *et al.*, 1986; Shimotsu *et al.*, 1986).

The selection of the mutually exclusive antiterminator/terminator secondary structures in the *B. subtilis trpEDCFBA* leader region is mediated by a *trans* acting factor, protein MtrB (Babitzke *et al.*, 1992). MtrB binding to the leader RNA was at first inferred *in vivo* by

showing that overexpression of the *trp* leader *in trans* resulted in constitutive expression of the operon and that this could be reversed by deletion of two homologous sequences (Kuroda *et al.*, 1986; Shimotsu *et al.*, 1986). *In vitro* studies have since shown that the *trp* attenuator stops transcription by Sp6 RNA polymerase and that tryptophan bound to MtrB causes gel retardation in non-denaturing polyacrylamide gels of the leader RNA species (Otridge and Gollnick, 1993).

A minimal *in vitro* system composed of template DNA, RNA polymerase, the four ribonucleoside triphosphates and purified MtrB was subsequently shown to attenuate transcription in the presence of tryptophan and two models have been proposed (Babitzke and Yanofsky, 1993). In model A, the RNA polymerase pauses after formation of the antiterminator structure and in the presence of MtrB, with bound tryptophan, this secondary structure is melted and transcription proceeds to form the transcription terminator. In model B, MtrB plus tryptophan binds successively to the 5' and 3' homologous sequences to stop formation of the antiterminator structure.

1.4.4.2 Regulation of catabolic operons by transcription antitermination

In addition to the widely studied sensor-kinase/response regulator systems in which the response regulator is controlled by phosphorylation, several systems have been found where transcription antiterminators are allosterically controlled by phosphate transfer from an environmentally responsive PEP-phosphotransferase sugar uptake system (PTS) (Amster-Choder and Wright, 1993). The best characterized are, the normally cryptic *bgl* operon of *E. coli* K12 (Mahadevan and Wright, 1987); the *sacPA* operon and *sacB* gene of *B. subtilis* (Debarbouille *et al.*, 1990; Crutz *et al.*, 1990)and the *arb* system of *Erwinia chrysanthemi* (El Hassouni *et al.*, 1992).

1.4.4.2.1 The *bgl* operon of *E. coli*

The *bgl* operon of *E. coli* is a catabolic system used for the uptake and utilization of aryl β-glucosides consisting of three genes *bglG, bglF,* and *bglB* (Fig 1.7) (Schnetz *et al.*, 1987; Mahadevan *et al.*, 1987). The first gene *bglG* is flanked by two rho-independent terminators that are known to be central to the regulation of the operon (Amster-Choder and Wright, 1993). The BglF protein phosphorylates β-glucosides during uptake and is homologous to the PTS family members EII^{GLC} and $EIII^{GLC}$ of *E. coli* which function in coupling transport of the substrate to phosphorylation (Bramley *et al.*, 1987; Bramley and Kornberg, 1987; Schnetz *et al.*, 1987). BglB was shown to be essential for hydrolysis of phosphorylated B-glucosides to yield glucose-6-phosphate.

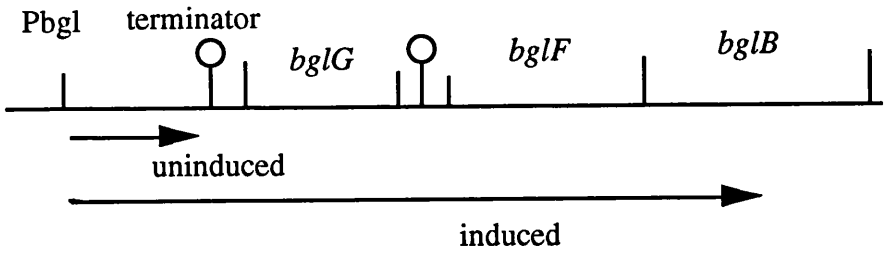


Fig 1.7 Schematic representation of the *bgl* operon

Positive regulation at the two terminators was initially shown by experiments where the promoter was changed to the *lac* promoter without any apparent change in the regulatory system and more directly by observation that mutations that destabilised the prospective upstream terminator or a terminator deletion led to constitutive expression of the operon without the requirement of the *bglG* product (Mahadevan and Wright, 1987). These experiments also showed that in the presence of functional BglF, induction required the presence of β-glucosides.

The function of BglG as a transcription antiterminator has subsequently been demonstrated and the binding site on the RNA transcript mapped (Houman *et al.*, 1990). It was initially demonstrated that a truncated leader transcript with an intact binding site could effect *in vivo* titration of BglG. In addition it was shown by using oligonucleotide competition, that the protein bound to an imperfect palindromic sequence of approximately 32 bases (Ribonucleic AntiTerminator sequence) which extends into the leading face of the terminator stem. A closely related BglG recognition site has also been found just upstream of the *bglG* downstream terminator. Evidence from studies using constructed mutations within the BglG binding region of the upstream terminator showed that BglG binding stabilized an alternative RNA secondary structure preventing the formation of the terminator loop (Amster-Choder and Wright, 1993). A model for *bgl* regulation has been proposed (Fig 1.8) in which the BglF, a membrane bound protein, is phosphorylated at a histidine

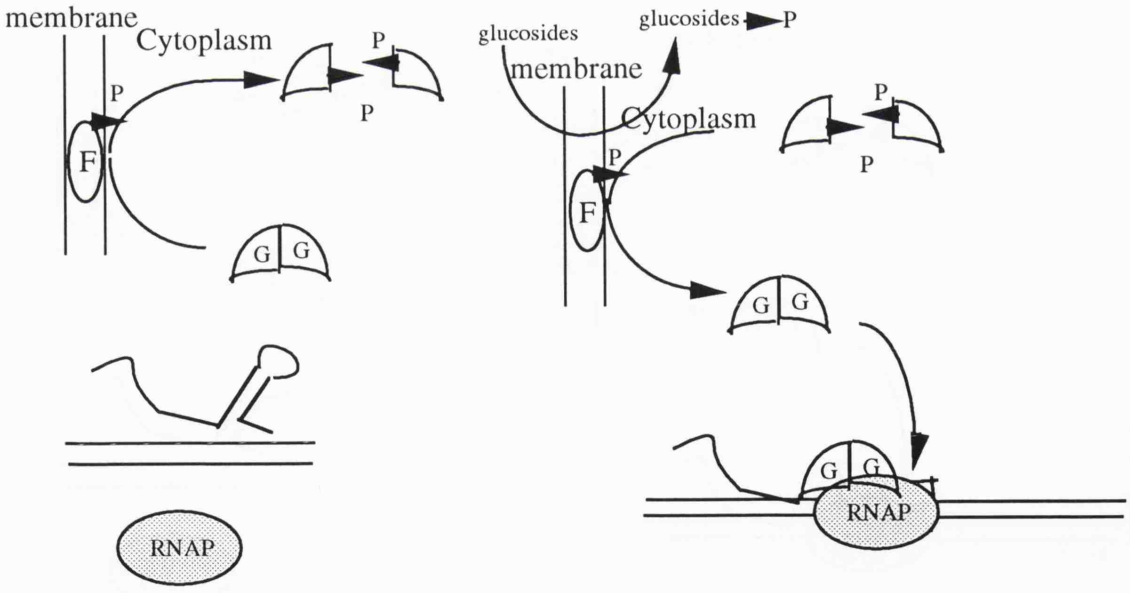


Fig 1.8 Proposed Mechanism for Transcription Antitermination by BglG

BglF (F) phosphorylates BglG (G) thus dissociating the dimer to the inactive form. RNAP represents the RNA polymerase.

residue by Hpr and acts as the substrate sensor by transferring this phosphate group either to the substrate, if present, or to BglG also at a histidine residue. In the absence of substrate, BglF phosphorylates BglG and this allosteric modification stops dimerization of BglG to the active form. In the presence of substrate, BglF dephosphorylates BglG-P by an as yet unknown mechanism, BglG dimerizes and binds to its recognition sequence and stabilizes the alternate RNA secondary structure. The binding of the BglG dimer and concomitant resumption of transcription elongation may also involve direct interaction with the transcription complex possibly via the Nus proteins (Amster-Choder and Wright, 1993).

1.4.4.2.2 *arb* genes of *Erwinia chrysanthemi*

The plant pathogenic bacterium *E. chrysanthemi* (Chartterjee and Vidaver, 1986) is capable of fermenting the natural β-glucosides arbutin and salicin through the *arb* pathway (El Hassouni *et al.*, 1990). The *arb* system consists of three genes *arbG*, *arbF*, and *arbB* transcribed in the same direction and all have been sequenced and analyzed (El Hassouni *et al.*, 1992). ArbF is highly homologous to the EII permeases in the PTS sugar transport system and has all of the functional motifs (Saier *et al.*, 1988). The ArbB amino acid sequence is highly homologous to BglB and other β-glycohydrolases (Henrisatt, 1991). The ArbG protein exhibits 60% identity with BglG and sequence alignment with other PTS associated antiterminators (Fig1.9) shows the presence of 58 conserved residues in all four proteins suggesting ArbG functions as an antitermination factor. Two Rho independent terminators have been identified flanking the *arbG* gene, STB1 upstream and STB2 downstream. Both terminators have partially overlapping sequences that are highly homologous to the BglG binding sequences (RAT), the STB2 sequence however, appears to have a duplication of the downstream half of the palindrome, boxB (El Hassouni *et*

al.,1992).

	BoxA		BoxB
STB1-*arb*	GGGTTGCTACTGCCAT	TG	GCAGGCAAAACCAGATGTTC
T1-bgl	GGATTGTTACTGCATT	C	GCAGGCAAAACCTGACATAA
BS*bgl-IR*	GGATTGTTACTGATAA	A	GCAGGCAAAACCTAAATTGC
BS*sacR/sacB*	GGTTTGTTACTGATAA	A	GCAGGCAAGACCTAAAATGT
BS*sac/R/sacP*	GGATTGTGACTGGTAA	A	GCAGGCAAGACCTAAAATTT
T2-*bgl*	GGATTGTTACCGCACT	AA	GCGGGCAAAACCTGAAAAAA
STB2-*arb*	GGATTGCGACTGTATA	TCCCTCA	GCGGGAAATACAGGCAAAAC

Fig 1.9 Sequence Alignment of RAT Sequences

1.4.4.2.3 Regulation of sucrose metabolism in *Bacillus subtilis*

Expression of the extracellular levansucrase enzyme, encoded by *sacB*, in *Bacillus subtilis* is regulated by the *sacXY* operon (Crutz *et al.*, 1990). The deduced amino acid sequence of SacY shows extensive homology to the transcription antiterminators BglG and SacT (Zukowski *et al.*, 1990). Genetic evidence suggests that SacX regulates SacY by a cascade involving the PTS system (Crutz *et al.*, 1990). SacY functions as a transcription antiterminator by allowing readthrough of a transcription terminator located between *sacB* and its promoter (Aymerich and Steinmetz; 1987; Shimotsu and Henner, 1986). SacX is a functional homologue of BglF and regulates SacY presumably by phosphorylation and dephosphorylation (Amster-Choder and Wright, 1993)

The *sacPA* operon codes for a membrane associated sucrose specific component of the PTS transport (*sacP*) and *sacA* the endocellular sucrase of *B. subtilis* (Aymerich and Steinmetz,

1992). Expression of *sacB* and *sacPA* is induced by sucrose and the two regulatory pathways for these saccharophytic enzymes are partially interchangeable (Steinmetz *et al.*, 1985).

SacT allows readthrough of a Rho-independent terminator located upstream of the *sacPA* operon possibly by binding to a 37 nucleotide imperfect palindrome that is highly homologous to the BglG binding site (RAT) and overlaps the terminator in a similar manner (Debarbouille *et al.*, 1990).

BglG, SacY, SacT and ArbG are all homologous and the conserved residues include the histidine and aspartate pair shown to be functionally important in BglG and SacT (Amster-Choder and Wright, 1993; Crutz *et al.*, 1990; Debarbouille *et al.*, 1990). The proposed (or for BglG proven) binding sites for these proteins, termed ribonucleic antiterminator (RAT) sequences are highly homologous (Aymerich and Steinmetz, 1992; reviewed in Amster-Choder and Wright, 1993).

A combination of mutation analysis and compromise folding of the RAT sequences for *sacB*, *sacPA* and *bglG* has yielded potential alternate secondary structure identical to one obtained independently for *bglG* (Houman *et al.*,1990; Aymerich and Steinmetz, 1992). The differences all lie in the bulges or loops in the secondary structures except for $bglT_2$ where there is a compensatory difference within the stem. Mutations that altered the proposed secondary structures reduced the inducibility of *sacB* and could be restored by compensatory mutations. In addition mutations that improved homology with the *bglG* RAT led to induction of *sacB* by BglG. These findings support a common regulatory mechanism and are corroborated by the findings that all these antiterminators except ArbG can substitute for BglG in *E. coli* but are devoid of control by BglF (Amster-Choder and

Wright, 1993).

1.4.4.3 Suppression of Rho-dependent transcription termination

In addition to termination/antitermination regulation of simple terminators described above, other systems have been described whereby the RNA polymerase is altered to override multiple rho-dependent and rho-independent terminators. The mechanism is best exemplified by the bacteriophage λ N and Q proteins which seem to modify the transcription complex at specific sites, *nut* on the transcript for N and *qut* on the template for Q, and become a part of the elongation complex allowing readthrough of multiple downstream terminators (Das, 1992; Friedman *et al.*, 1987; Yager and von Hippel, 1987; Roberts, 1993).

Upon infection of *E. coli* , λ phage DNA circularises by annealing of the 12 nucleotide overhangs followed by sealing of the nicks by the host DNA ligase. Transcription of the immediate early genes starts at the pL and pR promoters yielding the *N* gene transcript towards the left and *cro* gene transcript towards the right. The N protein encodes a transcription antiterminator that allows readthrough past the tLI terminator and the Cro tRI terminator and into the delayed early genes which direct phage replication and the Q gene whose product acts as an antiterminator for expression of the late genes encoding lysis functions and virion proteins. The decision for lytic or lysogenic development depends on the level of the two early genes *cro* and *cII*. CII activates transcription of the CI repressor favouring lysogenic development and Cro favours lytic development by preventing repressor expression from the pRM promoter.

1.4.4.3.1 Transcription antitermination by λ N gene product

During its role as a transcription antiterminator, the N protein forms a stable complex with the RNA polymerase at the *cis*-acting *nut* site positioned upstream of the two terminators tL1 and tR1 (Barik *et al.*, 1987; Horwitz *et al.*, 1987; reviewed in Das, 1992). The *nutL* and *nutR* sites have two components: BoxA defined as CGCTCTT (Olson *et al.*, 1982) followed by a 9 to 10 bp lambdoid phage evolutionary conserved region and known to be important for transcription antitermination by N (Morgan, 1986). This region is followed by Box B which is a 15 bp sequence with a hyphenated dyad symmetry proposed to form a hairpin loop in the transcript (Das, 1992).

Transcription antitermination by the N protein is known to require host factors NusA, NusB, S10 and NusG (Das 1993). The assembly of the transcription antitermination complex *in vitro* has been shown to be highly cooperative (Greenblatt, 1992). Studies by RNase footprinting and methylation experiments have shown that Box B is important for N binding to the RNA while Box A is important for the association of NusB and NusG with the elongation complex (Nodwell and Greenblatt, 1991; Friedman *et al.*, 1990). NusA has been shown to be sufficient for transcription antitermination by N at nearby intrinsic terminators, however, long distance antitermination requires the other host factors possibly to stabilize the complex (Das 1993). The actual mechanism of N antitermination is not known but a model, based on the determined protein/protein interactions, in which the *nut* site recognized by the N protein is carried along the surface of the RNA polymerase during elongation has been proposed (Fig 1.10) (Greenblatt, 1992).

1.5 Amidase operon of *Ps. aeruginosa*

Most species of the fluorescent *Pseudomonads* are able to grow on acetamide as the only

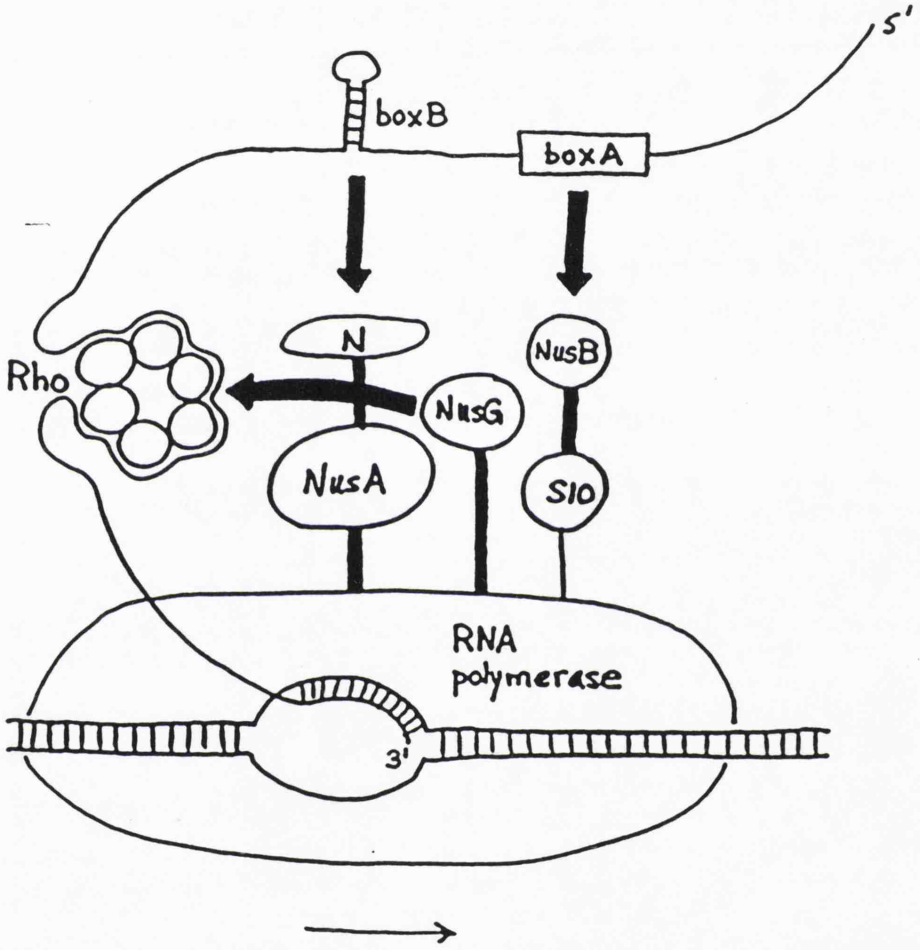


Fig 1.10 Model showing protein-protein interactions involved in antitermination by the bacteriophage λ N protein. The *nut* site (boxA + boxB) is made of RNA and is carried along the surface of RNA polymerase during elongation.

source of carbon and nitrogen by virtue of a chromosomally located acetamidase (Clarke and Ornston, 1975). Initial studies of the amidase system of *Ps. aeruginosa* PAC1 showed that enzyme activity was inducible and that the system showed different inducer/substrate specificities (Kelly and Clarke, 1962). The isolation of Ami- mutations (presumed to be structural gene mutations, *amiE*-) and Ami^{con} mutations (presumed to be in the regulatory gene *amiR*^{C}) allowed bacteriophage F116 transduction studies which showed that the two genes were closely linked (Brammar *et al.*, 1967). Later studies showed that *amiR* functioned as a positive regulator (Farin and Clarke, 1978) and that amidase expression was subject to catabolite repression by succinate and other TCA cycle intermediates (Smyth and Clarke, 1975). The *crc* (catabolite repressor) gene of *Pseudomonas aeruginosa* has been cloned recently although its site of action and mechanism have not yet been characterized (MacGregor *et al.*, 1992). Catabolite repression of enzymes for the degradation of N containing compounds in *Pseudomonas* can normally be relieved by N starvation (Van der Drift and Janssen, 1985) and although this effect has been reported in the amidase system of *Ps. aeruginosa* PAO1 (Janssen *et al.*, 1980), it has not been observed in *Ps. aeruginosa* PAC1 (Potts and Clarke, 1976).

The amidase genes from the high expressing constitutive mutant PAC433 were initially identified from a λ phage library, subcloned into the *E. coli* vector pBR322 and plasmid pJB950 characterised (Drew *et al.*, 1980; Clarke *et al.*, 1981). The PAC1 genes were subsequently cloned on a 5.381 kb DNA fragment in the same vector and the recombinant plasmid pAS20 (Fig 1.11) constructed and shown to have all the genes required for inducible amidase phenotype (Wilson and Drew, 1991). These two constructs have proved invaluable in the molecular study of amidase gene organisation and regulation.

1.5.1 Amidase structural gene *amiE*

Amidases with relatively relaxed specificities and showing substantial homologies to each other have been identified in *Ps. aeruginosa* (Clarke, 1984), *Ps. putida* (Clarke, 1972), *Arhtobacter* Spp (Asano *et al.*, 1982) and *Brevibacterium* species (Soubrier *et al.*, 1992). These amidases however show no sequence homology to other reported enantiomer selective amidases distinguishing themselves as a separate class of enzymes (Mayaux *et al.*, 1990; 1991).

The *Ps. aeruginosa* amidase has been purified, characterised and sequenced (Brown *et al.*, 1973; Ambler *et al.*, 1987). It is hexameric in its native form with an Mr. of 200,000 and catalyses hydolysis and transferase reactions typical of this type of amidase:

Reactions		Optimal substrate
Amide hydrolysis RCONH + H2O	RCOO' + NH_4^+	Propionamide
Amide transferase RCONH + NH2OH	RCONHOH + NH_4^+	Acetamide
Acid transferase RCOO' + NH2OH + H^+	RCONHOH + H2O	Acetate
Ester hydrolysis R' COOR + H2O	R'COO + H^+ + ROH	Ethyl acetate
Ester transferase R' COOR + NH2OH	R'CONHOH + ROH	Ethyl acetate

Fig 1.13 Amidase catalysed reactions

Under non-inducing conditions

Under inducing conditions

Fig 1.11 Amidase operon of *Ps. aeruginosa* PAC1
Restriction sites are **H** (*Hin*dIII), **P** (*Pst*I), **X** (*Xho*I), **K** (*Kpn*I), **E** (*Eco*RV), and **S** (*Sal*I)
Thick and thin lines show transcription from the *amiE* promoter.

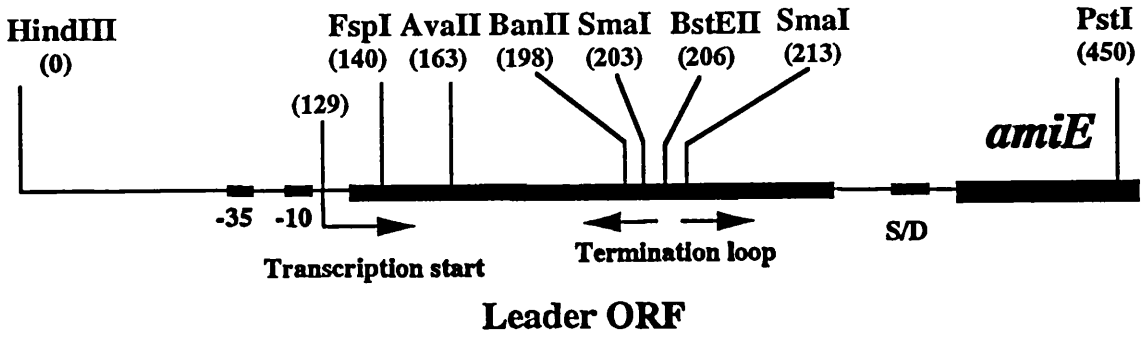


Fig 1.12 Amidase leader region organisation

The most fundamental and important aspect of the studies of the *Ps. aeruginosa* amidase has been the demonstration of experimental enzyme evolution by Clarke and others (Clarke and Drew, 1988). Using chemical mutagenesis and selective pressure, amidases with novel substrate specificities have been isolated and partially characterised. Mutants have been isolated that grow on butyramide and produce B-amidase, valeramide and produce V-amidase, phenylacetamide and produce Phe-amidase. All these enzymes are closely related to each other and are possibly point mutation products of the structural gene (Clarke and Ornston, 1975). The B-amidase in addition to hydrolysing butyramide also has a higher affinity for lactamide compared to the wild type A-amidase. Strain B6 which produced B-amidase was the starting point for a family of mutants producing other altered enzymes (Clarke and Slater, 1986). The altered specificity of B-amidase is due to a single point mutation within *amiE* which substitutes a phenylalanine for a serine at position 7 from the N-terminus (Clarke *et al.*, 1981).

The amidase gene has been sequenced and compared with the complete amino acid sequence of the purified enzyme (Brammar *et al.*, 1987; Ambler *et al.*, 1987). The codon usage is highly biased to a G or C in the third position characteristic of *Pseudomonas* genes (Brammar *et al.*, 1987).

1.5.2 *amiE* leader region

The DNA sequence upstream of *amiE* has been determined and studies of this region led to the proposal of the antitermination regulatory mechanism (Drew and Lowe, 1989). (Fig 1.12) Sequence studies showed that a potential σ^{70} promoter is located 150 bp upstream of the *amiE* initiation codon sequence, followed by a leader region containing a short ORF of

35 amino acids without a discernable Shine/Dalgarno sequence. Partly overlapping the 3' end of the ORF is a potential Rho-independent transcription terminator, and the Shine-Dalgarno sequence is located 9 bp upstream of the *amiE* initiation codon. It was shown that disruption of the terminator by a 10 bp deletion using fortuitously located *SmaI* targets led to AmiR independent constitutive amidase expression. The DNA sequence upstream of the transcription terminator was shown to have some homology to the sequences just upstream of the BglG regulated terminators. Subsequent investigations using primer extensions confirmed the location of the promoter, and showed that it functioned constitutively as demanded by an antitermination regulatory mechanism and extended the potential recognition sequence homology to the *sacB* upstream terminator (Wilson, 1991).

1.5.3 Positive regulator gene *amiR*

The *amiR* gene was initially defined phenotypically (incorrectly) by constitutive regulatory mutants selected from succinate/formamide plates (Brammar *et al.*, 1967). Later, mutants that produced a thermolabile regulator that had an amidase constitutive phenotype at low temperatures and amidase negative phenotype at high temperatures were isolated (Farin and Clarke, 1978). These results showed that amidase expression was positively regulated and correctly identified *amiR*. The positive control model was later confirmed by mobilising a cloned *amiE* gene into a variety of previously characterised *Ps. aeruginosa* amidase mutants (Drew, 1984). Studies performed in *E. coli* with constructed deletions of plasmid pJB950 localised *amiR* to within a 1.1 kb *ClaI*/*XhoI* DNA fragment more than 2 kb downstream of *amiE* and demonstrated that *amiE* and *amiR* were transcribed in the same direction (Cousens, 1985; Cousens *et al.*, 1987). This study also established high constitutive expression of amidase in *E. coli* by transcomplementing the *amiE* gene cloned on one plasmid (pDC5) with the *amiR* gene fragment (1.5 kb *XhoI*/*XhoI*) cloned on a second compatible plasmid (pDC35).

The DNA sequence of *amiR* from the constitutive mutant PAC433 has been determined and analyzed (Lowe *et al*, 1989). The codon usage resembles that of *amiE* with only 41 of the

possible 61 sense codons being used. The 590 bp gene codes for a protein with a predicted Mr. of 22000 and the gene is located within a polycistron with an upstream ORF terminator codon overlapping the *amiR* initiation codon by 1 bp. No sequence homology was observed between AmiR and other antitermination factors in the data base.

1.5.4 *amiC* a negative regulator gene

Deletion, insertion and rearrangement studies of the DNA sequences upstream of *amiR* were carried out in *E. coli* using the cloned genes in pJB950 which led to elevated constitutive amidase expression (Cousens, 1985; Cousens *et al.*, 1987; Wilson, 1991) and in pAS20 which led to low constitutive amidase synthesis (Wilson, 1991; Wilson and Drew, 1991) suggesting a negative regulatory role for this region. Sequencing of the region from pAS20 between the unique *Kpn*I and the *Pvu*II target within *amiR* revealed an open reading frame preceded by two potential σ^{54} promoters and a SD sequence (Wilson, 1991; Wilson and Drew, 1991). This ORF was then cloned, and shown to repress amidase synthesis from cloned *amiE* and *amiR* genes in *E. coli* , and in constitutive strains of *Ps. aeruginosa* showing that it encoded a new negative regulator of amidase expression. The protein product of this ORF, AmiC, has been purified and characterized (Wilson, 1991; Wilson and Drew, 1991). In recent studies, AmiC has been crystallized and the three dimensional structure with two domains has been determined (O'Hara, Wilson, Drew and Pearl, personal communication). In general terms AmiC appears to be closely related to a family of periplasmic binding proteins, particularly to the LivJ (leucine, isoleucine,

valine binding protein) of *E. coli.* AmiC has also been shown to bind aliphatic amides with a Kd similar to that of periplasmic branched-chain amino acid binding proteins and appears to be the sensory component of a sensor/regulator pair, with AmiR as the regulator (Wilson *et al.*, 1993).

1.5.5 *amiB* and *amiS* open reading frames

In their studies with mutator phage D3112, Rehmant and Shapiro (1983) isolated amidase negative mutants of *Ps. aeruginosa* PAO1 and determined the loci of insertion by Southern blotting. Insertions were found in *amiE*, *amiR*, and in the region immediately downstream of *amiE* the latter cases being described as 'leaky' due to background amidase expression. This region was investigated in two ways (Wilson, 1991). Attempts were made first to look for a protein product by making a plasmid construct carrying a DNA fragment containing this region downstream of a strong promoter with no success. Secondly, this region was sequenced and an ORF, *amiB* identified, which showed a large increase in the third position G+C bias and a relatively low usage of rare codons. *amiB* would code for a 371 amino acid (Mr 42000) protein. Homology searches revealed substantial homology of AmiB to the regulatory subunit of the Clp ATP-dependent proteases (Gottesman *et al.*, 1990).

The DNA sequencing of *amiR* included some downstream sequences and indicated a potential additional ORF (Lowe *et al.*, 1989). However, previous studies that involved deletions of this region led to no obvious change of phenotype (Cousens, 1985; Cousens *et al.*, 1987; Wilson, 1991). The DNA sequence of this region was determined and a 800 bp ORF, *amiS*, identified that would code for a highly hydrophobic protein with 6 potential transmembrane helices (Wilson, 1991). Recent structural analysis using 'threading' of the protein sequence to folds of known X-ray structures have indicated that AmiB/AmiS may

form components of a putative ABC-type transporter involved in amide uptake (Wilson *et al.*, in preparation).

1.6 Aims of the study

In broad terms, this study was intended to elucidate the mechanistic details of amidase induction and transcriptional regulation. Specifically,

1. To determine gene expression patterns under both inducing conditions and noninducing conditions with a view to identifying the necessary components for induction.
2. To study the transcription antitermination reaction *in vivo* and *in vitro* with a view to demonstrating and characterizing the protein/RNA interaction.

CHAPTER 2

Materials and Methods

2.1 Materials

2.1.1 *E. coli* strains

The *E. coli* strains used in this study are listed in Table 2.1(a). Recombination deficient *E. coli* hosts were used for propagating plasmids and bacteriophages in order to maintain partial diploids. Strain JA221 was used for cloning and *trans* complementation studies of amidase genes mainly because most of the previous work was done in this host (Drew, personal communication). Strain DH1 was used for *trans* complementation, cloning and as a donor in three strain plasmid mobilisation experiments with strain HB101 carrying plasmid pRK2013 as helper strain (Deretic *et al.*, 1986). *E. coli* strain GM2163 was used to propagate plasmids that were to be cut with *dam* methylase sensitive restriction enzymes. Strains JM107 and JM109 are deleted for the chromosomal *lac* operon and neighbouring genes for proline biosynthesis and they also carry F' (*traD*36 *proAB*+ *lacI*^{q} *lacZ*ΔM15), the *lac* mutation allows for α–complementation from a plasmid or phage with a functional N-terminus of *lacZ* gene (Yanish-Perron *et al.*, 1985). Recombinants for colour selection i.e. pUC18/19, M13mp18/19 and phagemid based constructs were propagated in these strains. Strain C600 was used for protein expression using the plasmid pT7.7/pGP1-2 system (Tabor 1990). For site directed mutagenesis, *in vitro* synthesised duplexes were transformed into strain BMH 71-18 *mutS* (Kramar *et al.*, 1984) prior to segregation in *E. coli* JM109.

2.1.2 *Ps. aeruginosa* strains

The *Ps. aeruginosa* strains used in this study are listed Table 2.1(b). *Ps. aeruginosa* strains were used in this study on the basis of their amidase phenotype. Strain PAC1 is the amidase wild type strain (Kelly and Clarke, 1962). Strain PAC111 is derived from PAC1 and is amidase constitutive, but butyramide repressible (Brammar *et al.*, 1967). Strain PAC452 (*ami*Δ161) is an amidase deletion mutant (R.E.Drew personal communication). Strain PAC327 is a regulatory defect mutant of PAC1 (Drew, 1984).

Table 2.1 Bacterial strains

a) *E. coli*

Strain	Genotype	Reference
JA221	*hsdR recA trp leu*	Clarke and Carbon (1978)
C600	*thr leu thi*	Harayama *et al* (1980)
JM107	*thi* Δ (*lac- proAB*) F'*lacZ* ΔM15, *proAB*	Yanisch-Perron *et al*(1985)
JM109	*thi* Δ (*lac- proAB*) F'*lacZ* ΔM15, *proAB*	Yanisch-Perron *et al*(1985)
HB101	$hsdS_B$ $recA_1$ *endA*1 *thi-1* $relA_1$,*proA*	Boyer and Roulland-Dussoix (1969)
DH1	*hsdR, recA, thi*	Hanahan (1983)
GM2163	*dcm*-6, *dam*-13	
BMH 71-18 mutS	*thi, supE,* Δ(*lac-ProAB*) [*mutS*::Tn10]	Promega, protocols and applications manual
JM109 (DE3)	*thi* Δ (*lac- proAB*) F'*lacZ* ΔM15, *proAB* λ(DE3)	Promega Protocols manual

b) *Ps. aeruginosa*

Strain	Genotype	Amidase Phenotype	Reference
PAC1	*amiE+ amiR+*	inducible	KellyandClarke(1962)
PAC111	*amiE+amiR*11	constitutive	Brammar *et al* (1967)
PAC327	*amiE* + *amiC*35 *amiR*33 *crp7*	Ami-	Brown (1969)
PAC452	*amiΔ*	Ami-	Day (1975)

2.1.3 Plasmids and Bacteriophages

Plasmids used in this study are listed in Table 2.2 (a) and new plasmids constructed are listed in table 2.2(c). Plasmids pUC18/19 were used to obtain large quantities of DNA fragments of interest because of their high copy number and ease of selection of recombinants using the blue/white colour screen with the chromogenic substrate X-gal after IPTG induction. Plasmids pGEM3Z/4Z have the additional advantage of having opposing bacteriophage T7 and SP6 promoters at the multiple cloning site which could be used for synthesis of high specific activity RNA (Promega). Plasmids pT7.7 and pGP1-2 are respectively a custom made protein expression vector and T7 RNA polymerase delivery plasmid (Tabor 1990). The pALTER phagemid (Promega) was used for all site directed mutagenesis. Bacteriophages M13mp18/19 (Messing, 1983) were used for cloning DNA fragments to be used for SS DNA sequencing (section 2.2.2.15). DNA fragments to be mobilised into *Ps. aeruginosa* were cloned into the IncQ-derived broad-host-range vectors pKT231 (Bagdasarian *et al.*, 1981) and pMMB66H/E (E/H) (Morales *et al.*, 1990).

Table 2.2 Plasmids

a)

Plasmid	Size	Markers	Reference
pBR322	4.4	Ap^RTc^R	Bolivar *et al* (1977)
pUC18	2.7	Ap^R	Yanisch-Perron *et al* (1975)
pUC19	2.7	Ap^R	Yanisch-Perron *et al* (1975)
pBGS19-	4.4	Km^R	Spratt *et al* (1986)
pRK2013	-	Km^R	Figurski and Helinski (1979)
pKT231	13.0	Sm^R Km^R	Bagdasarian *et al* (1981)
pMMB66HE	8.9	Ap^R	Morales *et al* (1990)
pAS20	9.1	Ap^R *ami* ind	Wilson and Drew (1991)
pJB950	9.1	Ap^R *ami* con	Clarke *et al* (1981)
pSW24	4.2	Ap^R *amiR*	Wilson (1991)
pSW35	14.5	Sm^R *amiR*	Wilson (1991)
pTM1	7.0	Cm^R *amiE*	Wilson (1991)
pTM2	5.9	Km^R *amiR*	Wilson (1991)
pSW41	10.3	Ap^R *amiC*	Wilson and Drew (1991)
pSW101	18.1	Sm^R *ami*	Wilson (1991)
pALTER™	5.68	Tet^R, f1, *lacZ*ΔM15	Promega protocols and applications manual
pT7-7	2.47	Ap^R pT7	Tabor (1990)
pGP1-2	7.14	Km^R T7RNAP	Tabor (1990)

b) Phages

Phage	Markers	Reference
M13mp18	*lacZ*ΔM15	Yanisch-Perron *et al* (1985)
M13mp19	*lacZ*ΔM15	Yanisch-Perron *et al* (1985)
M13KO7		Promega protocols and applications manual
R408		Promega protocols and applications manual

c) Plasmids constructed in this study

Plasmid	Markers	Plasmid	Marker
pMW1	ApR *amiECRS*	pMW30	KmR *amiR*
pMW6	ApR TetR *amiE* LeaderM	pMW40	ApR *amiR*
pMW9	ApR TetR *amiE*	pMW41	ApR *amiR*
pMW10	ApR TetR *amiE* LeaderM	pMW42	ApR *amiE* Leader
pMW11	ApR TetR *amiE* LeaderM	pMW44	ApR *amiE* leader
pMW12	ApR TetR *amiE* LeaderM	pMW45	ApR *amiR*
pMW13	ApR TetR *amiE* LeaderM	pMW40L	ApR *ami*lead.*amiR*
pMW14	ApR TetR *amiE* LeaderM	pMW46	ApR *ami*lead.*amiR*
pMW15	ApR TetR *amiE* LeaderM	pMW47	ApR *ami*lead.*amiR*
pMW16	ApR TetR *amiE* LeaderM	pMW48	ApR *ami*lead.*amiR*
pMW21	ApR *amiCRS*	pMW49	ApR *ami*lead.*amiR*
pMW22	SmR *amiECRS*	pMW50	ApR *amiR*
pMW25	ApR TetR *amiE* LeaderM	pMW51	ApR *amiR*
pMW26	ApR TetR *amiE* LeaderM	pSB100	ApR *ami*Lead.
pMW27	ApR TetR *amiE* LeaderM	pSB101	ApR TetR *amiE*
pMW28	ApR TetR *amiE* LeaderM	pSB102	ApR TetR *amiE* leaderM
		pSB103	ApR TetR *amiE* LeaderM

2.1.4 Bacteria growth media recipes

a) Liquid media

i) Nutrient broth

The nutrient broth used was either Oxoid No.2 broth prepared as recommended or L broth; Bacto tryptone 10 g/l, Bacto yeast extract 5 g/l, sodium chloride 5 g/l and glucose 1 g/l (pH7.2) made up in deionised water and sterilized by autoclaving for 20 minutes at 15 psi Terrific broth was prepared by mixing in 900 ml deionised water; 12 g/l Bacto tryptone, 24 g/l Bacto yeast extract and 4 ml/l glycerol, autoclaving for 20 minutes at 15 psi. on

liquid cycle and adding 100 ml of a sterile solution of 0.17 M KH_2PO_4, 0.72 M K_2HPO_4 (Sambrook *et al.*, 1989). 2 x YT medium was prepared by dissolving 16 g bacto-tryptone, 10 g bacto-yeast extract and 5 g NaCl into 900 ml of water, adjusting the pH to 7.0 with 5 N NaOH, adjusting the volume to 1 litre with deionised water and autoclaving as described above.

ii) Minimal medium

E. coli basal medium with nitrogen contained: 4.5 g/l Na_2HPO_4, 2.0 g/l KH_2PO_4, 1.0 g/l $(NH_4)_2SO_4$, 1.0 g/l NH4Cl, 0.5 g/l KNO_3, 0.8 % (v/v) sodium lactate, 1.0 % (v/v) *E. coli* trace elements, pH 7.0-7.2 (Clarke and Laverack, 1983).

E. coli basal minimal medium without nitrogen was prepared as follows : 4.5 g/l Na_2HPO_4, 2.25 g/l KH_2PO_4, 2.5 g/l Na_2SO_4, 1 g/l NaCl, 1.0 % *E. coli* trace elements, pH 7.0-7.2. The media was supplemented with 0.01 % L-tryptophan and 0.01 % L-leucine as sources of nitrogen (P. Laverack, personal communication).

The *E. coli* trace element solution contained: 40 g/l Na citrate, 20 g/l $MgSO_4.7H_2O$, 2 g/l $FeSO_4.7H_2O$, 1.0 g/l $MnCl_2.2H_2O$ and $CaCl_2.2H_2O$ and 0.02 g/l of each of $ZnSO_4.7H_2O$, $Na_2B_4O_7.7H_2O$, $(NH_4)_6Mo_7O_{24}.7H_2O$, $CuSO_4.5H_2O$ and $CoCl_2$.

Pseudomonas basal medium was that of Brammar and Clarke (1964): 12.5 g/l K2HPO4, 3.8 g/l KH2PO4, 1.0 g/l (NH4)2SO4, 0.1 g/l MgSO4.7H2O and 0.5 % trace elements solution (pH 7.2).

Pseudomonas trace elements solution was that described by Kelly and Clarke (1962): 0.23 g/l H_3BO_3, 0.174 g/l $ZnSO_4.7H_2O$, 0.116 g/l $FeSO_4$ $(NH_4)_2SO_4.6H_2O$, 0.096 $CoSO_4.7H_2O$, 0.022 g/l $(NH_4)_6Mo_7O_{24}.4H_2O$, 0.008 g/l $CuSO_4.5H_2O$, .008 g/l $MnSO_4.4H_2O$.

All solutions were autoclaved at 15 psi for 15 minutes and stored at room temperature and the media supplemented with sterile carbon sources at 0.5 % for *E. coli* and 1.0 % for *Pseudomonas* as indicated in the respective experiments.

b) Solid support media

For plates and storage stabs, 1.2 % (w/v) Difco agar was added to either L-broth or Oxoid No. 2 or to minimal media prior to autoclaving. Top lay agar was made similarly except that 0.6 % agar was used.

2.1.5 Antibiotics

The antibiotics were added to the appropriate bacterial growth media at the following final concentrations:

	E. coli	*P. aeruginosa*
Ampicillin	100 µg/ml	
Carbenicillin		500 µg/ml
Kanamycin	100 µg/ml	250 µg/ml
Streptomycin	100 µg/ml	500 µg/ml
Chloramphenicol	10 µg/ml	

2.1.6 Dilution buffer

Where necessary, bacteria were diluted in buffer prepared as described by Brammar *et al*(1967): 3.0 g/l KH_2PO_4, 7.0 g/l Na_2HPO_4, 4.0 g/l NaCl, 0.2 g/l $MgSO_4.7H_2O$, pH 7.2.

2.1.7 Phenol:Chloroform

Phenol:chloroform for miniplasmid preparation and other DNA purification procedures was prepared as follows (Sambrook *et al.*, 1989): Analytical grade phenol was equilibrated with 100 mM Tris-HCl (pH 8.0), mixed with equal volume of chloroform:isoamyl alcohol (24:1) and stored in a light-tight container at -20 °C under the equilibration buffer.

2.1.8 DNAase free RNAase

Pancreatic Ribonuclease A was dissolved at a concentration of 10 mg/ml in 10 mM Tris.Cl (pH 7.5), 15 mM NaCl. It was then heated at 100 °C for 15 minutes, cooled slowly to room temperature and stored in aliquots at -20°C.

2.1.9 Reagents

The following reagents and kits were used in the study:

Reagent/kit	Supplier
Acetic acid	James Burroughs
Acrylamide	Sigma
Agarose	Sigma
Ampicillin	Beecham

Bacto agar	Difco
Bacto tryptone	Difco
Carbenicillin	Sigma
RNAase free DNAase	Promega
Carbenicillin	Sigma
DNA modifying enzymes	Anglian Biotechnology Ltd
DNA sequencing kits	Pharmacia
Ethanol	James Burroughs Ltd
Geanclean kit	BIO 101 Inc.
Hydroxylamine	Sigma
Immunochemicals	Sigma
IPTG	Pharmacia
Kanamycin	Sigma
Lactamide	Sigma
Lysozyme	Sigma
Methanol	James Burrough Ltd
3MM paper	Whatman Ltd
NAP-5 columns	Pharmacia
Oligonucleotides	Kings college London
Qiagen DNA purification kit	Qiagen Inc.
Ribonuclease A	Sigma
Radiochemicals	NEN, Sigma
Restriction enzymes	Anglian Biotechnology Ltd
Reverse sequencing primer	Promega
Riboprobe transcription system	Promega
Ribonuclease inhibitor	Pharmacia
Streptomycin	Sigma
T7 sequencing primer	Promega
X-gal	Northumbrian Biotech. Ltd
X-ray film	Fuji Photo Film Co.
Bacto yeast extract	Difco

2.2 Methods

2.2.1 Microbiological Methods

2.2.1.1 Bacterial growth conditions

Bacteria were routinely grown in 20 ml universals containing 5 ml media on a reciprocating shaker at 37°C. Large scale growth was carried out in conical flasks of at least 5 times the volume of culture. Growth was monitored by measuring the optical density at 450 nm for *E. coli* and 670 nm for *Ps. aeruginosa.*

2.2.1.2 Strain storage

Ps. aeruginosa and *E. coli* strains for long term storage were aliquoted into 40 % sterile glycerol and stored at -70°C. For short term storage, an inoculating loop with the bacteria was stabbed into sterile L/agar bottles and stored in the dark at room temperature as described in Sambrook *et al* (1989).

2.2.1.3 Amidase assays

Amidase assays were performed on whole cells by the hydroxamate transferase method of Brammar and Clarke (1964) using acetamide as substrate. The bacteria were grown overnight on most occasions either in minimal medium or nutrient broth under the conditions specified in the respective results sections.

0.1 ml of appropriately diluted bacterial suspension were added to 0.9 ml mixed substrate (2 volumes of 100 mM Tris-HCl buffer, pH 7.2, 1 volume of 0.4 M acetamide, 1 volume of 2 M hydroxylamine neutralised to pH7.2 with NaOH) at 37°C and incubated for 10 minutes. The reaction was stopped by addition of 2 ml of ferric chloride/HCl reagent (100 ml of 60 % ferric chloride solution and 57 ml of concentrated HCl made up to 1 litre with

distilled water). The tubes were vortexed to disperse nitrogen bubbles and the A_{500} determined in a Unicam SP600 spectrophotometer against a reagent blank.

The units of amidase activity expressed as µmoles of acetylhydroxamate formed/min/mg of bacteria were calculated from the following formula:

$$\frac{A_{500} - (0.08 \times A_{670(450)}) \times 3.5}{A_{670(450)} \times 0.56}$$

2.2.1.4 Competent cell preparation and transformation of *E. coli*

Competent cells were prepared by $CaCl_2$ treatment. A single colony from a fresh plate was grown overnight in 5 ml LB medium and 1 ml then used to inoculate 50 ml LB at 37 °C. Aeration was continued at 37°C to an $O.D_{450}$ of 0.4-0.5, the cells chilled on ice for 10 minutes and collected by centrifugation at 5000 rpm in a Sorvall RC2-B centrifuge at 4 °C. The cells were resuspended in 25 ml of ice cold 100 mM $MgCl_2$ and collected as above. The cell pellet was then resuspended in 2.5 ml of ice cold 100 mM $CaCl_2$ and held on ice for 1hr.

Transformation was carried out by adding 100 ng-1 µg of plasmid DNA in 100 µl SSC:$CaCl_2$ (3:4) to 200 µl of the competent cells. The transformation mixture was held on ice for 45 minutes, heat shocked at 37°C for 5 minutes, chilled again on ice for 45 minutes and then 0.7 ml of LB medium at 37°C added. The cells were allowed to recover by growing for at least 1 hr without shaking at 37°C and plated by overlaying 300 µl in

0.6 % top agar onto antibiotic selective plates. The plates were incubated at 37°C overnight except for the heat inducible plasmids whereby the plates were incubated at 30 °C for 16-18 hours.

2.2.1.5 Transfection of M13mp18/19 recombinants into *E. coli*

E. coli strain JM109 for propagation of M13mp18/19 single strand DNA were transformed with DS DNA as described above except that; the overnight cultures were grown from minimal medium plate colonies, the transformants after heat induction were placed on ice and plated as soon as possible by overlaying in top agar containing, 200 µl of lawn cells (mid log phase cells of the same culture), X-gal (20 µl of 50 mg/ml stock) and IPTG (100 µl of 100 mM stock). The plates were incubated overnight at 37 °.

2.2.1.6 Mobilisation of plasmids from *E. coli* into *Ps. aeruginosa.*

E. coli strain HB101 carrying plasmid pRK2013 was used as the helper strain in triparental matings (Deretic *et al.*, 1986). Overnight cultures of the donor carrying a broad-host-range vector based construct, the helper strain, and the recipient strain were plated together on nutrient agar plates and incubated at 37°C in an upright position for six hours. The lawn of bacteria was harvested and washed several times in dilution buffer before being plated onto antibiotic selective plates.

2.2.2 DNA methods

2.2.2.1 Isolation of plasmid DNA

Single colonies were patched out onto antibiotic selective plates and the patches used to

inoculate 5 ml LB media in universal bottles. Plasmid DNA was isolated by the alkaline lysis method as described in Sambrook *et al* (1989). The liquid cultures were grown overnight at 37°C with vigorous shaking and 1.5 ml of the culture harvested by centrifugation at 12000 x g in an Eppendorf microcentrifuge for 1 minute. The media was aspirated leaving the pellet as dry as possible and the pellet resuspendend by vortexing in ice cold miniprep lysis buffer (Solution I) (25 mM Tris-HCl, pH8.0, 10 mM EDTA, 50 mM glucose, 4 mg/ml lysozyme). The mixture was incubated at room temperature for 5 minutes and 200 µl of freshly prepared miniprep Solution II (0.2 N NaOH, 1 % SDS) added. The tube was inverted several times to mix the contents and placed on ice for 5 minutes. The mixture was rapidly neutralised by addition of 150 µl of miniprep Solution III (60 ml 5 M potassium acetate mixed with 11.5 ml glacial acetic acid and 28.5 ml water) was added and the tube vortexed in an inverted position. The tube was placed on ice for a further 5 minutes then centrifuged at 12000 x g for 5 minutes. The supernatant was transferred to a fresh tube, an equal volume of phenol:chloroform added and vortexed for 1 minute followed by centrifugation for 2 minutes at 12000 x g. The upper aqueous phase was transferred to a fresh tube and the DNA precipitated by addition of 2 volumes of ethanol chilled at -20°C. The tube was held on ice for 10 minutes and centrifuged at 12000 x g for 15 minutes at 4 °C to collect the DNA. The pellet was washed by addition of 70 % ethanol centrifuged for 10 minutes, the ethanol removed by aspiration and the pellet dried at room temperature before resuspension in an appropriate volume of TE buffer (10 mM Tris-HCl, pH8.0, 1 mM EDTA) containing RNAase A at a concentration of 20 µg/ml. This procedure was also used to prepare plasmid DNA from *Ps. aeruginosa.*

2.2.2.2 Column purification of plasmid DNA

Plasmid DNA for sequencing was sometimes purified from anion exchange columns (Qiagen Inc). The cells required to give up to 100 µg of plasmid DNA were harvested and resuspended in 4 ml buffer P1 (100 µg/ml RNAase A, 50 mM Tris/HCl, 10 mM EDTA, pH8.0), 4 ml of buffer P2 (200 mM NaOH, 1% SDS) was added, mixed and then incubated at room temperature for 5 minutes. 4 ml of chilled buffer P3 (3.0 M potassium acetate, pH5.5) was added, mixed, and incubated for 15 minutes on ice. The bacterial lysate was centrifuged at 30,000 x g and the supernatant removed promptly then loaded on a Qiagen-tip 100 pre-equilibrated with 4 ml buffer QBT (750 mM NaCl, 50 mM MOPS, 15 % ethanol, pH 7.0, 0.15 % Triton X-100). The column was washed with 2 x 10 ml of buffer QC (1.0 M NaCl, 50 mM MOPS, 15 % ethanol, pH 7.0) then eluted with 5 ml buffer QF (1.25 M NaCl, 50 mM Tris-HCl, 15 % ethanol, pH 8.5). The DNA was concentrated by precipitation with 0.7 volumes isopropanol, recovered by centrifugation at 12000 x g, washed with 70 % ethanol and resuspended in an appropriate volume of sterile water.

2.2.2.3 Large scale isolation of plasmid DNA and CsCl gradient purification

Large scale purification of plasmid DNA was carried out by a modification of the alkaline lysis method and purified on CsCl as described in Sambrook *et al* (1989). The cells were grown overnight at 37°C in 500 ml cultures, chilled on ice for 10 minutes and harvested by centrifugation at 5000 rpm for 15 minutes at 4°C in a Sorvall GS3 rotor. The supernatant was discarded and the centrifuge tube stood upside down to drain all the supernatant. The bacterial pellet was resuspended in 20 ml Solution I containing lysozyme at 1 % w/v and incubated at room temperature for 5 minutes. 50 ml of freshly prepared Solution II was

added, mixed by inverting the tube several times and stored at room temperature for a further 5 minutes. 50 ml of ice cold Solution III was then added and mixed by shaking the tube several times. The tube was stored on ice for 10 minutes to precipitate chromosomal DNA, high molecular weight RNA, membranes and proteins. The bacterial lysate was then centrifuged at 7000 rpm for 30 minutes at 4°C in a Sorvall GS3 rotor and the rotor allowed to stop without braking. The supernatant was filtered through four layers of cheese cloth and the nucleic acids precipitated by addition and mixing of 50 ml of 50 % polyethylene glycol (mwt 6000) and leaving on ice for 1 hour. The DNA was recovered by centrifugation at 15000 rpm in a sorvall GSA rotor for 30 minutes at 4°C. The supernatant was drained off making sure to remove all droplets of PEG, dried and resuspended in 5 ml TE buffer. The DNA was extracted with phenol:chloroform twice and the volume made up to 5 ml with water before addition of CsCl at 1.1 g/ml and 200 µl of ethidium bromide (10 mg/ml in water). The solution was transferred into a Beckman quick-seal pollyallomer tube and centrifuged in a Beckman L-7 ultracentrifuge Ti70.1 rotor at 20°C for twenty hours at 50,000 rpm. The Lower band of closed circular DNA was aspirated into a syringe fitted with a 21 gauge hypodermic needle being visualised under a long wave UV lamp. The ethidium bromide was removed by extraction with CsCl saturated isopropanol several times. The DNA/CsCl solution was diluted with 3 volumes of water and the DNA precipitated with two volumes of ethanol and the DNA concentration determined spectrophotometrically.

2.2.2.4 Preparation of single-stranded phage DNA

Well isolated plaques of potentially recombinant phages as determined by their colourless background from the blue/white screen (above) were used for SS DNA isolation as described in Promega protocols and applications guide. 5 ml TYP media was inoculated

with 200 µl of an overnight culture of *E. coli* JM109, phage from a clear well isolated plaque and grown with good aeration for between 6 and 8 hours. The cells from 1.5 ml of culture were removed by centrifugation at 12000 x g for 15 minutes in an Eppendorf microcentrifuge and the supernatant centrifuged again. 1.2 ml of this culture was transferred to a fresh microfuge tube and the phage precipitated by addition of 0.25 volume of phage precipitation solution (3.75 M ammonium acetate, pH 7.5, 20 % polyethylene glycol [mwt 8000]) and standing on ice for 30 minutes. The phage pellet was recovered by centrifugation at 12000 x g for 15 minutes and resuspended in 400 µl TE buffer. The phage was lysed by addition of 0.4 ml of chloroform:isoamyl alcohol (24:1) and vortexing for 1 minute. The phage lysate was centrifuged at 12000 x g for 5 minutes and the upper aqueous phase transferred to a fresh tube and phenol:chloroform extracted until no material appeared at the interface. The aqueous phase was then transferred to a fresh tube mixed with 0.5 volumes of 7.5 M ammonium acetate, 2 volumes of ethanol, stored at -20 ^{o}C for 30 minutes and centrifuged at 12000 x g for minutes. The ethanol was discarded, the pellet washed with chilled 70 % ethanol and dried under vacuum before being resuspended in 20 µl sterile water.

2.2.2.5 Preparation of single-strand phagemid DNA

The protocol for preparation of single strand phagemid DNA was identical to that of M13 except that *E. coli* cells carrying the phagemid were infected with helper phage to switch the phagemid from plasmid replication mode to f1 replication mode. 100 µl of an overnight culture of cells containing the phagemid were used to inoculate 5 ml TYP and the culture aerated for 30 minutes at 37 ^{o}C. The cells were then infected with helper phage R408 or M13KO7 at a multiplicity of infection (m.o.i) of 10 and growth continued for a

further 8 hours to overnight. The cells were pelleted and the phage isolated from the supernatant, and the DNA extracted as described above.

2.2.2.6 Estimation of DNA concentration

DNA concentrations were determined spectrophotometrically by reading the absorbance at 260 nm whereby double-stranded DNA has an absorbance of 1 at a concentration of 50 μg/ml and single-stranded DNA has the same absorbance at 40 μg/ml. The purity was determined by measuring the ratio of A_{260}/A_{280} with sequencing quality single-stranded DNA having a ratio of greater than 1.7. On occasions, the concentration was estimated by running an aliquot of DNA on agarose gels alongside markers of known concentration and staining with ethidium bromide.

2.2.2.7 Agarose gel electrophoresis of DNA

Agarose gels of between 0.6% to 2.0% were prepared in TBE buffer (0.045 M Tris-borate, pH 8.0, 0.001 M EDTA). The agarose suspension was melted in a loosely capped Duran bottle and cooled to 60°C before addition of ethidium bromide to a final concentration of 0.5 μg/ml before pouring. Gels were run submerged in TBE buffer and DNA samples were mixed with gel loading (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) at a ratio of 5:1 v/v before loading into the wells alongside molecular weight markers. Electrophoresis was carried out at a voltage of 1-5 V/cm. The DNA fragments were visualised under UV light and photographed.

2.2.2.8 Isolation of plasmid DNA from agarose gels

DNA fragments were purified from agarose gels by the freeze squeeze method as described by Tautz and Renz (1983) and by elution from a silica matrix (Glassmilk) using the

Geneclean kit of BIO 101 Inc. In the former technique, the band of interest was excised under long wave UV light, frozen and centrifuged through a silane coated glass wool plug. The eluate was phenol:chloroform extracted twice, ethanol precipitated and the yield estimated by running an aliquot of the DNA on an agarose gel.

In the Geneclean technique, the excised gel band was weighed and mixed with 0.5 volumes of TBE modifier and 4.5 volumes of NaI reagent. The mixture was incubated at 55°C for 5 minutes with occasional mixing to dissolve the agarose. 2 µl glass milk vortexed to obtain a homogeneous suspension was added, mixed and incubated on ice for 5 minutes with occasional mixing to bind the DNA. The mixture was centrifuged for 5 seconds at 12,000 x g and the supernatant discarded. The pellet was washed by resuspending 3 x in wash buffer and care was taken to remove all liquid in the final wash. The glass milk was resuspended in 5 µl sterile water and incubated at 55°C for 3 minutes to elute the DNA and the mixture centrifuged for 30 seconds. The elution step was repeated 3 x and the eluate pooled. This method gave consistently high recoveries of DNA compared to the freeze/squeeze method.

2.2.2.9 DNA cloning

All the techniques for cloning into plasmid vectors were as described in Sambrook *et al* (1989) unless otherwise stated. Plasmids and DNA fragments were routinely digested with restriction enzymes in 20 µl volumes using buffer and temperature conditions recommended by the supplier. On some occasions it was necessary to ligate DNA fragments with non-complementary protruding termini and the following methods were used to modify the ends.

2.2.2.10 Filling recessed 3' termini

0.2-5 μg of DNA was digested to completion with the appropriate restriction enzyme in a 20 μl reaction and a solution containing the desired dNTPs at a concentration of 1 mM added. 1 unit of Klenow fragment of *E. coli* DNA polymerase I was added for each μg of DNA in the reaction and the mixture incubated for 15 minutes at room temperature. For DNA fragments purified on agarose gels or digested with restriction enzymes that were not heat inactivated, the DNA was resuspended in TE buffer pH 7.6 and $MgCl_2$ added to a concentration of 5 mM before being treated as above. The Klenow polymerase was then inactivated by heating to 75°C for 10 minutes or extraction with phenol:chloroform after making up the volume to 100 μl with TE buffer (pH 7.6) and precipitated with ethanol.

2.2.2.11 Removing 3' termini

Protruding 3' termini were removed using the 3'->5' exonuclease activity of bacteriophage T4 DNA polymerase. 0.2-5 μg of DNA was digested in a 20 μl reaction volume and 1 μl containing 2 mM of dNTPs added. 1-2 units of T4 DNA polymerase was added for each μg of DNA and the reaction incubated at 12°C for 15 minutes. The T4 DNA polymerase was inactivated by heating at 75°C for 10 minutes or by phenol:chloroform extraction after making up the reaction volume to 100 μl and the DNA precipitated with ethanol.

2.2.2.12 Generation of blunt ends using mung-bean nuclease

Plasmid DNA cleaved with *Nde*I and *Eco*RI that generate 5' and 3' protruding ends respectively was made blunt ended by digestion with mung-bean nuclease as described in Sambrook *et al* (1989). After restriction enzyme digestion, phenol:chloroform extraction,

ethanol precipitation and drying 5 µg of the linear DNA was resuspended in 2 µl of 10 x mung-bean nuclease buffer (300 mM sodium acetate, pH 4.5, 500 mM NaCl, 10 mM $ZnCl_2$ and 50% glycerol), 2.5 µl mung-bean nuclease (0.5 u/µl) and 5.5 µl distilled water. The reaction was incubated at 37°C for 1 hour, extracted once with phenol:chloroform, ethanol precipitated, dried and resuspended in ligation buffer.

2.2.2.13 Removal of 5' phosphate groups

To prevent recircularization of plasmid DNA cleaved with only one enzyme during cloning or prior to 5' terminal end labelling, the 5' phosphate groups were removed by calf intestinal phosphatase (CIAP) as described in Promega protocols and applications guide. To a 20 µl digest was added 10 µl of 10 x calf intestinal phosphate buffer (500 mM Tris-HCl, pH 9.0, 10 mM $MgCl_2$, 1 mM $ZnCl_2$ 10 mM spermidine), CIAP at 0.01 u/ρmol ends, the volume was made up to 100 µl and the reaction incubated at 37°C for 1 hour. The reaction was stopped by addition of 2 µl of 0.5 M EDTA, phenol:chloroform extracted and ethanol precipitated as described elsewhere.

2.2.2.14 Ligation of vector and insert DNA

Ligation of blunt ended and cohesive ended DNA fragments was carried out at 12°C overnight in a 10 µl reaction volume containing 1 µl of 10 x bacteriophage T4 DNA ligase buffer (200 mM Tris.Cl, pH7.6, 50 mM $MgCl_2$ 50 mM dithiothreitol, 2 mM ATP, 500 µg/ml bovine serum albumin), insert and vector DNA at a maximum ratio of 3:1, 1-5 Weiss units of T4 ligase. Ligation mixes were either transformed into fresh competent *E.*

coli cells or stored at -20°C.

2.2.2.15 DNA sequencing

DNA sequencing was carried out by the chain termination method of Sanger *et al* (1977) on both double stranded plasmid and single stranded M13/phagemid templates using the T7 sequencing™ kit of Pharmacia. 10 µl of single stranded DNA templates prepared as described above with a A_{260}/A_{280} of at least 1.7 and containing 1.5-2 µg of DNA was mixed with 2 µl of universal primer at 4.44 µg/ml (1-2 ρmol for other primers) and 2 µl of annealing buffer. The tubes were vortexed gently, centrifuged briefly and incubated at 60°C for 10 minutes and then held at room temperature for at least 10 minutes to complete the annealing. The labelling reactions were carried out by adding 3 µl of labelling mix-dATP (1.375 µM each dCTP dGTP and dTTP and 333.5 mM NaCl), 1µl of [α–^{35}S]dATP--S and 2 µl of diluted T7 DNA polymerase to the annealed primer and template. The components were mixed by gentle pipetting, collected at the bottom of the tube by brief centrifugation and incubated at room temperature for 5 minutes. In the termination reactions, 4.5 µl of the labelling reaction mix was transferred to each of four 2.5 µl aliquots of prewarmed sequencing mixes i.e. A mix-short (840 µm each dCTP, dGTP and dTTP; 93.5 µM dATP; 14 µM ddATP; 40 mM Tris-HCl, pH7.6, and 50 mM NaCl), C mix-short (840 mM each dATP, dGTP and dTTP; 93.5 µM dCTP; 17 µM ddCTP; 40 mM Tris-HCl, pH7.6,and 50 mM NaCl), G mix-short (40 µM each dATP, dCTP and dTTP; 93.5 µM dGTP; 14 µMddGTP; 40 mM Tris-HCl, pH 7.6, and 50 mM

NaCl) and T mix-short (840 µM each dATP, dCTP and dGTP; 93.5 µMdTTP; 14 µM ddTTP; 40 mM Tris-HCl, pH7.6, and 50 mM NaCl). Incubation was continued at 37°C for another 5 minutes and stopped by addition of 5 µl of stop solution (0.3 % each bromophenol blue and xylene cyanol FF; 10 mM EDTA, pH7.5, and 97.5 % deionised formamide). The tubes were spun briefly to collect the contents at the bottom of the tubes and 3 µl aliquots transferred to another tube for denaturation and loading onto sequencing gels. The remainder was stored at -20°C for a second run which was carried out within one week.

Double-stranded sequencing was identical to that described above except that the template was denatured by addition of 8 µl of 2 M NaOH to 1-2 µg DNA in a total of 32 µl of water. The tubes were vortexed briefly and the contents collected at the bottom of the tube by gentle centrifugation. The mixture was incubated at room temperature for 10 minutes and neutralised by addition of 7 µl 3 M sodium acetate (pH4.8) and 4 µl of distilled water. 120 µl of 100 % ethanol was added and the tubes held at -70°C for 15 minutes to precipitate the denatured DNA. The DNA was collected by centrifugation for 15 minutes at 12,000 x g at 4°C, washed with 70 % ethanol and dried briefly under vacuum. The DNA was dissolved in 10 µl distilled water and annealing carried out as described above.

2.2.2.16 Denaturing sequencing gels

The sequencing products were separated on 8 M urea sequencing gels using uniform 0.4 mm or 0.2 mm to 0.6 mm wedge-shaped spacers and shark tooth combs. The gel mixture was prepared by mixing 15 ml of 40% acrylamide stock solution (38 g acrylamide, 2 g

N,N'-methylenebisacrylamide and distilled water to make up to 100 ml), 20 ml 5 x TBE (0.445 M Tris-borate, 2 mM EDTA, pH 8.0), 23 g urea and made up to 100 ml with distilled water. The mixture was degassed and polymerization initiated by addition of 200 µl of 10 % ammonium persulphate, mixing, then adding 75 µl TEMED (N,N,N',N'-tetramethylenediamine) mixing and pouring. The comb was inserted and the gel mould laid flat on the bench to allow the gel to polymerize for 45-60 minutes.

The comb was removed after polymerization of the gel and the sealing tape removed from the bottom. The gel was then set up in a BRL sequencing apparatus and prewarmed by running at 1200 V for the wedge-shaped, and 1700 V for the uniform gel, for 45-60 minutes. The wells were flushed with TBE before inserting the comb and the samples loaded after heating to 90°C for 2 minutes. The gels were run at constant voltage (above) sometimes with several loadings. At the end of the run, the gels were fixed in 10 % acetic acid, 10 % methanol in distilled water for 15 minutes, dried on the sequencing cycle on a BRL gel drier and exposed to X-ray film for 16-48 hours.

2.2.2.17 DNA End Labelling

pBR322 DNA *Hpa*II DNA fragments were dephosphorylated as described above and used as substrate for forward reaction end labelling with bacteriophage T4 polynucleotide kinase. 1-50 ρmoles of DNA, 5 µl of bacteriophage T4 polynucleotide kinase 10 X buffer (0.5 M Tris-HCl, pH 7.6, 0.1 M $MgCl_2$, 50 mM dithiothreitol, 1 mM spermidine HCl and 1 mM EDTA, pH 8.0), 50 pmoles [γ-^{32}P]ATP (3000 Ci/mmole; 10 µCi/µl) and 10-20 units of bacteriophage T4 polynucleotide kinase were mixed in a total volume of 50 µl and incubated for 30 minutes at 37°C. The reaction was stopped by the addition of sequencing

stop mix (above) and used as marker in the RNA purification.

2.2.2.18 5' Phosphorylation of oligonucleotides

Oligonucleotides were 5' phosphorylated as described in the Promega protocols and applications guide. 100 pmoles oligonucleotide, 2.5 µl of 10 x kinase buffer (500 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 50 mM DTT, 1.0 mM spermidine and 10 mM ATP) and 5 units of T4 polynucleotide kinase were mixed and the volume made up to 25 µl with distilled water. The reaction was incubated at 37°C for 30 minutes and then heated to 70°C for 10 minutes to inactivate the kinase.

2.2.2.19 *In vitro* Site-Directed mutagenesis

Site-Directed mutagenesis was carried out using the Altered Sites™ System of Promega. The DNA to be mutated was cloned into the pAlter phagemid (Promega) and single stranded templates prepared as described. In the annealing reaction, 0.05 ρmol of recombinant pALTER SS DNA, 0.25 ρmol ampicillin repair oligonucleotide, 1.25 ρmol mutagenic oligonucleotide and 2 µl 10 x annealing buffer were mixed and the volume made up to 20 µl with sterile water. The mixture was heated to 70°C for 5 minutes and allowed to cool slowly to room temperature. The reaction mixture was then placed on ice and the following additions made; 3 µl 10 x synthesis buffer, 1µl T4 DNA polymerase (10 u/µl), 1µl T4 DNA ligase (2 u/µl) and 5 µl sterile water. The synthesis/ ligation reaction was incubated at 37°C for 90 minutes and transformed into *E. coli* BMH 71-18 *mutS*.

MPP DNA was made from the transformants and retransformed into *E. coli* JM109 with selection on ampicillin plates. The required mutants were analyzed by restriction enzyme mapping and DNA sequencing.

2.2.3 RNA methods

2.2.3.1 General procedures

Sterile disposable plastic ware was used whenever possible. Glassware was baked at 180°C for 8 hours and covered in aluminium foil. Distilled water used to make buffers and other solutions was incubated at 37°C with 0.1% DEPC for 12 hours then autoclaved as described earlier to destroy the DEPC.

2.2.3.2 *In vitro* Transcription

In vitro RNA synthesis was carried out using the Promega Riboprobe kit. DNA fragments for *in vitro* transcription were cloned downstream of either the T7 or SP6 promoter in either plasmid pGEM4Z or pALTER. Runoff transcripts were made from plasmids linearized with enzymes that yielded either recessed 3' termini or blunt ends. Where no such sites were available, 3' protruding ends were made blunt using bacteriophage T4 DNA polymerase.

2.2.3.3 Labelled RNA synthesis

In the standard transcription protocol, the following components were added at room temperature in the order listed:

4 µl 5 x Transcription buffer (200 mM Tris-HCl, pH 7.5, 30 mM $MgCl_2$, 10 mM spermidine, 50 mM NaCl); 2 µl 100 mM DTT; 20 u RNaisin ribonuclease inhibitor; 4 µl of

a solution of 2.5 mM each of ATP, GTP and UTP, 2.4 µl 100 µM CTP; linearized template DNA (0.2-1.0 mg/ml in water), 5 µl [α–^{32}P]CTP (50 µCi at 10 mCi/ml) and 1µl of either SP6 or T7 RNA polymerase (at 15-20 u/µl). The transcription reaction was incubated at 37-40^{o}C for 60 minutes. At the end of the synthesis reaction, 5 µl of sequencing stop mix was added and the RNA purified on a 6 M urea sequencing gel with ^{32}P end labelled pBR322/*HpaII* fragments as markers. At the end of the run, three dissimilar pieces of Whatman-3MM paper were soaked with a small amount of [α–^{32}P]CTP and placed asymmetrically on the gel. The was covered with Saran wrap, exposed to an X-ray film for 5-15 minutes and the film developed using a X-Ograph processor. The film was then orientated on the gel using the Whatman-3MM pieces of paper and the RNA bands excised, crushed in an microfuge tube and incubated overnight with 2 volumes of elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, pH8.0, 0.1% SDS) with shaking. The acrylamide was removed by centrifuging through siliconized glass wool and the RNA solution phenol:chloroform extracted twice, ethanol precipitated and dried before resuspension in 100 µl DEPC treated sterile water. The specific activity of the RNA was determined by liquid scintillation counting and adjusted to 10,000 counts/µl for RNA bandshifts.

2.2.3.4 Cold RNA Synthesis

In the large scale synthesis of cold RNA, the following components were mixed in the order shown: 20 µl 5 x Transcription buffer; 10 µl 100 mM DTT; 100 u rRNasin ribonuclease inhibitor; 20 µl of a solution of each of ATP, CTP, GTP, and UTP at 2.5 mM

each; 2 µl of linearized template DNA (2-5 µg in water); 2 µl of SP6 or T7 RNA polymerase (at 15-20 u/µl) and the volume made up to 100 µl with sterile nuclease free water. The reaction mixture was incubated at 37°C for 60-120 minutes and the RNA transcripts either purified on a denaturing gel as described above or treated with RQ 1 RNAase-free DNAase to remove the template DNA, phenol:chloroform extracted twice and ethanol precipitated as described above.

2.2.3.5 RNA gel shift assays

The RNA/protein complex formation was assayed by running reaction mixtures on non-denaturing polyacrylamide gels. Samples were set up by preincubating 1.5 µl of 10 X buffer (1.5 mM Tris-HCl, pH 8.0, 50 mM NaCl and 50 mM KCl); 40 u rRNasin ribonuclease inhibitor and 0.5-8 µl protein extract for 10 minutes at room temperature followed by addition of 40-120 µg total yeast RNA, 10,000 counts of probe RNA and DEPC treated sterile water to a total of 15 µl. The reaction was incubated for 15 minutes at room temperature and 3 µl of sterile glycerol added before separation at 4°C on a 5% native polyacrylamide gel preelectrophoresed for 30-60 minutes at 35 mA constant current in 1% TBE.

The 5% polyacrylamide gel for the resolution of complexes was run on a BioRad electrophoresis apparatus and contained: 16.6 ml 30% acrylamide solution (20 g acrylamide, 1 g N,N'-methylenebisacrylamide and deionised water to 100 ml); 62.7 ml deionised water and 20.0 ml 5 x TBE (0.445 M Tris-borate, 2 mM EDTA, pH 8.0). The solution was degassed followed by addition of 0.7 ml 10% ammonium persulphate and 35

µl TEMED and poured.

The gel was exposed without drying in an X-ray film cassette at -70°C with an intensifying screen overnight and the film developed in an X-Ograph developer.

2.2.3.6 5'-End labelling of RNA

In vitro synthesized cold RNA was dephosphorylated using CIAP as described for DNA above except that, 0.01 unit of CIAP per pmole of 5' termini was used and the incubation was carried out for 15 minutes at 37°C followed by a 30 minute incubation at 55°C. Phosphorylation was carried out using bacteriophage PNK as described for DNA above.

2.2.4 Protein methods

2.2.4.1 Expression using the T7 RNA polymerase/promoter system

The *amiR* gene was cloned downstream of the T7 promoter sequence and Shine-Dalgarno sequence in plasmid pT7.7 (Fig 4.4). Plasmid pMW51 was then transformed into *E. coli* C600 carrying the T7 RNA polymerase gene under the control of λp_L promoter (plasmid pGP1-2) (Tabor and Richardson, 1985) and protein expression was carried out as described by Tabor (1990). Transformants carrying both plasmids were plated on LB/ampicillin/kanamycin plates and grown overnight at 30°C. A single colony was picked and inoculated into 5 ml Terrific broth/ampicillin/kanamycin medium and grown overnight at 30°C. The overnight culture was diluted 1:40 into fresh Terrific broth/ampicillin/kanamycin medium and grown for several hours at 30°C to an OD_{450} of 0.5 and the temperature raised quickly to 42°C for 30 minutes to induce the gene under pT7 control. The temperature was reduced to 37°C and the cells grown for an additional

90 minutes with shaking. The cells were chilled on ice for 10 minutes then centrifuged at 4°C for 5 minutes in a Sorvall GS-3 rotor at 5000 rpm. The supernatant was discarded, the cells resuspended in ice cold TES buffer (6.055 g Tris-HCl, pH8.0, 1.861 g EDTA, 2.922 g NaCl) and harvested as above. The cells were resuspended in 1/50 of original culture volume in cold lysis buffer (20 mM Tris-HCl, pH 8.0, 1 mM DTT, 0.5 mM PMSF, 5 mM EDTA). The tube containing the cells was put on ice and disrupted by sonication in an MSE Soniprep for ten 15 second bursts at amplitude 14 with 15 second cooling intervals. The cell debris was removed by centrifugation at 15,000 rpm for 15 minutes at 4°C and the supernatant stored at -70°C in 15% sterile glycerol.

Plasmids pMW40 and pMW41 which have the *amiR* gene cloned downstream of a T7 promoter were transformed into *E .coli* JM109(DE3) which carries a single T7 RNA polymerase gene in the chromosome under the control of a *lac*UV5 promoter (Studier and Moffat, 1986). The transformants for protein expression were aerated at 37°C to an $O.D_{450}$ of between 0.4-0.5 and induced with 0.4 mM IPTG. The cells were harvested after 3 hours and cell free extracts prepared as described above.

2.2.4.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out by the discontinuous system described by Laemmli (1970) using a Biorad protein electrophoresis or Miniprotean II gel electrophoresis apparatus as recommended by the manufacturers. A 15% SDS-polyacrylamide resolving gel (minigel) was prepared by mixing 2.5 ml 4 x resolving gel buffer (1.5 M Tris-HCl, pH 8.8, 0.4% SDS), 2.5 ml deionised water, 5.0 acrylamide mix (30% acrylamide, 0.8% bisacrylamide in deionised water), 0.05 ml 10% APS and 5 µl TEMED. The gel casting mixture was poured into the mould leaving 2 cm height for the stacking gel and layered with water

saturated n-butanol. After polymerisation of the resolving gel, the n-butanol overlay was poured off and 4.5% stacking gel mixture prepared as follows poured in: 1.25 ml 4 x stacking gel buffer (0.5 M Tris-HCl, 0.4% SDS, pH 6.8), 3.0 ml deionised water, 0.75 ml acrylamide mix (above) , 0.1 ml 10% APS and 10 µl TEMED.

The protein samples together with molecular weight markers were diluted in 2 x loading buffer (25% 4 x stacking gel buffer, 2% SDS, 5% β–mercaptoethanol, 20% glycerol and 0.0025% bromophenol blue) and boiled for 5 minutes prior to loading. Electrophoresis was carried out in Tris-glycine electrophoresis buffer diluted from 4 x stock (0.1 M Tris, pH 8.3, 0.77 M glycine, 0.4% SDS) at 35 mA constant current until the dye front reached the bottom of the gel. The gel was stained by immersing in at least 5 volumes of Coomassie stain (prepared by dissolving 0.25 g Coomassie Brilliant Blue R 250 in 90 ml of 1:1 v/v methanol:water mixture and adding 10 ml of glacial acetic acid) and shaking on a rotating platform for 4 hours at room temperature. The stain was removed and the gel destained by soaking in the methanol/acetic acid solution without dye (above) on a slowly rotating platform for 4-8 hours with three or four changes of destaining solution.

2.2.4.3 Immunoblotting

Proteins were electrophoretically transferred from SDS-polyacrylamide gels to nitrocellulose filters (0.45 µm pore size) for detection with antibody by western blotting (Towbin *et al.*, 1979; Burnette, 1981) using a NovaBlot Kit together with a Multiphor II electrophoresis system (LKB). A layer of six Whatman-3MM filter papers was soaked in Anode Solution 1 (0.3 M Tris, pH 10.4, 20% v/v methanol in distilled water) and placed on the anode electrode plate. Another three filter papers were soaked in Anode Solution 2

(25 mM Tris, pH 10.4, 20% v/v methanol in distilled water) and placed on top of the first six filter papers taking care not to trap any air bubbles. The membrane filter was soaked in the same electrode solution and carefully placed on top of the layer of filter papers. The gel was slowly lowered onto the immobilizing membrane and any trapped air bubbles removed by wetting the surface of the gel with electrode buffer and gently pushing them out. Three filter papers were soaked with cathode solution (40 mM 6-amino-n-hexanoic acid, pH 7.6, 20% v/v methanol in distilled water) and carefully placed on top of the gel. A further six filter papers were soaked in cathode solution and these made the contact with the electrode. The NovaBlot Kit was assembled and the transfer carried out at 100 mA for 1 hour.

The blots were blocked after the transfer by incubating with gentle agitation in blocking solution (2% w/v nonfat dried milk in TBS pH 7.2 [150 mM NaCl, 25 mM Tris-Cl]) for 20 minutes. The filter was transferred into fresh blocking solution containing suitably diluted rabbit antisera (1:1000) raised against MalE-AmiR fusion protein and incubated with shaking at room temperature for up to 18 hours. The primary antibody solution was discarded and the filter was washed three times for 10 minutes in TBS (Tris buffered saline, pH7.2) and immediately incubated at room temperature with fresh blocking solution containing horseradish labelled goat anti-rabbit IgG (Sigma) diluted 1:2000 in TBS for 2 hours. The filter was washed three times by 10 minute gentle agitations in a tray containing TBS (to cover the filter) and the colour developed by the addition of chloronapthol substrate (Sigma). Developed blots were washed in TBS and dried.

2.2.4.4 N-terminal sequencing

N-terminal sequencing was carried out on protein transferred onto 'Protoblot' membrane using NovaBlot Kit as described above. After transfer, the blots were stained with amido black for 2 minutes, destained in 40% methanol until the bands were visible with no

background and the band of interest cut out and sequenced. The first twenty amino acids were sequenced by Edman degradation using an Applied Biosystems 470A Gas Phase Sequencer with on line detection of amino acid phenylthiohydantoin derivatives using 120A HPLC (Applied Biosystems, Warrington U.K.).

2.2.5 Protein purification methods

2.2.5.1 Ammonium sulphate fractionation

A cell free extract prepared from C600, pGP1-2, pMW51 as described above was placed in a beaker on ice and transferred to the cold room on a magnetic stirrer. The amount of solid ammonium sulphate required to give 20%, 30%, 40%, 50%, 60%, 70% and 80% saturation was determined from published tables (Harris and Angal ed. 1989). At each step ammonium sulphate was added slowly with stirring and then stirred for a further 1 hour. The resulting suspension was centrifuged at 15,000 rpm in a Sorvall SS-34 rotor at 4^{o}C and the pellet resuspended into cell lysis buffer (above). Any material that did not go into solution was removed by centrifugation. The supernatant was taken to the next ammonium sulphate fractionation step and the procedure repeated. The fractionation was followed by SDS-PAGE electrophoresis as described above.

2.2.5.2 Ion exchange chromatography

Ion exchange chromatography was carried out using an FPLC system III machine (Pharmacia). 200 µl (17.3 µg protein) of clarified crude extract (prepared as above) was applied to a MonoQHR 5/5 ion exchange column prewashed with elution buffer (20 mM Tris-Cl, pH8.0, 1 mM DTT, 1mM EDTA). The column was eluted with a linear gradient of 0-1 M NaCl at a flow rate of 1 ml/min. Fractions were collected with a Frac-100

fraction collector (Pharmacia) and stored at 4°C for prior to analysis by SDS-PAGE.

DEAE cellulose ion exchange chromatography was carried out using a DE52 ion exchange column washed with elution buffer (20 mM Tris-Cl, pH8.0, 1 mM DTT, 1mM EDTA) and eluted with a linear gradient of 0-1 M NaCl. The fractions were collected as above and analyzed by SDS-PAGE.

2.2.5.3 Gel filtration

Gel filtration was carried out using the FPLC system described above. A clarified fraction from the 20% ammonium sulphate fraction that contained AmiR was loaded onto a Superdex 200 FPLC column. The column was eluted with EB (20 mM Tris-Cl, pH8.0, 1 mM DTT, 1mM EDTA), and 2 ml fractions collected for analysis by SDS-PAGE and stained with Coomassie blue or by Western blotting.

CHAPTER 3

Studies on Amidase induction

3.1 Introduction

Initial studies on the regulation of amidase expression in *Ps. aeruginosa* were performed by Kelly and Clarke (1962) in the wake of the classic Jacob and Monod (1961) *lac* operon work. They observed that N-substituted amides gave gratuitous induction, and different substitutions or carbon chains could eliminate induction implying a size exclusion at the inducer pocket. In addition thioacetamide inhibited induction but had no effect on the catalytic activity of the enzyme and other amide derivatives had varying activities as substrates without correlation with their inducer/inhibitor activity implying distinct inducer and substrate binding sites. A large number of regulatory mutants were later isolated including constitutives which were classified as magnoconstitutive if they produced high levels of amidase all the time, and semiconstitutive if they could be induced to higher levels.

Investigations with cloned amidase genes from PAC433 and PAC1 have identified *amiE*, the structural gene (Brammar *et al.*, 1987); *amiR*, the positive regulator (Lowe *et al.*,1989); *amiC*, the negative regulator (Wilson and Drew, 1991) and the open reading frames *amiB* and *amiS* (Wilson, 1991). An earlier study of the insertion and replication of the *Pseudomonas* mutator phage D3112 used the amidase operon of strain PAO1 because of the well characterised genetic system. Amidase negative mutants were selected on fluoracetamide plates and Southern blotting used to map the site of the insertions (Rehmat

and Shapiro, 1983). Insertions into *amiE* or *amiR* produced amidase negative phenotypes and insertions into what has now been identified as the *amiB* open reading frame produced a 'leaky' amidase negative phenotype. To investigate the role of *amiB* in amidase expression an *amiB* deletion mutant was constructed and analyzed.

3.2 Role of *amiB* in Amidase Induction

3.2.1 Studies on *amiB*- Phenotype in *E.coli*

E.coli contains no amidase activity and thus studies with cloned amidase genes are free of extraneous activity. The wild type amidase genes had been cloned on a 5.3 kb *Hin*dIII/*Sal*GI DNA fragment into pBR322 to make plasmid pAS20 (Fig 3.1) (Wilson and Drew, 1991) and the *amiB* gene identified and sequenced (Wilson, 1991) by the time I commenced the project. Two 'unique' *Apa*I targets were identified 423bp apart within *amiB* which provided a convenient way for making a deletion in plasmid pAS20 which could then be assessed for any change in phenotype.

Initial attempts to digest plasmid pAS20 isolated from a methylase positive *E.coli* strain (JA221) with *Apa*I resulted in linear molecules. The *Apa*I restriction refractory site was identified by fine restriction mapping with *Apa*I and *Pvu*II as position 1722 (the upstream site) which is 1bp downstream of a *dcm* methylase recognition site. Plasmid pAS20 was transformed into *E.coli* GM2163 (*dcm*-6, *dam*-13). A single transformant was grown up in 100 ml NB O/N and plasmid DNA isolated by the PEG precipitation method described earlier. Subsequent *Apa*I digestion went to completion as shown by AGE confirming the earlier suspicion that methylation of the juxtaposed *dcm* site inhibited the enzyme.

About 2 μg of plasmid pAS20 DNA was digested with *Apa*I and the 8.7 kb fragment was

isolated from the agarose gel, ligated and transformed into *E.coli* JA221 to Ampicillin resistance. The desired plasmid pMW1 (Fig 3.1) was identified from resistant colonies by MPP and restriction enzyme analysis with *Hin*dIII and *Xho*I.

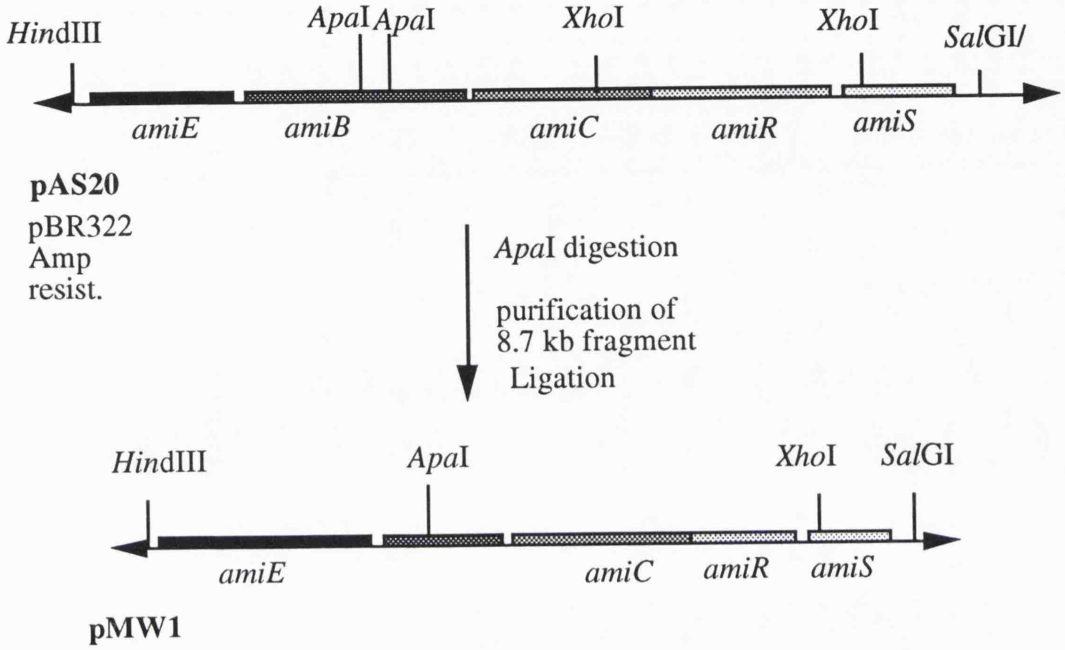


Fig 3.1 Construction of plasmid pMW1. Plasmid pAS20 was digested with *Apa*I and ligated.

Expression of amidase in *E.coli* carrying plasmids pMW1, pAS20 and pJB950 was measured using whole cells grown in minimal medium under inducing, non-inducing and repressing conditions with either glucose or lactate as a carbon source lactamide was used as the inducing amide and butyramide repression was measured with lactate as the carbon source. In addition experiments were carried out under nitrogen excess and nitrogen limiting conditions. The results are shown in Table 3.1A. The results for pAS20 and pJB950 were as reported previously (Wilson, 1991). Under nitrogen excess conditions pAS20 is devoid of amidase activity but under nitrogen limiting conditions shows inducible

amidase expression. Under nitrogen excess conditions pJB950 shows low constitutive amidase activity (average 3.7 u) which increases dramatically under N-limiting conditions to a high constitutive but further inducible level (average 38.0 u). Plasmid pMW1 carrying the *amiB* deletion shows similar results to pAS20 with very low constitutive expression for cells grown in excess nitrogen and low inducible expression under N-limition. The results for pAS20 and pMW1 were not identical but no real change was seen with the *amiB* mutation in *E. coli*. Poor heterologous amidase expression in *E. coli* has been reported previously (Cousens *et al* 1987; Drew and Lowe, 1989) but the differences seen are fortunately only in magnitude and not principle as most of the studies have been carried out in this organism. In fact *amiC* mediated repression and nitrogen limitation activation of amidase expression were first observed in *E.coli* (Wilson, 1991) as well as localization of the *amiR* coding region (Cousens *et al.*,1987). Since very high inducible amidase activity has been reported previously with the PAC1 gene fragment in *Ps. aeruginosa* (Wilson and Drew, 1991) it was decided to investigate the *amiB* phenotype in *Ps. aeruginosa* because of the higher amidase expression in the original host.

3.2.2 Amidase expression from cloned *amiB*- wild type genes in *Ps aeruginosa*

Plasmid pMW1 (above) is based on pBR322 which carries pMB9 origin of replication and cannot replicate in *Ps. aeruginosa*. To assessthe effect of the *amiB* 423bp *Apa*I deletion in *Ps. aeruginosa* a further subcloning was required. The 4.8 kb *Hin*dIII/*Sal*GI DNA fragment was excised from plasmid pMW1 and cloned into the *Hin*dIII/*Xho*I sitesof the broad host range vector pKT231 (Bagdsarian *et al.*, 1981) creating a *Sal*GI/*Xho*I hybrid site at the 3' end of the amidase gene fragment. Potential recombinants were transformed into *E.coli* DH1 and selected for streptomycin resistance. Plasmid DNA was isolated from the transformants by the MPP protocol described earlier and restriction enzyme mapped

with *Hin*dIII/*Xho*I and *Pvu*II to identify plasmid pMW22 (Fig 3.2).

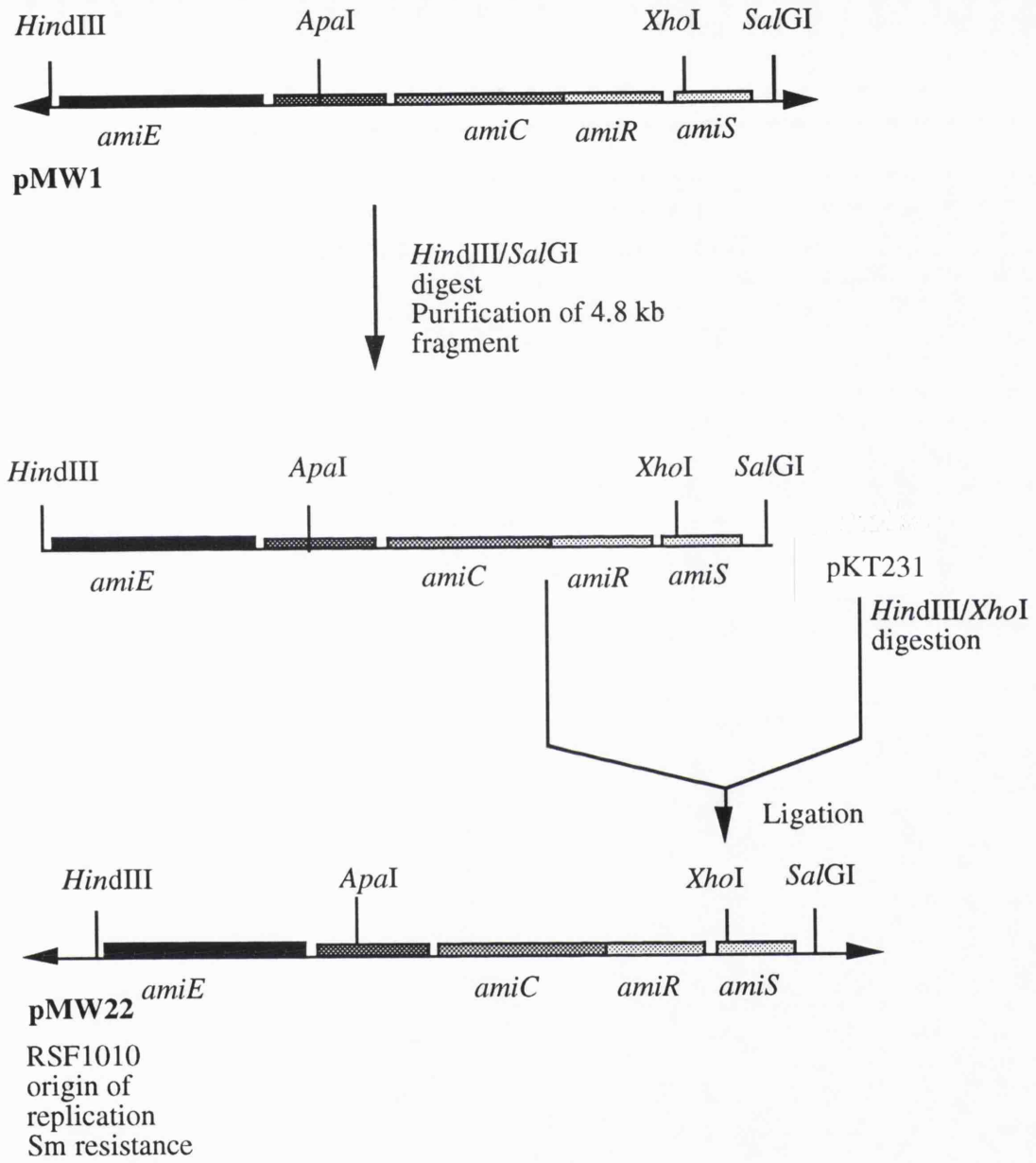


Fig 3.2 Construction of Plasmid pMW22.
The 4.8 kb amidase genes fragment from pMW1 was subcloned into pKT231

A.

Plasmid	Nitrogen +/-	Growth Conditions: Glucose	Glucose/ Lactamide	Succinate	Succinate/ Lactamide	Succinate/ Butyramide
pAS20	+	0	0.1	0	0	0
pAS20	-	0	6.2	0.4	4.9	0.1
pJB950	+	3.8	4.0	3.2	3.8	3.5
pJB950	-	21.6	43.8	24.7	61.0	38.6
pMW1	+	0.2	0.2	0.2	0.2	0.1
pMW1	-	0.6	1.3	0.5	0.9	0.4

Comparision of amidase activities between the constitutive amidase genes, the WT genes, and the *amiB* deletion mutant in *E. coli* JA221. The results are given as μmoles acetylhydroxamate/minute/mg bacteria.

B.

Plasmid	Growth Conditions: Succinate	Succinate Lactamide	Lactate	Lactate/ Lactamide	Lactate/ Butyramide
pMW22	1.0	54.8	0.5	96.1	0.5
pSW101	1.1	51.9	0.6	30.7	0.9

Amidase assays in *Ps. aeruginosa* PAC452pMW22 and PAC452pSW101. pSW101 data was obtained from Wilson (1991) The results are given as μmoles acetylhydroxamate/minute/mg bacteria

Table 3.1 Amidase assays of *amiB* deletion mutation in *E.coli* and *Ps. aeruginosa*

Plasmid pMW22 was mobilised into the *ami* deletion strain PAC452 (Drew, personal communication) by triparental mating using *E.coli* HB101 pRK2013 (Deretic *et al.*, 1986) as helper strain and *E.coli* DH1 pMW22 as the donor. Streptomycin resistant transconjugants were selected on *Pseudomonas* basal minimal medium plates (Brammar and Clarke, 1964) with citrate as the carbon source. This selects against *E.coli* which is unable to use citrate as an energy source and for the more nutritionally versatile *Ps. aeruginosa.*

Amidase assays were carried out using whole cells under inducing, non-inducing and repressing conditions. The results are shown in Table 3.1B alongside those obtained with PAC452 pSW101 which carries the isogenic *amiB+* fragment (Wilson and Drew, 1991). Low constitutive amidase expression is obtained in absence of inducer which rises to 54.8 units and 96.1 units with succinate and lactate respectively under inducing conditions and compares very well to the wild type genes. A similar low constitutive expression is also obtained with pSW101 (row 2) and is most likely a result of gene dosage caused by the copy number of the plasmid. This finding ruled out a role for *amiB* or its product in the induction process.

3.3 Coordinate expression of the regulator genes *amiC* and *amiR*

Previous studies had shown that *E. coli* pTM1 (*amiE*) expressed amidase at a very low constitutive level (0.7 u); *E. coli* pTM1, pTM2 (*amiR*) showed high constitutive amidase expression (33.3 u) and *E. coli* pTM1, pTM2, pSW41 (*amiC*) again showed a low constitutive (non-inducible) level of amidase expression (0.9 u). In plasmids pTM2 and pSW41 *amiR* and *amiC* were expressed from heterologous vector promoters. This ruled

out the possibility of transcriptional regulation of the amidase operon by AmiC. Previously reported complementation studies in *Ps. aeruginosa* showed that the WT strain PAC1 carrying an *amiR* expression vector (pSW40) expressed amidase constitutively and PAC101 (amidase constitutive) carrying the *amiC* expression vector pSW41 produced very low (non-inducible) constitutive levels of amidase activity. These studies indicated that AmiC regulation of amidase activity was likely to be by interaction with AmiR but was unlikely to be by covalent modification. From these two sets of results it appears that an alternative mode of AmiC amide dependent regulation of amidase expression might be protein/protein interaction between AmiC and AmiR and involves a stoichiometric relationship. In this case over expression of AmiR in PAC1 would produce inducer independent constitutive amidase expression; and over expression of AmiC in PAC101 would produce inducer independent repression of amidase expression. The *amiC* termination codon overlaps the *amiR* initiation codon, an arrangement normally seen with genes whose expression is linked (Dunn and Studier, 1983). There was then a real possibility that the ratio of the two gene products is tightly controlled and coordinate expression programs inducibility. To test for this possibility an expression system for the two genes was devised. The wild type*amiC/R* gene fragment was excised from plasmid pAS20 on a *Kpn*I/*Sal*GI fragment and initially cloned into plasmid pGEM3Z (Promega) to make plasmid pMW43 (Fig 3.3). Potential recombinants were transformed into *E.coli* JA221 and plated onto IPTG/X-gal/Ap NA plates. Plasmid DNA was isolated from white colonies followed by restriction enzyme mapping with *Hin*dIII/*Xho*I and *Hin*dIII/*Eco*RI to confirm presence of the insert. This step which introduces *Eco*RI and *Hin*dIII sites at both ends of the *amiC/R* DNA fragment was used as an alternative to blunt end cloning because the desired broad host range vector pMMB66EH (Morales *et al.*, 1990) does not have a *Kpn*I recognition sequence in the multiple cloning site.

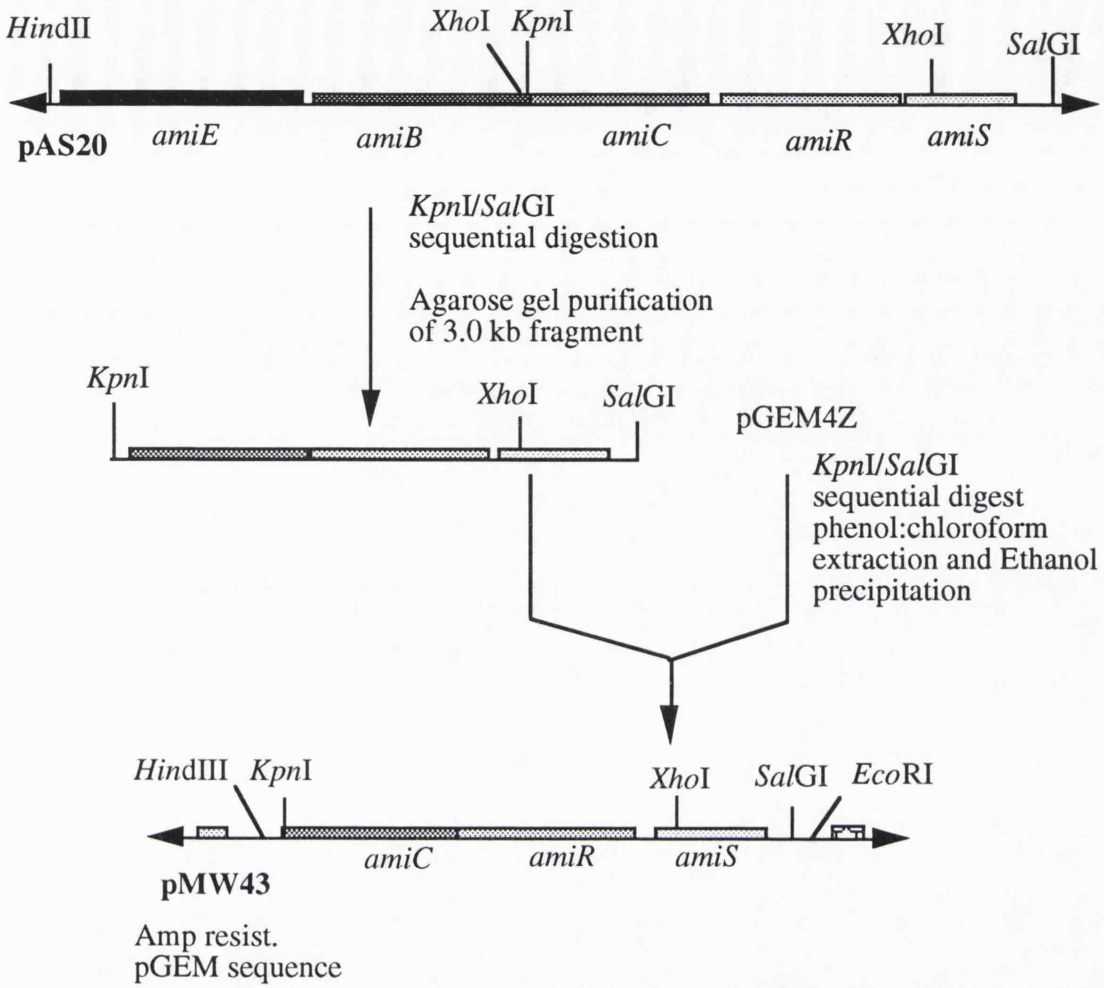


Fig 3.3 Construction of plasmid pMW43. The 3.0 kb *Kpn*I/*Sal*GI DNA fragment from pAS20 was cloned into pGEM4Z. Figure not drawn to scale.

Plasmid pMW43 was double digested with *Hin*dIII /*Eco*RI and the 3.0 kb insert purified on an agarose gel and ligated to *Hin*dIII/*Eco*RI cut pMMB66EH. The ligation mixture was transformed into competent *E.coli* JA221 and selected on Ap/NA plates. Plasmid pMW21 (Fig 3.4) was identified from the Ampicillin resistant colonies by restriction enzyme mapping MPP DNA with *Hin*dIII and *Xho*I.

amiC/R transcription in pMW21 is expected to start at the vector *tac* promoter which is in the same orientation as *amiC* in pSW41 (Wilson and Drew, 1991; Wilson *et al* 1993) Plasmid pMW21 was transformed into competent *E.coli* JA221pTM1 (*amiE*) and plated on Cm/Ap plates to select for both plasmids. The resistant colonies were subjected to MPP and the plasmid DNA digested with *Hin*dIII followed by AGE to confirm presence of both plasmids. Strain JA221 pTM1,pMW21 was streaked out to single colonies on a Cm/Ap/NA plate for use in amidase assays. The results of the assays are presented in Table 3.2a and were carried out both in the presence and absence of IPTG under the five growth conditions defined previously (Drew, 1984). In the absence of IPTG, low constitutive (0.8 units) expression was observed during growth on either glucose or succinate comparable to pTM1 alone (0.7 u). In the presence of the inducer, lactamide, expression rises to 5 and 6.4 units respectively. This reflects leaky transcription from the *tac* promoter even in the absence of its inducer. The assays performed in the presence of IPTG show the same trend albeit in a more dramatic way. Growth on glucose and succinate without inducer gave 3 and 3.9 units of activity respectively which is higher than above and implies that some active AmiR is present as long as the two genes are transcribed. Glucose/lactamide gave 22.3 units and succinate/lactamide 40.4 units of activity. These activities are of the same order of magnitude as obtained with cloned *amiE* and *amiR* (Wilson, 1991) and therefore reflect an amide dependent inactivation of *amiC* control over *amiR*.

Plasmid pMW21 was mobilised into *Ps. aeruginosa* strains PAC1 (wild type), PAC452 (*ami* deletion), PAC111 (ami^{con}) and PAC327 (*amiC*$^{-}$ *amiR*$^{-}$) by triparental mating using the helper strain HB101 pRK2103 as described earlier. Transconjugants were selected on *Pseudomonas* basal minimal medium plates with citrate as carbon source in the presence of

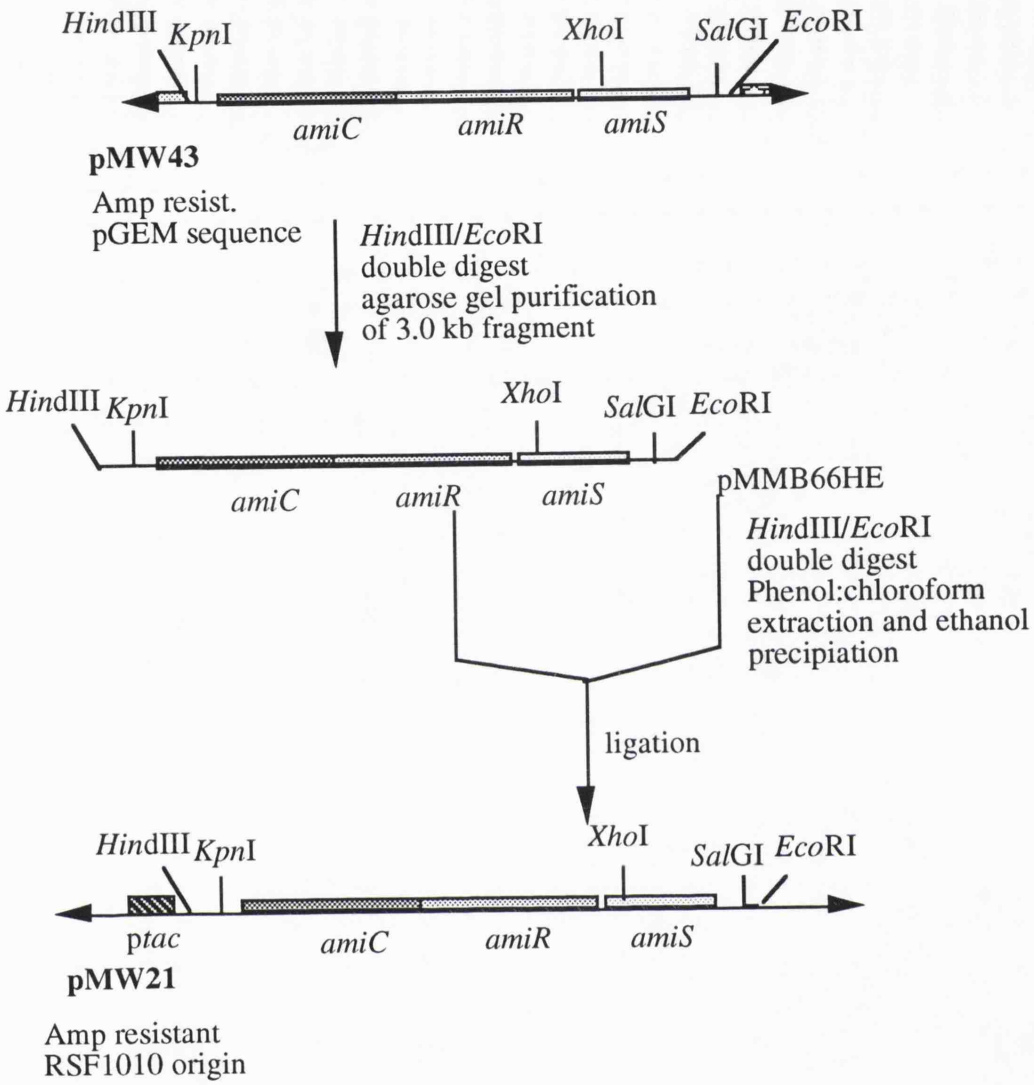


Fig 3.4 Construction of plasmid pMW21.
The 3.0 kb *amiC/R/S* DNA fragment from pAS20 was cloned into pGEM4Z to make pMW43 and then subcloned into pMMB66HE

Carbenicillin at 500 μg/ml. Plasmid DNA was prepared from the transconjugants and pMW21 transfer confirmed by restriction enzyme analysis with *Xho*I. The effect of pMW21 on the above strains except PAC452 was investigated by carrying out amidase assays under the five conditions described in the presence of IPTG (Table 3.2.b). PAC1

shows inducible amidase expression, PAC111 shows constitutive amidase expression which is sensitive to butyramide repression and PAC327 is amidase negative under all growth conditions. The presence of pMW21 in PAC1 showed no change in amidase activities measured except a higher background under non-inducing conditions. The presence of pMW21 in PAC111 was seen to convert the constitutive expression of the parental strain to inducible. Thus 0.7 u and 0.5 u were found with succinate and lactate rising to 3.2 and 9.4 units respectively in the presence of lactamide. PAC327 is an *ami*R- mutant derived from an original constitutive (*amiC*-) mutant strain PAC142. PAC327 pMW21 showed a mixed semiconstitutive phenotype. Low amidase expression was seen under all growth conditions (1.0-1.3 units) which was inducible by lactamide to 3.6 u (succinate) and 2.2 u (lactate).

In all of these assays the presence of a multicopy *amiC/amiR* expression vector has converted expression to an inducible phenotype. However the multiple copies of *amiC,amiR* present on the vector has not led to substantially more amidase being produced. Thus either the *trans* complementation is not very efficient or expression is being limited by the *amiE* promoter. To try and estimate the relative amounts of AmiC and AmiR expressed, PAC452 pMW21 was grown in LB at 37^{0}C O/N with vigorous shaking in the presence of 100 mM IPTG. The cells were harvested and cytosolic extracts made as described earlier. Fig 3.5 shows a Coomassie stained 15% SDS-polyacrylamide gel (Laemmli 1970) loaded with molecular weight markers in lane 1, a PAC452 extract in lane 2, PAC452 pMMB66EH extract in lane 3, IPTG induced PAC452 pMW21 extract in lane 4, and IPTG induced PAC452 pMW21 extract from cells grown in the presence of lactamide in lane 5. Lane 1 and 2 are identical, however, a new band running just below the 45 kDA marker (est. 43 kDA) is evident in lanes 4 and 5. This corresponds to the expected AmiC

band (Wilson and Drew, 1991) and no band that could be attributed to AmiR (22 kDa) is observed. This implies that AmiR is produced at considerably lower levels or it is highly labile. Overexpression studies of *amiR* are further described in Chapter 4.

		Growth Conditions				
Plasmids	IPTG +/-	Glucose	Glucose/ Lactamide	Succinate	Succinate/ Lactamide	Succinate/ Butyramide
pTM1,pMW21	-	0.8	5.0	0.8	6.4	0.3
pTM1,pMW21	+	3.0	22.3	3.9	40.4	0.5

Table 3.2 a Amidase activity of coordinately expressed *amiC/R* genes in *E. coli* DH1pTM1. The cultures were grown in *E. coli* minimal medium supplemented with thiamine. The results are given as µmoles acetylhydroxamate/minute/mg bacteria . pTM1carries the *amiE* gene, and pMW21 carries the *amiC/R* genes.

		Growth Conditions				
Strain	Plasmid	Succinate	Succinate/ Lactamide	Lactate	Lactate/ Lactamide	Lactate/ Butyramide
PAC1	-	0.1	8.1	0.1	2.5	0.0
PAC1	pMW21	0.3	3.5	0.3	7.1	0.3
PAC111	-	6.7	9.7	2.7	3.4	0.1
PAC111	pMW21	0.7	3.2	0.5	9.4	0.4
PAC327	-	0.0	0.0	0.0	0.0	0.0
PAC327	pMW21	1.0	3.6	1.3	2.2	0.7

Table 3.2 b Amidase activities of *Ps. aeruginosa* ; PAC1 (wild-type), PAC111 (*amiC*-,constitutive mutant), and PAC327 (*amiC*-, *amiR*-) mutant with and without pMW21. The results are given as µmoles acetylhydroxamate/minute/mg bacteria

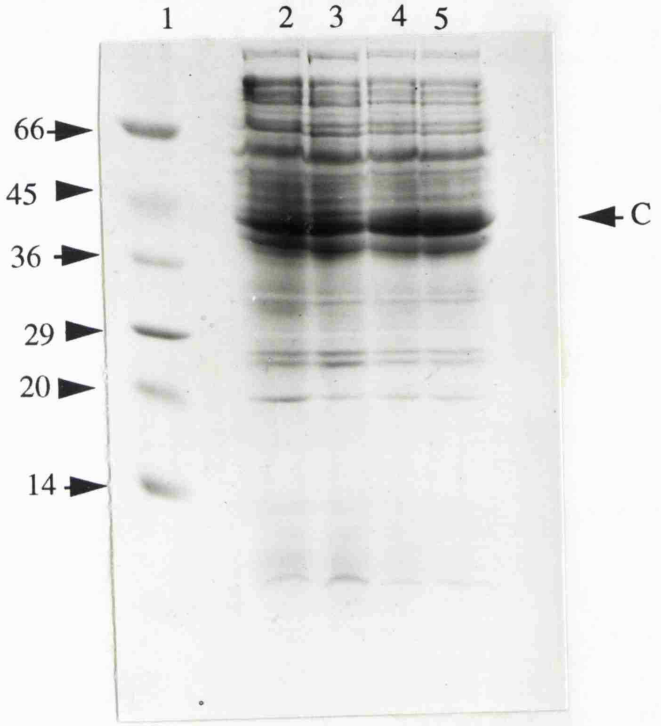
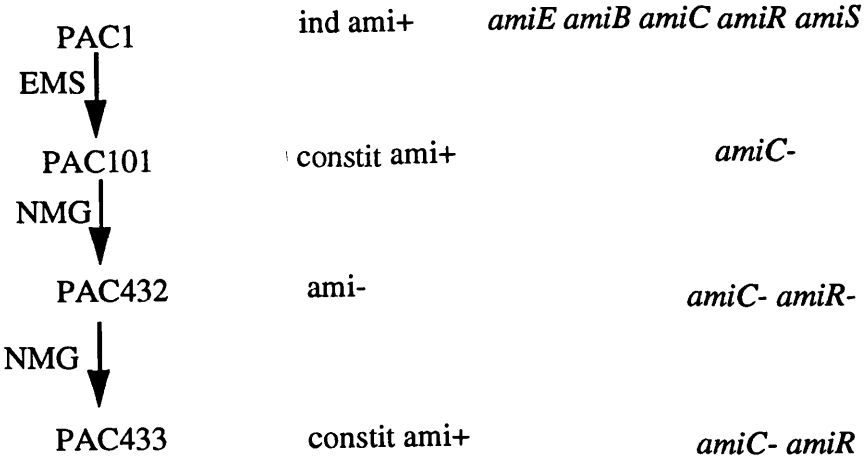


Fig 3.5 SDS-polyacrylamide gel of extracts of PAC452 pMW21.
Analysis of regulator gene products from plasmid pMW21. Whole cell extracts were separated on a 15% SDS-polyacrylamide gel and stained Coomassie blue. Lanes: 1) molecular weight markers expressed in kDa; 2) *Ps. aeruginosa* PAC 452 extract; 3)*Ps. aeruginosa* PAC 452 (pMMB66EH) extract ; 4)*Ps. aeruginosa* PAC 452(pMW21), 5 mM IPTG extract; 5) *Ps. aeruginosa* PAC 452(pMW21), 5 mMIPTG extract; 0.2 % lactamide. A diffuse new band is indicated by arrow C

3.4 Comparative studies of PAC1 and PAC433 regulator genes

Plasmid pJB950 (Clarke *et al.*, 1981) carries genes from the magnoconstitutive mutant PAC433 (Smyth and Clarke, 1974) and *E. coli* pJB950 express amidase at low constitutive levels (Cousens *et al.*, 1987). PAC433 was isolated in a series of steps from the wild type PAC1:



The nature of the *amiC* mutation in PAC101 has been investigated by trans complementation with the *amiC* expression vector pSW41 (Wilson and Drew, 1991). PAC101,pSW41 showed low constitutive, noninducible amidase expression under all growth conditions showing that the mutation to constitutive expression in PAC101 was in *amiC*. The second mutation to generate the *ami*- mutant PAC432 has been located to *amiR* by construction of PAC432 pDC35 (*amiR* expression vector) which shows low constitutive amidase expression (R. Drew, personal communication). The *ami*+ revertant PAC433 shows high AmiR dependent amidase expression and is thus *amiR*+ (Cousens *et al.*, 1987).

3.4.1 Transcomplementation Studies with PAC433 *amiR*

To test whether the PAC433 *amiR* gene could be repressed by the WT *amiC*, plasmid pJB950 was digested with *Xho*I and the 1.5 kb *Xho*I/*Xho*I gene fragment band excised from an agarose gel, purified and ligated to *Sal*GI linearized pBGS19-. The ligation mixture was used to transform *E. coli* JA221 to kanamycin resistance. Plasmid pMW30 (Fig 3.6) was identified from resistant colonies by MPP isolation of plasmid DNA followed by restriction enzyme mapping with *Hin*dIII and *Hin*dIII/*Bsc*I. The plasmid was transformed into *E. coli* JA221 pTM1 (*amiE*) and JA221 pTM1 pSW41 (*amiC*) and transformants selected on Cm/Km and Cm/Km/Ap NA plates followed by plasmid DNA isolation by MPP and restriction enzyme analysis to confirm the presence of the plasmids.

The newly constructed strains were grown and streaked to single colonies for the amidase assays shown in Table 3.3. Plasmid pTM1 gives a background activity of 0.7 units under all growth conditions (Wilson *et al.*, 1993). With pTM1, pMW30, constitutive amidase expression with an average of 20.2 units was obtained. The presence of *amiC* (pTM1, pMW30 and pSW41) abrogated amidase expression in a manner similar to pTM1 pTM2 (PAC1 *amiR*) pSW41 (Wilson, 1991) to give a basal 0.3 units and 1.1 units in absence and presence of IPTG respectively. This result confirms the location of the constitutive mutation in PAC433 to the *amiC* gene. The DNA sequence of the amidase operon available is a combination of PAC1 and PAC433 sequences (Appendix 1). The *amiE* leader region (*Hin*dIII to *Hin*dII) is from both PAC1 and PAC433 (Drew and Lowe, 1989); the *amiE* sequence is from PAC433 (Brammar *et al.*, 1987); the *amiB* sequence is from PAC1 (Wilson *et al.*, manuscript submitted); the *amiC* sequence is from PAC1 (Wilson and Drew, 1991); the *amiR* gene (including the C-terminal coding region of *amiC* and N-terminal region of *amiS*) is from PAC433 (Lowe *et al.*, 1989) and the *amiS* gene sequence is from

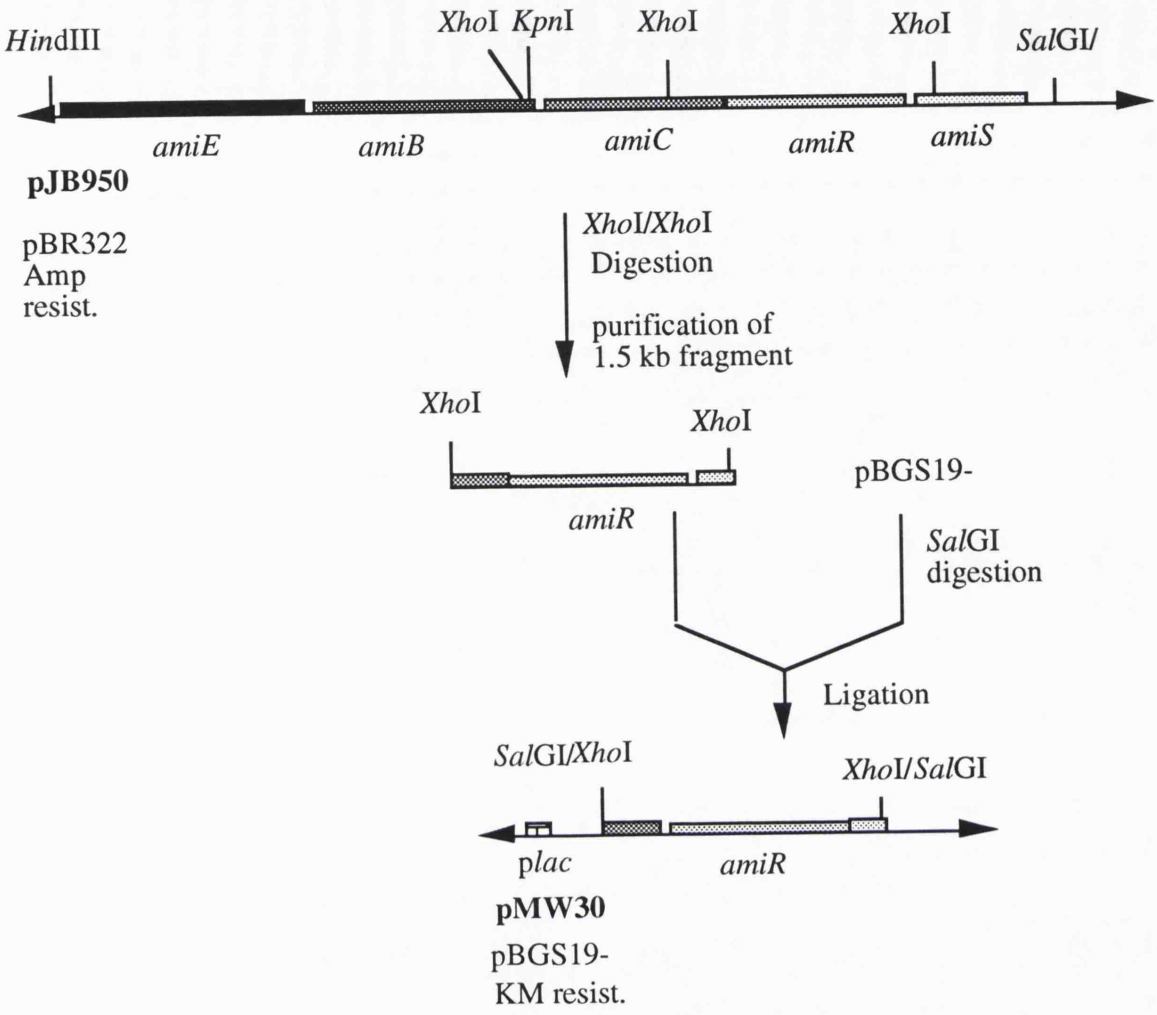


Fig 3.6 Construction of plasmid pMW30. The *amiR* gene from plasmid pJB950 was cloned into the *Sal*GI site of plasmid pBGS19- on a 1.5 kb *Xho*I DNA fragment

Plasmids	IPTG +/-	Growth Conditions				
		Glucose	Glucose/ Lactamide	Succinate	Succinate/ Lactamide	Succinate/ Butyramide
pTM1,pTM2	-	30.8	31.8	31.0	33.6	39.1
pTM1,pMW30	-	23.2	30.7	18.4	15.0	9.8
pTM1,pMW30	+	36.7	26.8	10.3	11.0	16.2
pTM1,pMW30, pSW41	-	0.2	0.4	0.2	0.2	0.3
pTM1,pMW30, pSW41	+	0.4	0.4	0.4	1.4	1.4

Table 3.3 Transcomplementation studies with cloned PAC433 (amidase constitutive) *amiR* gene in *E. coli* JA221. The results are given as μmoles acetylhydroxamate/minute/mg bacteria

PAC1 (Wilson *et al.*, manuscript submitted). Thus the complete DNA sequence of *amiC* from PAC1 is available and the C-terminal part of the gene from PAC433. Some changes have been identified between the sequences which may account for the constitutive phenotype (Wilson, 1991). To formally locate all of the changes to constitutivity, the N-terminal region of *amiC* from PAC433 has been sequenced.

The DNA sequencing of the *amiC* gene from PAC1 included running off the 3' end of the gene into *amiR*. Bearing in mind the three mutagenic steps between PAC1 and PAC433 and the fact that the parent of PAC433, PAC432, is *amiR*-, the remainder of the *amiR* gene from PAC1 was sequenced to compare with the PAC433 *amiR* sequence available.

3.4.2 DNA Sequence Analysis of the *amiC* gene from PAC433

To sequence the 652 *Xho*I/*Xho*I PAC433 *amiC* gene fragment, plasmid pJB950 was digested with *Xho*I and the fragment agarose gel purified before cloning into the *Sal*GI site of plasmid pUC19 (Fig 3.7). Two plasmids with opposite orientations were identified by restriction enzyme analysis with *Kpn*I whose restriction site is located asymmetrically within the insert. The two plasmids pMW650a and b were purified on CsCl gradients for double stranded sequencing. The plasmid DNA was denatured with NaOH and sequenced by the dideoxy chain termination method of Sanger *et al* (1977) using a T7 polymerase kit (Pharmacia). The sequence obtained was of limited quality and it was therefore decided to try single stranded sequencing. The 652bp insert from plasmid pMW650a was removed on a *Hin*dIII/*Eco*RI fragment and cloned into *Hin*dIII/*Eco*RI digested M13mp18 and 19. The ligation mixture was transformed into *E. coli* JM109 and plaques with insert identified histochemically as described elsewhere. Preparation of single stranded template from SJW 1/2 for sequencing was as described earlier and the concentration of DNA was estimated by AGE alongside M13mp18 SS of known concentration. Approximately 2 μg aliquots of these DNAs were subjected to dideoxy chain termination sequencing. The sequence obtained was identical to the PAC1 sequence and is shown in the Appendix (2389-3041) (Wilson and Drew, 1991). Thus the only change in *amiC* which lead to constitutive amidase expression is that identified previously (Wilson, 1991), the C to T at 3291 which causes a Pro to Leu change at residue 255 in AmiC.

3.4.3 DNA sequencing Studies of the *amiR* gene from PAC1

The strategy for sequencing the wild type *amiR* gene is shown on Fig 3.9. The 954 bp *Xho*I/*Bcl*I fragment was cloned into the *Sal*GI/*Bam*HI site of pUC19 to make plasmid pMW954. Plasmid DNA was isolated from JM109 pMW954 and Column purified for

sequencing as described above. This attempt only yielded 90 bp of clear sequence which did not cover the region of interest. To make SS templates, phages SJW3/4 were made by cloning the 381 bp *Bsc*I/*Eco*RI fragment from pMW954 into the *Acc*I/*Eco*RI sites of M13mp18/19 respectively and recombinants identified by AGE of SS DNA by running M13mp18/19 control alongside.

SJW5/6 were made by isolating the 476 bp *Bcl*I fragment from plasmid pAS20 (Fig 3.9) and cloning into first the *Bam*HI site of plasmid pUC19 and then into M13mp18/19 on a *Hin*dIII/*Eco*RI fragment. Phages SJW7/8/9/10 were made by digesting plasmid pMW40 (Chapter 4) with *Rsa*I and purifying the fragments on 12% agarose gel before cloning into *Sma*I linearized M13mp19.

The sequence obtained is identical to that obtained for the PAC433 *amiR*. This result shows that the revertant PAC433 carries an identical *amiR* gene to the wild type and the uncharacterised *amiR* mutation in PAC432 has changed back to the wild type sequence.

3.5 Confirmation of the *amiS* ORF

The presence of a potential ORF downstream of the *amiR* gene was first recognized at the time of sequencing the 1.515 *Xho*I/*Xho*I *amiR* gene fragment (Lowe *et al* 1989). The sequence between the downstream *Xho*I and the *Sal*GI site that delimits the 3' end of the amidase genes was subsequently determined for the non coding strand and partially for the complementary strand (Wilson 1991). Analysis of the deduced amino acid sequence revealed a highly hydrophobic protein, *amiS*, with six potential transmembrane helices and

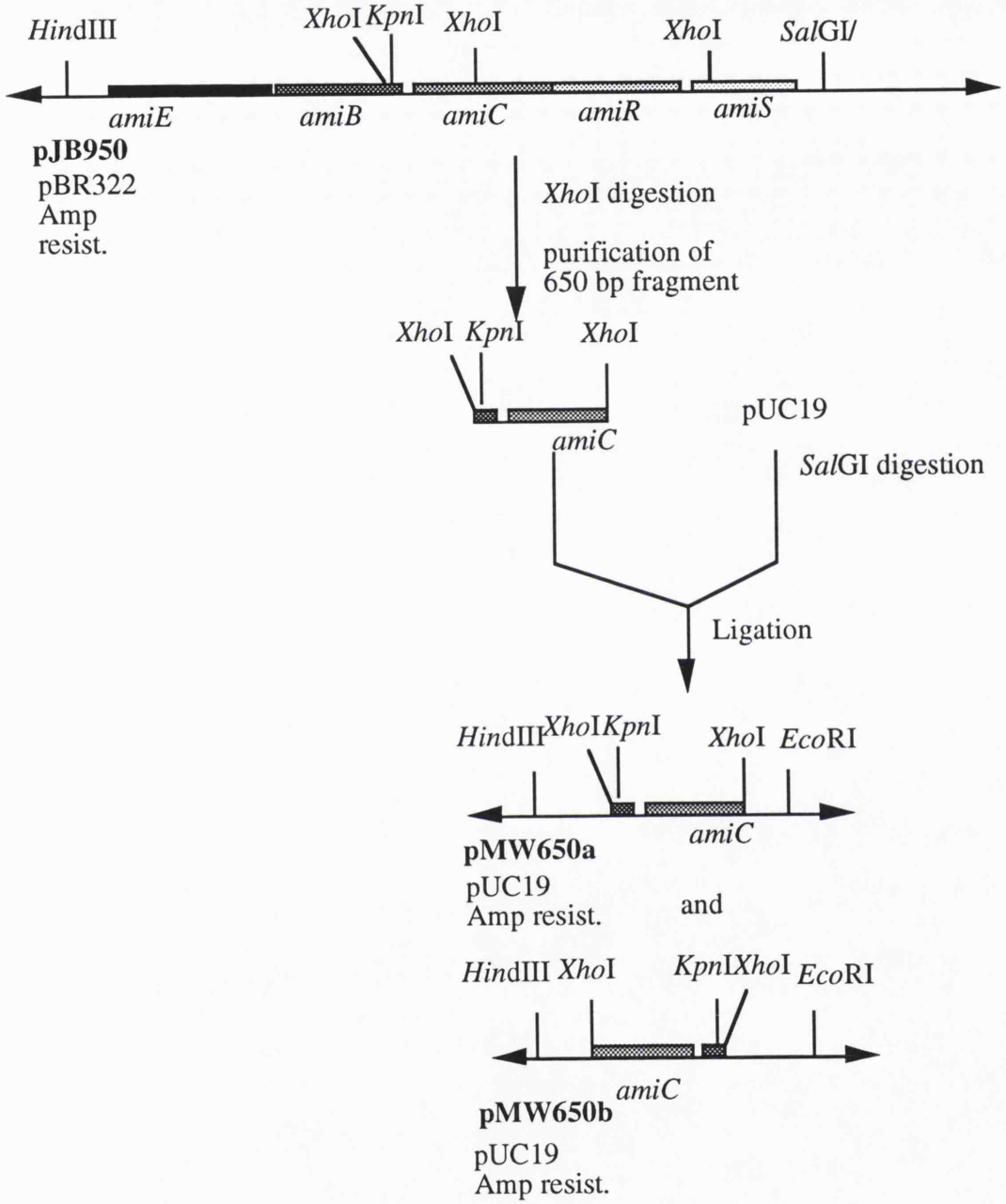


Fig 3.7 Construction of sequencing templates for the 650 bp *XhoI*/*Xho*I PAC433 fragment.The orientation of the inserts in plasmids pMW650a and pMW650b was determined by *Kpn*I digestion.

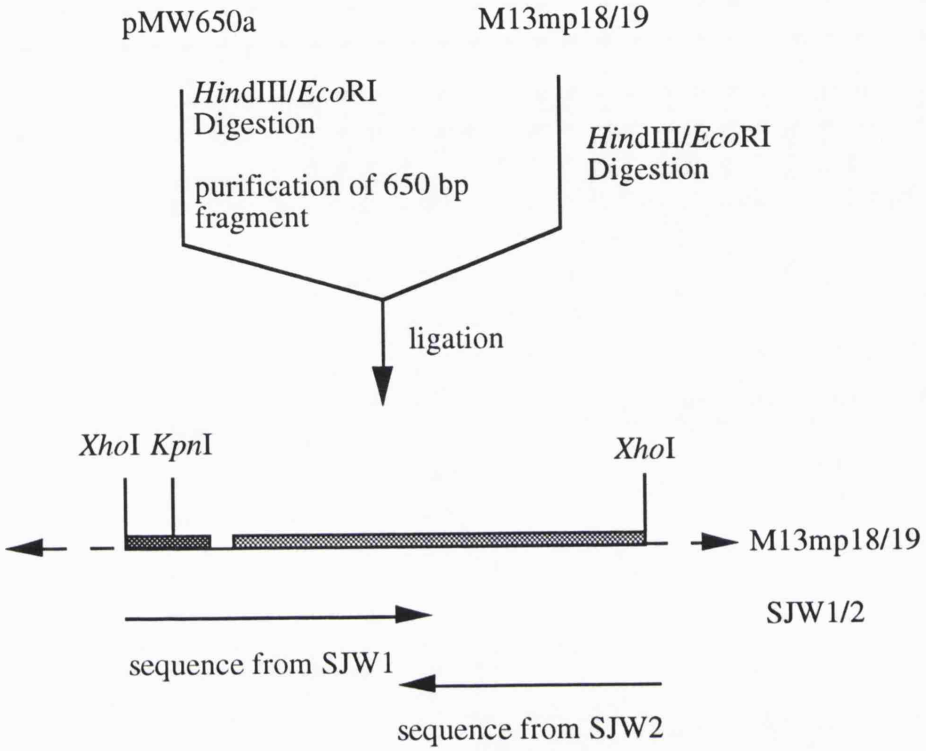


Fig 3.8 Sequencing strategy for *Xho*I/*Xho*I DNA fragment from PAC433.
Single stranded templates were used due to limited quality of sequence obtained with double stranded templates. Figure not drawn to scale

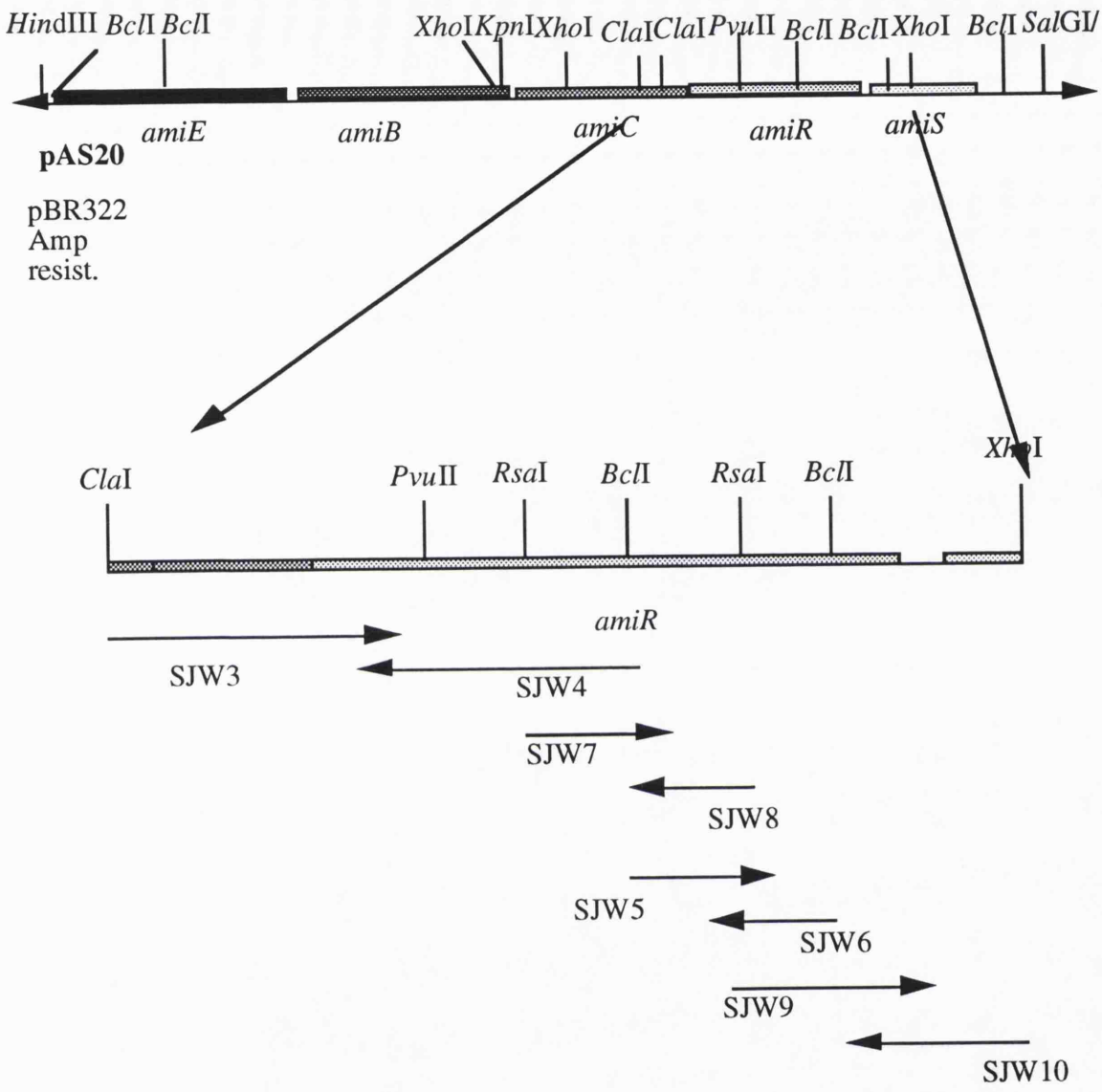


Fig 3.9 Sequencing strategy for the wild type *amiR* gene using M13mp18/19 recombinant phages. Details for the construction of the sequencing templates are given in the text. Figure not drawn to scale

in this one respect, *amiS* product resembles the integral membrane component of a periplasmic-binding protein-dependent transport systems (Tam and Saier, 1993). Its proposed function being a component of an amide uptake system (Wilson, 1991; Drew and Wilson, 1992). Many of this family of proteins also participate in transmitting signals from their cognate periplasmic binding proteins to cytoplasmic regulator proteins which in turn effects gene expression (Tam and Saier, 1993).

To confirm the carboxyl terminal coding region for *amiS*, plasmid pAS20 (Fig 1.11) purified on a CsCl gradient was denatured with NaOH and annealed to a synthetic oligonucleotide which is complementary to 4679 to 4889 bp region. Sequencing was carried out using the dideoxy chain termination method of Sanger *et al* (1977) using the sequenase version 2.0 kit of USB as described. The sequence obtained confirmed the 4665 to 4883 bp region of the *amiS* ORF as reported (Wilson, 1991) and is shown in the Appendix.

3.6 Summary

Work presented in this chapter has unequivocally ruled out the possibility of *amiB* being involved in the amidase induction process on two counts. Firstly, the *amiB* deletion construct plasmid pMW22 shows an amidase inducible phenotype indistinguishable from a similar wild type construct plasmid, pSW101 (Wilson and Drew, 1991). Secondly, plasmid pMW21 which carries the *KpnI/Sal*GI fragment from strain PACI and as such does not carry the *amiB* sequence confers an amidase inducible phenotype to the *amiE* gene in *trans-* to plasmid pTMI, in *E. coli*, and to the amidase constitutive strain PAC111 and amidase negative strain PAC327 in *Pseudomonas.*

Sequence analysis of the *amiB* gene product has shown a potential ATP binding motif (Walker *et al.*, 1983) and it is possible that it functions as an energy transducing component of an active amide transport system of the ATP binding cassette-type (ABC-type) (Higgins *et al.*, 1986) (Drew and Wilson, 1992 ; Wilson, 1991; Wilson and others, manuscript submitted). Whereas these experiments did not shed any new light to this proposal, they suggested that passive diffusion or non specific transport of amides into the cell triggers induction of the amidase operon and that regulation of the amidase operon is unlinked to transport.

AmiC shows homology both in sequence and structure to the periplasmic branched chain amino acids binding proteins (LIVJ) of *E. coli* and *C. freundii* (Tam and Saier, 1993; Wilson *et al.*,1993), but lacks the N-terminal secretory signal peptide which is consistent with its cytoplasmic localization and function as an amide dependent repressor of the amidase operon (Wilson *et al.*,1993). The mechanistic details of the AmiC function have not been fully characterised, however an *in vitro* AmiR/leader RNA gel shift assay in the presence of; (i) AmiC, (ii) AmiC and inducing amide acetamide, and (iii) AmiC with repressing amide butyramide did not show any effect of AmiC on the RNA binding activity of AmiR (data not presented).

DNA sequencing studies of the wild type *amiR* gene and the 650 bp *Xho*I/*Xho*I PAC433 *amiC* fragment has confirmed that the constitutive phenotype in PAC433 is as result of the carboxyl terminal *amiC* mutation described previously (Wilson, 1991) and that the *amiR* gene is identical to the wild type. The results agree with the PAC433 *amiR* (plasmid pMW30) phenotype in the presence and absence of *amiC* (Table 3.3).

CHAPTER 4

AmiR expression and purification studies

4.1 Introduction

One of the long term aims of this investigation was to study the mechanism of the transcription antitermination process (Chapter 5). An important aspect of this study would be the purification and characterization of AmiR for *in vitro* experiments. However, this type of study is frequently difficult because many regulatory proteins are normally produced in low levels, assays are often difficult which means that the monitoring of active fractions during purification is costly and overproduction of some regulatory proteins can lead to loss of cell viability.

Initial attempts to identify AmiR in crude cell free extracts of various PAC strains were all unsuccessful (Farin, 1976). Three types of investigations were performed.

i) Equilibrium dialysis of phosphocellulose column fractions of crude extracts of PAC1 which bound ^{14}C acetamide. This initially showed encouraging results but the pooling of the amide binding fractions and testing by equilibrium dialysis was unsuccessful.

ii) Dual isotope labelling studies of PAC1 and PAC453 (*ami* Δ) with ^{3}H leucine and ^{14}C leucine. Many extra peaks and troughs were observed when the ^{3}H/^{14}C ratio of each fraction was determined.

iii) Dual isotope labelling studies of PAC438 *ami*R^{ts} grown at 25^{o}C and heat shocked to 42^{o}C for 10 minutes. Cells were mixed broken and fractionated on a phosphocellulose P11 column. Differences were seen between the heat shocked and non-

heat shocked extracts but nothing that could be identified as AmiR was found.

These experiments were designed on the premise that AmiR would be inducible and that it bound acetamide and DNA in a mechanism analogous to the allosteric modulation of the *E. coli lac* repressor or AraC in its regulatory function. It was not until the amidase genes were cloned (Drew *et al.*,1980) , the *amiR* locus determined (Cousens *et al.*, 1987) and sequenced (Lowe *et al.*,1989) that a protein, of about 23kDa the predicted size of AmiR, was visualized on an SDS - polyacrylamide gel as a product of an *in vitro* transcription/translation reaction (Cousens, 1985).

The *amiR* gene was located to the 1.5 kb *XhoI/XhoI* DNA fragment and this and smaller derivatives were used in subsequent studies. The 1.5kb *XhoI/XhoI amiR* gene fragment was cloned into the high copy number vector pUC19 for expression in *E. coli* (Wilson 1991). This construct pSW24 and the broad host range pMMB66HE (Bagdasarian *et al.*,1981) derivative pSW40 carrying the same insert showed AmiR activity in *trans* complementation studies but did not express amounts of AmiR detectable on SDS - polyacrylamide gels in either *E. coli* or *Ps.aeruginosa* (Wilson, 1991). It is of interest to note that the negative regulator, AmiC was expressed from both vectors as a major intracellular protein in *E. coli* and *Ps. aeruginosa* (Wilson and Drew 1991).

4.2 Protein expression from the T7 promoter

The following sections describe attempts to over-produce AmiR using systems that are based on T7 RNA polymerase expression systems. Bacteriophages T7, T3, T4 and SP6 have a related family of RNA polymerases that are highly specific for their respective promoters (Sousa et al., 1992). The T7 enzyme is a single 91 kDa polypeptide. It initiates RNA

synthesis very efficiently and has an elongation rate 5 times greater than that of the *E. coli* RNA polymerase (Studier and Moffat, 1986). The advantage of T7 systems is that the T7 promoter is not recognized by host RNA polymerases, and the T7 RNA polymerase transcribes only the cognate promoter sequence and not from the host promoters (Studier and Moffat, 1986). Thus custom made T7 expression systems have been developed (Studier and Moffat, 1986; Tabor and Richardson, 1985).

Use of these T7 expression systems involve: cloning of the gene of interest downstream of a T7 promoter carrying the appropriate translation signals, and controlled production or delivery of the T7 RNA polymerase enzyme. The enzyme can be supplied by infection of cells carrying the appropriately cloned gene of interest with T7 phage. This however, also introduces competition from the other T7 promoters for the enzyme and rapid cell lysis then occurs (Studier and Moffat, 1986). The T7 gene 1 coding for the RNA polymerase has been cloned into Lambda derivatives and lysogenic hosts are available (Studier and Moffat, 1986) and also into plasmids under the control of regulated *E. coli* RNA polymerase promoters (Tabor and Richardson, 1985). Using these systems transcription of the T7 RNA polymerase can be controlled and hence the preferential expression of the gene of interest can be regulated.

4.2.1 Construction of plasmids pMW40 and pMW41

The following studies were designed to place the *amiR* gene downstream of a T7 promoter and to maximize transcription by providing T7 RNA polymerase from a DE3 lysogenic host. The *amiR* gene fragment was excised on a 1.5 kb *Hin*dIII/*Eco*RI DNA fragment from plasmid pSW24 (Wilson, 1991; Fig 4.1). The DNA fragment was purified by agarose gel electrophoresis using the freeze squeeze method. The fragment was then ligated into *Hin*dIII/*Eco*RI digested pGEM4Z (Promega), phenol: chloroform extracted and

ethanol precipitated. The ligation mix was transformed into fresh competent *E. coli* JM109 and recombinants selected by colour screen on IPTG/X-gal nutrient agar plates containing ampicillin. White colonies were patched onto NA/Amp plates and the patches grown in 5 ml N-broth cultures for DNA isolation by the miniplasmid preparation protocol. The isolated plasmid DNA was restriction enzyme mapped with *Hin*dIII and *Eco*RI and plasmid pMW40 with the 1.5 kb insert was identified (Fig 4.1).

To make an *amiR* expression vector carrying a shorter region upstream of the gene, pMW41 was constructed. The *amiR* gene was excised from pSW24 on a 924 bp DNA fragment with *Bsc*I and *Eco*RI and ligated to *Acc*I and *Eco*RI digested pGEM4Z. The recombinant would therefore have an *Acc*I/*Bsc*I hybrid site at the 5'end of insert. The ligation mixture was again transformed into *E.coli* JM109 and transformants selected as described above. Plasmid pMW41 (Fig 4.2) with the 954 bp insert was identified by *Hin*dIII/*Eco*RI restriction mapping.

Plasmids pMW40 and pMW41 both have the *amiR* gene cloned downstream of the vector T7 RNA polymerase promoter and the *lac* promoter. These two plasmids differ only in that pMW40 has more of the upstream sequence between the promoters and the coding sequence. The translation signal was expected to be the wild type *amiR* Shine Dalgarno sequence. Plasmid pGEM4Z was not designed an expression vector, however, T7 RNA polymerase *in vivo* is known to circumscribe plasmids bearing the promoter several times producing relatively stable mRNAs that can accumulate to levels of rRNA (Studier and Moffat, 1986).

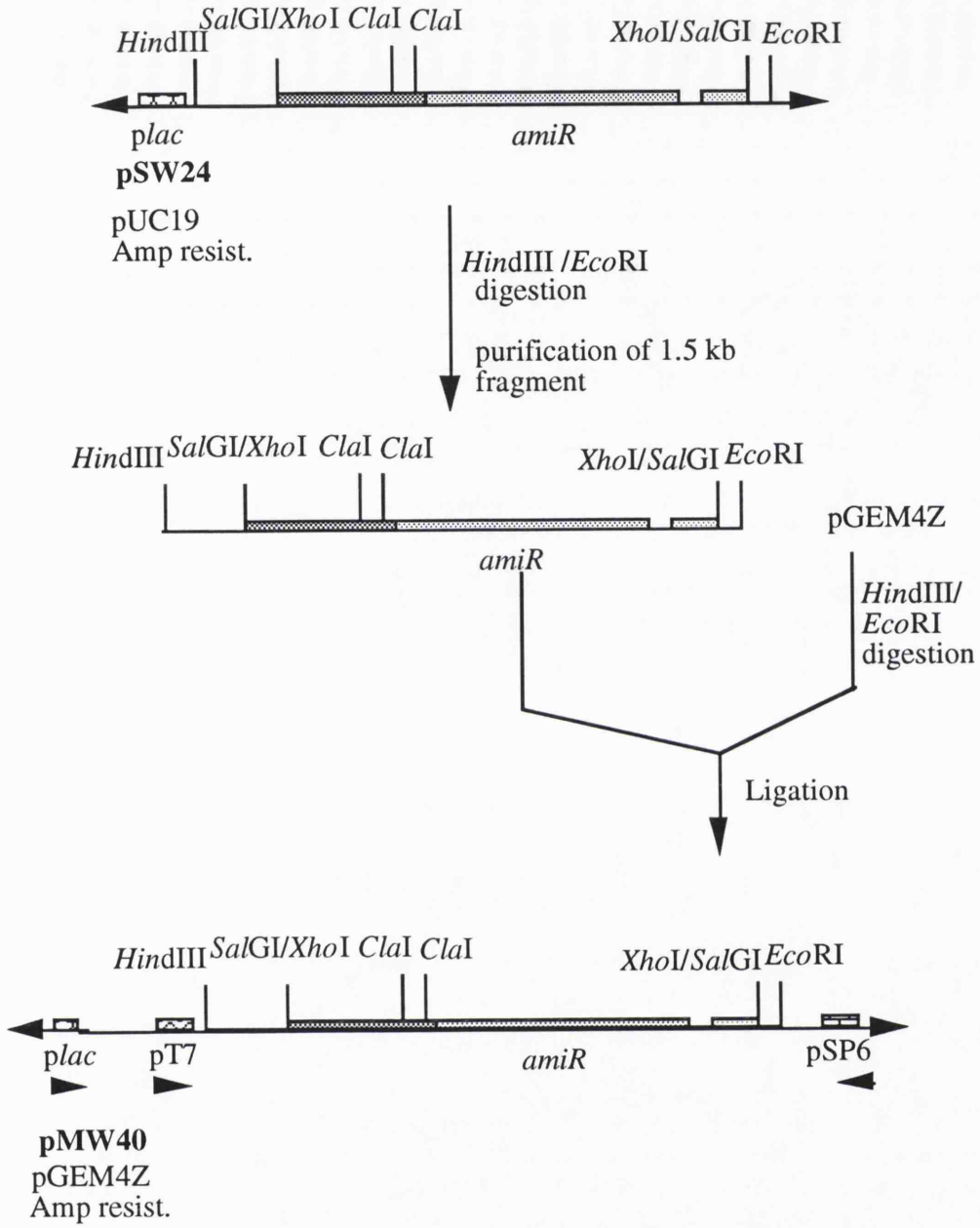


Fig 4.1 Construction of plasmid pMW40. The 1.5 kb *amiR* gene fragment from pSW24 was subcloned into pGEM4Z.

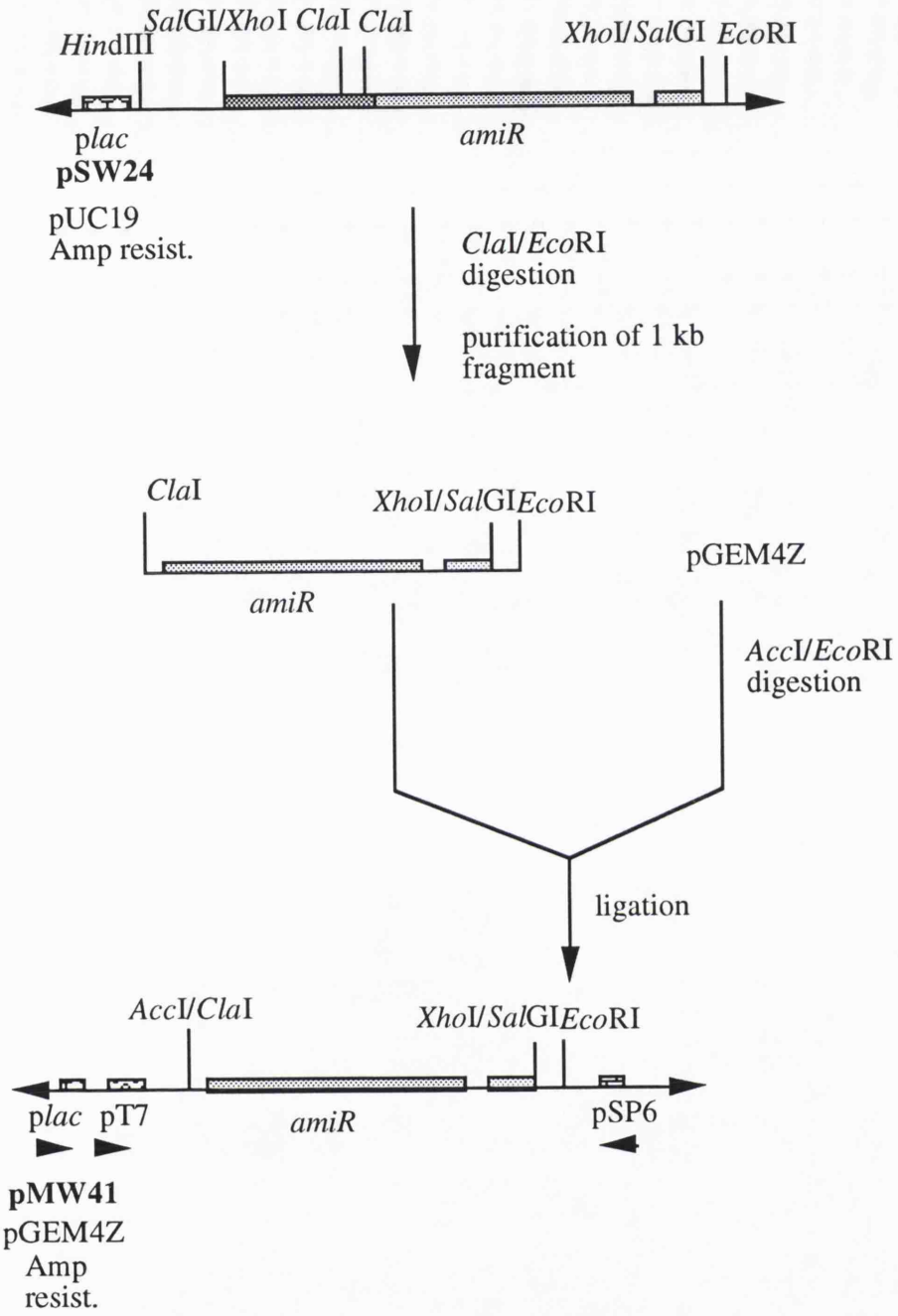


Fig 4.2 Construction of plasmid pMW41. The 954 bp *Acc*I/EcoRI DNA fragment from pSW24 was subcloned into pGEM4Z.

4.2.2 Protein expression from pMW40 and pMW41

Plasmids pMW40 and pMW41 were transformed into competent *E. coli* JM109 (DE3) (Promega) with selection on ampicillin nutrient agar plates. Transformants were grown in liquid media and DNA isolated by the miniplasmid protocol. The plasmids were checked by restriction enzyme mapping with *Hin*dIII and *Eco*RI. As a control for subsequent protein expression analysis, pGEM4Z was transformed into the same strain.

E.coli JM109 (DE3) has a single copy of the T7 gene 1 in the chromosome under *lac* UV5 promoter control. The T7 gene is inserted into an engineered *Bam*HI site within the *int* gene of the lambda cloning vector D69 and the Integrase function during strain construction was provided by a helper lysogen of heterologous immunity (Studier and Moffat, 1986).

Total cell extracts were prepared from 50 ml cultures of JM109 (DE3) pGEM4Z, JM109 (DE3) pMW40 and JM109 (DE3) pMW41 as described in Materials and Methods. The cells were sonicated for 5 minutes (15 seconds bursts followed by 15 seconds cooling intervals) at amplitude 14 and then centrifuged to remove membranes and cell debris. The extracts were analyzed on a 15% SDS-polyacrylamide gel after a 1:1 dilution in loading buffer and denaturation. The Coomassie stained gel is presented as Fig 4.3. A new 23 kDa AmiR band, the calculated size from the DNA sequence, was anticipated in the fourth and fifth tracts which were loaded with extracts from JM109 (DE3) pMW40 and JM109(DE3) pMW41 respectively. However, these lanes were identical to the controls. To eliminate the possibility that AmiR was associating with the cell debris in an insoluble form, whole cell extracts were made by boiling pellets from cultures grown in an identical manner to the ones above in 1/10 of culture volume, 1 x gel loading buffer and once again run on a similar gel. No new band was observed that could represent AmiR (data not

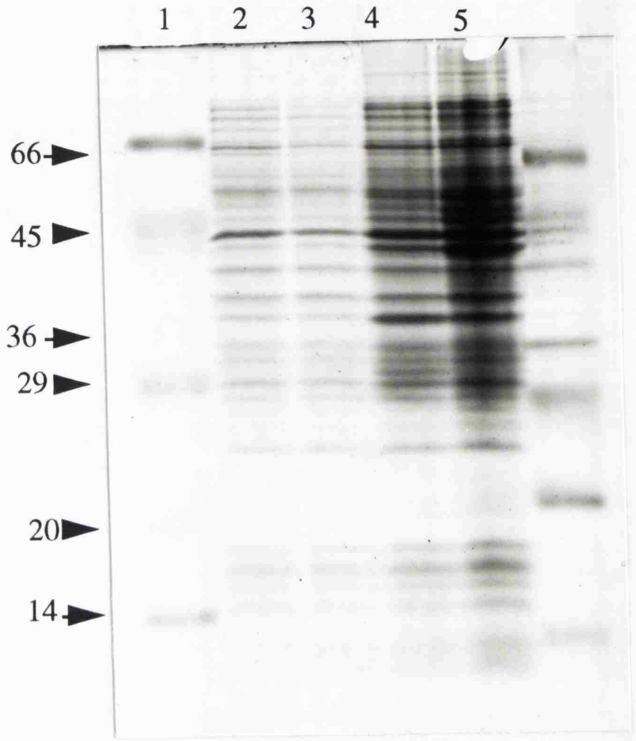


Fig 4.3 SDS-polyacrylamide gel of extracts of JM109(DE3).

Cell free extracts were analysed on a 15% SDS-polyacrylamide gel and stained Coomassie blue. Lanes: 1) molecular weight markers; 2) JM109(DE3) extract; 3) JM109(DE3) (pGEM4Z) extract; 4) JM109 (DE3) (pMW40) extract; 5) JM109(DE3) (pMW41) extract. The molecular weight markers are expressed in kDa.

shown).

To determine whether these constructs were expressing AmiR, plasmid pTM1 (*ami*E) (Wilson *et al* 1993) was transformed into the strain JM109 (DE3) pMW40 and JM109 (DE3) pMW41. Potential transformants were selected on chloramphenicol/ampicillin plates and resistant colonies patched onto the same medium then grown in liquid media for miniplasmid preparation. Isolated DNA was digested with *Hin*dIII and run on an agarose gel. Double transformants were thus identified and used for amidase assays. Plasmids pMW40 and pMW41 were also transformed into *E. coli* DH1 previously transformed with pTM1 and the strains identified in a similar manner.

Table 4.1 shows results of amidase assays performed on overnight nutrient broth cultures of these strains together with the controls in the presence and absence of IPTG. Strain DH1 pTM1 gives 0.7 units of activity. DH1 pTM1 pMW40 gave high constitutive levels (29.8 u) of amidase which rose only modestly in IPTG induced culture (40.8 u). This result is not seen in strain JM109 (DE3). JM109 (DE3) pTM1 gave 0 units. JM109 (DE3) pTM1, pMW40 gave 1.2 units and 0.2 units with and without IPTG respectively. DH1pTM1, pMW41 gave 0.9 units and 0.8 units with and without IPTG respectively.

Strain	- IPTG	+ IPTG
JM109 (DE3) pTM1	0.0	0.0
JM109 (DE3) pTM1 pMW40	1.2	0.2
JM109 (DE3) pTM1 pMW41	0.1	1.2
DH1 pTM1	0.7	-
DH1 pTM1 pMW40	29.8	40.8
DH1 pTM1 pMW41	0.9	0.8

Table 4.1 Amidase *trans* complementation, bioassays, for AmiR expression by T7 promoter constructs pMW40 and pMW41. The cultures were grown in nutrient broth. The results are expressed as µmoles acetylhydroxamate/minute/mg bacteria

Plasmid pMW41 also showed poor transcomplementation in JM109 (DE3) giving 0.1 units without IPTG and 1.2 units with IPTG. The cell density was similar in all the cases as determined by measuring the O.D. at 450 nm. In strain DH1, the transcription initiates entirely from the pGEM4Z *lac*Z promoter which is located upstream from the T7 RNA polymerase promoter. In strain JM109 (DE3), some transcription would be expected to start from *lacZ* promoter, however, due to the high activity of T7 RNA polymerase the vast majority of transcripts would be expected to initiate from the T7 RNA polymerase promoter both in the presence and absence of IPTG because of the leaky nature of the *lac*UV5 promoter. The *amiR* coding sequence is located more than 700 bp from the transcription start site of both promoters. Sequence analysis of both constructs indicates that there are unlikely to be any alternate reading frames in the *amiR* upstream region that would lead to translation of an alternative protein.

One possibility for low expression of AmiR from these constructs could have been due to a inefficient translation initiation. A potential SD sequence has been identified upstream of *amiR* which has several mismatches with the 3' end of 16S rRNA (Lowe *et al.*,1989; Wilson, 1991). The organization of the *amiR* and *amiC* genes is striking in that the termination codon of the latter overlaps the initiator codon of the former a property usually seen where the relative levels of two gene products is tightly controlled by programmed ribosome frame shifting whereby only a fraction of the initial ribosomes continue into the downstream gene (Studier and Studier, 1983).

4.2.3 Cloning of the *amiR* coding sequence.

The wild type *ami*R gene was PCR amplified with oligonucleotides which generated *Eco*RI (5') and *Xba*I (3') restriction sites on the ends of the gene. This PCR product was subcloned into the maltose binding protein fusion vector pmal CRI to generate plasmid

pSW50. Large quantities of pure AmiR-MBP fusion were produced by amylose affinity chromatography but following cleavage of the fusion with Factor Xa, the AmiR precipitated and was not refolded efficiently (S.A Wilson, personal communication).

The PCR[d] *amiR* gene was excised from plasmid pSW50 with *Eco*RI and *Hin*dIII and purified from an agarose gel. It was then ligated into *Hin*dIII/*Eco*RI cut pGEM3Z (Promega) and the ligation mixture transformed into competent *E. coli* JM109. The transformants were plated onto X-gal/IPTG/Amp plates. White colonies were grown for MPP and plasmid pMW45 was identified by restriction enzyme mapping with *Hin*dIII and *Eco*RI.

4.2.4 Construction of pMW50

The next set of over expression experiments were designed to provide an optimal RBS for *amiR* in addition to selective transcription by the T7 RNA polymerase. Plasmid pMW45 was double digested with *Eco*RI/*Hin*dIII and the 600 bp *amiR* DNA fragment purified by agarose gel electrophoresis. The fragment was cloned into *Eco*RI/*Hin*dIIIcut vector pT7.7 (Fig 4.4) (Tabor, 1990) . The ligation mixture was used to transform *E.coli* JM109 to ampicillin resistance. DNA was isolated from transformants by the MPP protocol and characterised by restriction mapping with *Hin*dIII and *Eco*RI. The recombinant plasmid pMW50 was identified (Fig 4.5).

The T7 expression vector pT7.7 contains the ϕ 10 T7 promoter and the translation start site for T7 gene 10 protein. By cloning with different enzymes within the engineered multiple cloning site, fusions to the gene 10 protein can be expressed (Tabor, 1990). The *amiR* gene in pMW50 is out of frame for expression in this way and required further

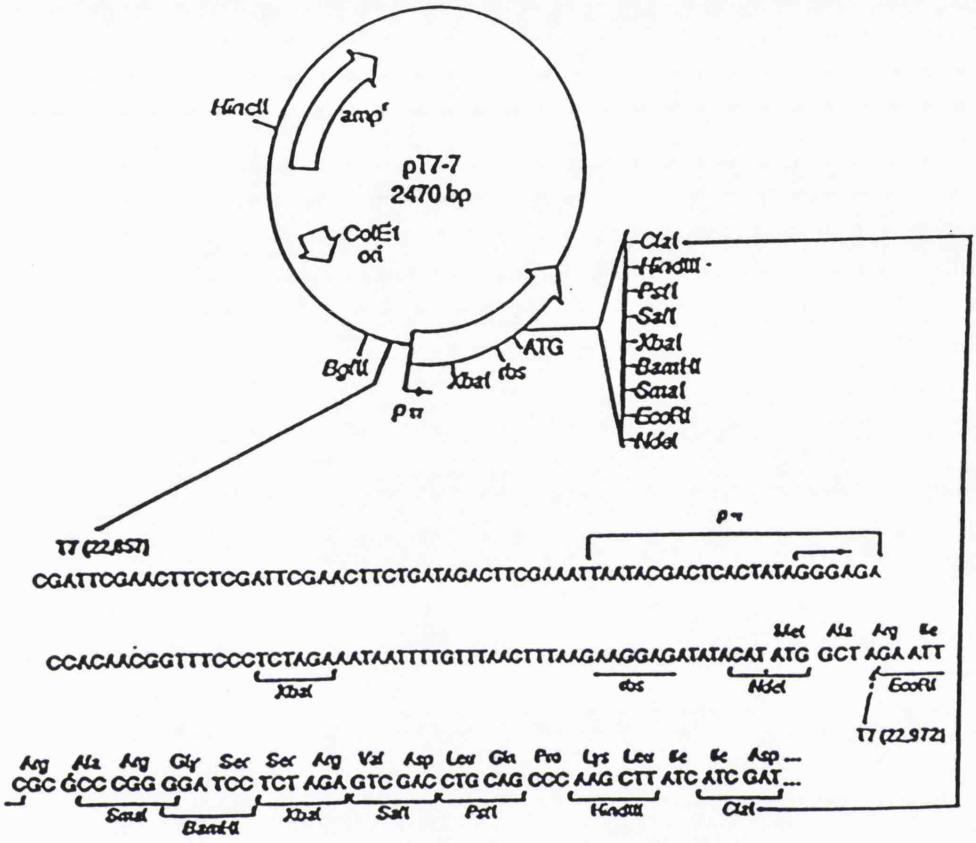


Fig 4.4 T7 expression vector pT7.7 contains a T7 RNA polymerase promoter and a SD sequence upstream of the multiple cloning site.

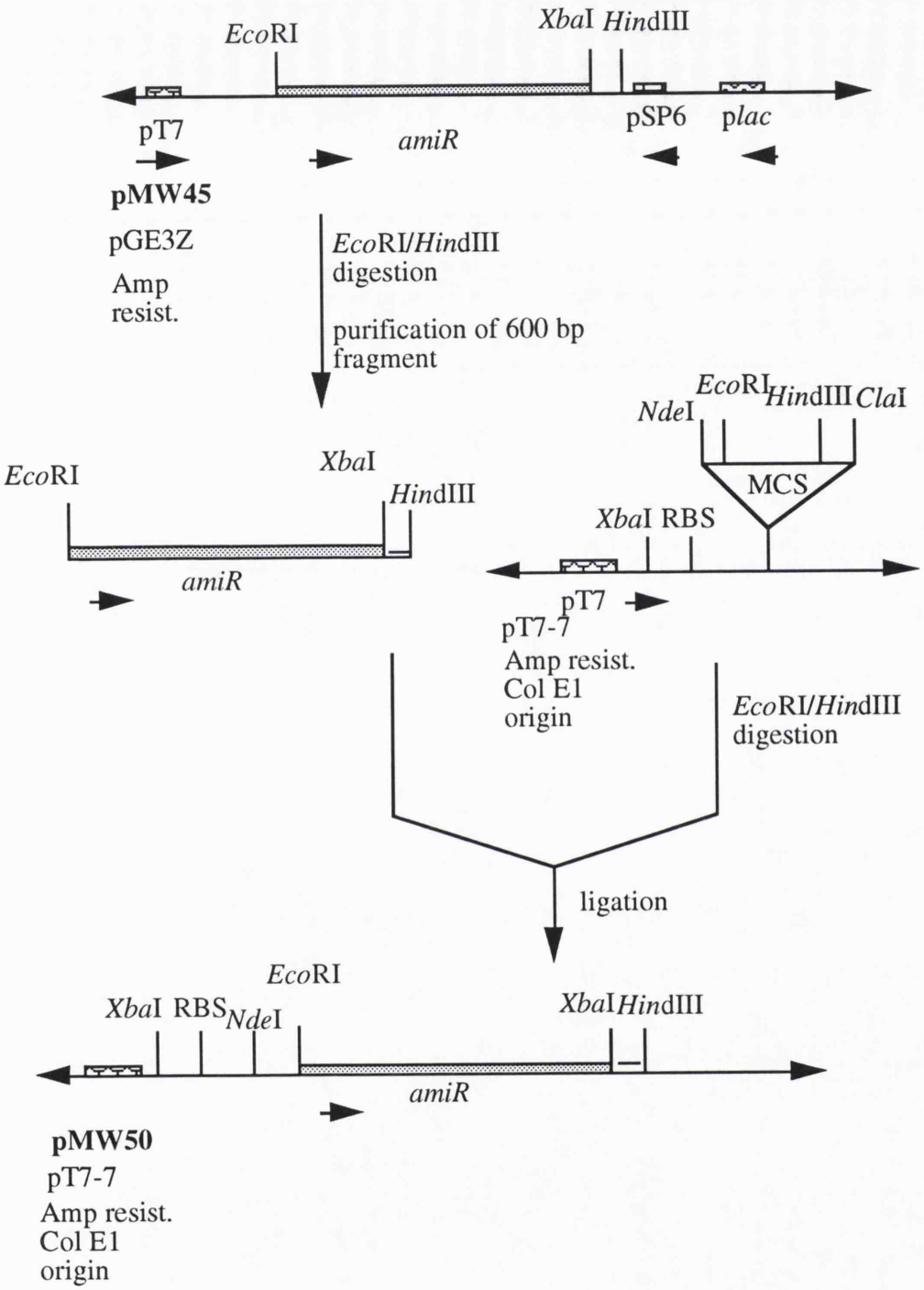


Fig 4.5 Construction of plasmid pMW50. The 600 bp *amiR* coding sequence was subcloned into pT7.7 downstream of the vector T7 promoter and SD sequence

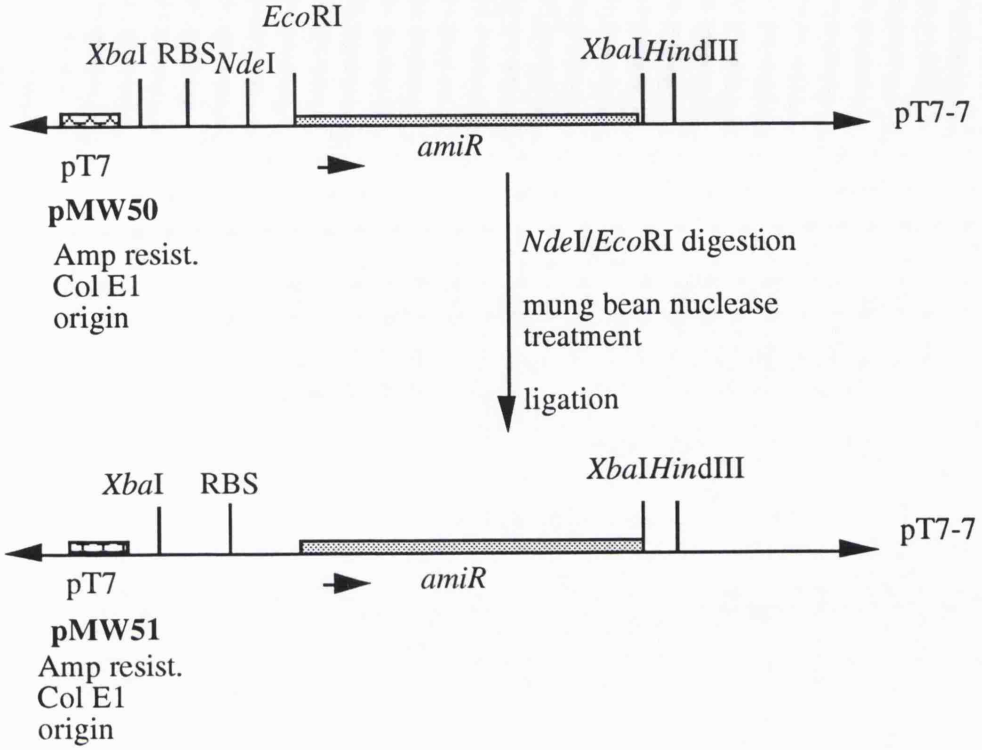


Fig 4.6 Construction of plasmid pMW51. pMW50 was modified to put *amiR* in the correct reading frame.

The junction was sequenced. – AAG**AAGGAG**ATATAC**ATG**AGCGCC—

SD *amiR*

manipulations (Fig 4.6). The plasmid was double digested with *Nde*I and *Eco*RI followed by trimming of the overhangs with mung bean nuclease. To ensure that both enzyme digests went to completion, pilot digests with each were set up and analysed on an agarose gel alongside an aliquot of the double digest. The digest was phenol:chloroform extracted and ethanol precipitated prior to setting up the ligation. The ligation mixture was transformed into *E.coli* JM109 and ampicillin resistant transformants were selected.The constructs were initially analyzed by digestion with *EcoRI/Hin*dIII and *NdeI/Hin*dIII for the absence of the *EcoR*I and *Nde*I sites.

To check that the modified plasmid was correct, the junction of one of the constructs (pMW51) was sequenced. Plasmid DNA was purified from 100 ml culture using the Qiagen plasmid Midi protocol (Qiagen Inc). The double stranded DNA was denatured by alkali before annealing a 20-mer T7 promoter primer (Promega). Sequencing was performed with the sequenase version 2.0 kit (USB).

The sequence obtained at the join has the correct reading frame with the AGGA SD sequence located 11-7 bp upstream of the AmiR initiation codon AUG. The last C and A from the 3' and 5' ends of the join respectively were evidently lost during the manipulations. This is not expected to affect expression because the optimal SD location is centred 11-7 bp upstream of the initiation codon (Shine and Dalgarno, 1975).

In the following experiments, the T7 RNA polymerase was supplied on a second plasmid pGP1-2 (Tabor and Richardson, 1985). This plasmid has a p15A origin of replication and is hence compatible with pMW51 which is Col E1 based. The T7 RNA polymerase is expressed from the λ P_L promoter under the control of the thermolabile λ repressor CI-

857. The repressor in turn being expressed constitutively from a *lac* promoter.

Plasmids pT7.7, pMW51, pGEM4Z and pMW40 were used to transform *E.coli* C600 pGP1-2. Potential double transformants were selected on kanamycin/ampicillin plates. Single well isolated colonies were patched and plasmid DNA isolated for restriction enzyme analysis to confirm presence of the plasmids. Crude cell free extracts were prepared from C600 pGP1-2, pT7.7; pMW51; pGEM4Z; pMW40 as described earlier (Tabor, 1990) . 50 ml TB broth media at 30°C was seeded with 1 ml overnight cultures. The cultures were aerated at 30°C to an OD_{450} of 0.4-0.5 and induced by incubation with occasional shaking at 42°C for 30 minutes. The cultures were then grown with vigorous shaking at 37°C for a further 90 minutes and harvested at 4°C. The cells were broken by sonication as described elsewhere and the supernatants diluted in 2 X SDS-PAGE loading buffer. The samples were denatured by boiling for 2 minutes and loaded onto a 15% SDS-polyacrylamide gel (Fig 4.7). Lane 2 was loaded with the pGP1-2 pT7.7 extract and lane 3 with pGP1-2 pMW51 extract. A single major new band of 23.17 kDa, as estimated from a standard curve of log mwt Vs distance of migration is seen on lane 3. This compares well with the predicted size of AmiR (21.8 kDa) from DNA sequence studies (Lowe *et al.*, 1989) and with the 23 kDa radiolabelled product from *in vitro* transcription/translation (Cousens 1985). The new band accounts for approximately 20% of the total soluble protein as determined by densitometry scanning and was presumed to be AmiR. Lanes 4 and 5 loaded with pGP1-2 pGEM4Z and its *amiR* derivative pMW40 are virtually identical and contain no visible AmiR band. This result is similar to that where the T7 RNA polymerase was provided by a DE3 lysogen (above). This is in spite of good *trans* complementation by pMW40. The critical difference between pMW51 and pMW40 is that in the latter, the *amiR* coding sequence is located more than 700 bp downstream of the T7

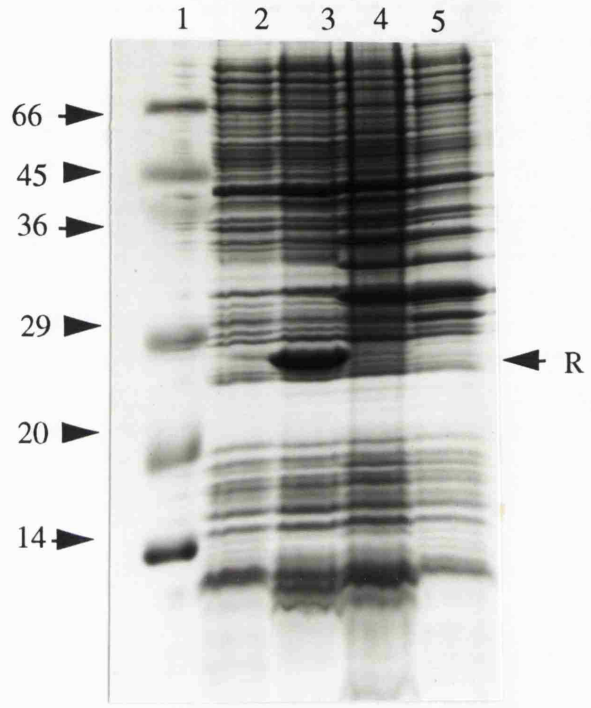


Fig 4.7 SDS-polyacrylamide gel of extracts of *E. coli* C600.
AmiR overexpression using T7 RNA polymerase/promoter two plasmid system. Whole cell extracts were separated on a 15% SDS-polyacrylamide gel and stained with Coomassie blue. Lanes: 1) molecular weight markers expressed in kDa; 2) C600 (pGP1-2)(pT7.7) extract; 3) C600 (pGP1-2) (pMW51) extract ; 4) C600 (pGP1-2) (pGEM4Z) extract; 5) C600 (pGP1-2) (pMW40) A major new band in lane 3 is indicated by arrow R .

promoter and translation would utilize the WT *amiR* SD. This result suggests that the WT AmiR translation signal is rather inefficient.

4.3 AmiR purification attempts

The isolation and characterization of AmiR has been a long term aim of the investigations of the amidase regulatory system. Previously encountered problems have included the lack of a suitable *in vitro* assay and lack of strains able to overproduce the polypeptide (Farin, 1976; Wilson 1991). The T7 expression system described above was promising in that, due to the relatively large quantities of the protein produced, the purification process could be followed with Coomassie stained SDS-Polyacrylamide gels. 500 ml cultures were grown in 3 litre shakeflasks as described elsewhere with the induction and growth conditions being identical to those described above.

4.3.1 Ammonium sulphate fractionation

The cells were harvested by centrifuging at 5000 rpm in a Sorvall GSA rotor at 4^{o}C , washed once in TES buffer (50 mM Tris-Cl, pH8.0, 2 mM EDTA, 20 mM NaCl) and sonicated in 10-20 ml cell lysis buffer (described in Materials and Methods) followed by clarification by centrifugation in a Sorval SS34 rotor at 15000rpm for 15 mins at 4^{o}C. The supernatant was transferred to an Erlenmeyer flask on ice and fractionated with ammonium sulphate in 10% steps as detailed in Materials and Methods.

The AmiR fractionated almost exclusively in the 20% precipitate (Fig 4.8) and remained as a suspension even in copious amounts of buffer. At this stage it was impossible to tell whether the ammonium sulphate treatment was denaturing the protein rendering it insoluble or whether AmiR was aggregated within the cells.

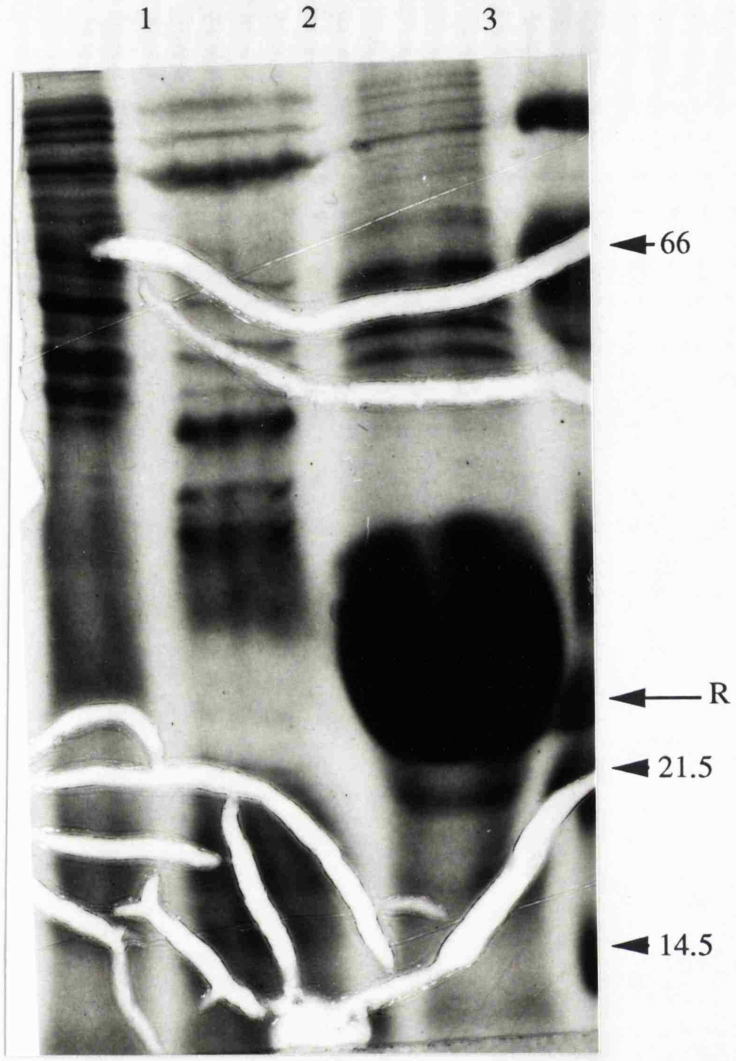


Fig 4.8 SDS-polyacrylamide gel of AmiR ammonium sulphate fractions stained with Coomassie blue. Ammonium sulphate fractionation of C600 (pGP1-2) (pMW51) extract. The fractions were analysed on a 15%SDS-polyacrylamide minigel. Lanes: 1) 30% ammonium sulphate pellet; 2) 20% ammonium sulphate supernatant; 3) 20% ammonium sulphate pellet. The relative molecular weights are expressed in kDa. arrow R indicates AmiR.

4.3.2 Ion exchange chromotography I

Another attempt at AmiR isolation was made by fractionation of a crude cell free extract prepared as above by ion exchange fast protein liquid chromatography (FPLC) (Pharmacia). The protein concentration in the crude extract was assayed by the Biorad miniprotein assay and 200 µl (17.3 µg) was loaded onto a monoQ HR 5/5 ion exchange column prewashed with elution buffer [200 mM Tris/HCl pH7.5, 1mM EDTA, 0.5 mM PMSF]. The column was eluted with a linear salt gradient of 0-1 M NaCl at a flow rate of 1 ml/min. The fractions representing the major peaks as estimated with a UV monitor were run on 15% SDS-polyacrylamide gels and stained with Coomassie blue. AmiR was not detected in any of these fractions although as estimated from the starting material it should have been the major peak. The flow through was not investigated because AmiR was expected to bind given its calculated pI of 6.8.

4.3.3 Ion exchange chromatography II

A third attempt to fractionate the crude extract using a +ve ion exchange resin was made by first binding AmiR to a DE52 column which also removes lipids and carbohydrates which might interfere with the performance of Mono-Q column. A fresh pellet was sonicated in Tris buffer [20 mM Tris/HCL pH7.2, 1mM PMSF] centrifuged and the supernatant loaded onto a pre-equilibrated DE52 column. The column was eluted with a linear gradient of 0-1 M NaCl. 50 µl samples from fractions with a significant OD at 280 were run on 15% SDS-polyacrylamide gels and those that contained a significant amount of AmiR were pooled. The experiment did not however result in an enrichment of the protein despite the fact that AmiR with a pI of 6.8 was expected to bind to DE52 resin at a pH of 7.2. The pooled fractions were precipitated with ammonium sulphate as described above and the fractions run on a 15% SDS-polyacrylamide gel. Protein with an appropriate molecular weight of

AmiR was spread over a wide range of fractions and as such no purification was achieved. Evidently most of the AmiR was not binding to the column presumably because of aggregation.

4.3.4 Ammonium sulphate/gel filtration fractionation

In the next purification attempt, it was decided to fractionate only the soluble material from the 20% ammonium sulphate pellet. A 2.7g cell pellet was resuspended in 20 ml of cell lysis buffer (described in Materials and Methods), sonicated and clarified by centrifugation as described above. The supernatant was fractionated with ammonium sulphate as described above and the fractions analyzed on a 15% SDS-polyacrylamide gel. AmiR was again identified in the 20% cut and this material was resuspended in gel filtration buffer [50 mM Tris/HCL pH7.2, 150 mM NaCl, 1mM DTT and 1mM EDTA] and clarified by centrifuging at 12000 rpm in a microfuge at 4^{o}C. The supernatant was loaded onto a Superdex 200 FPLC gel filtration column and 2 ml fractions collected. Fractions covering tubes 10-80 were selected and 20 µl samples were run on a 15% SDS-polyacrylamide gel alongside the crude extract, ammonium sulphate 20% pellet and supernatant. We did not detect significant amounts of AmiR in any of the fractions collected in significant amounts indicating that almost all of the 20% ammonium sulphate precipitated AmiR was insoluble.

At this point it was clear that the over expression system was in some way leading to aggregation of the protein. In some cultures most of the protein appeared in the pellet after sonication and centrifugation and some intermediate situations were also observed.

4.4 Denaturation and refolding of AmiR

It was assumed that the AmiR was present in cells in an aggregated, denatured form. At

this time a rabbit antisera raised against the purified MalE-AmiR fusion became available (S.A. Wilson personal communication). The antisera was used to monitor purification of any AmiR that remained in solution. A 20% ammonium sulphate pellet was made as described above from 3 g of cells and resuspended in gel filtration buffer (above) then denatured by addition of urea to 6 M. The protein was refolded by 100 fold dilution in Tris buffer [50 mM Tris/HCL pH8.0, 150 mM NaCl, 15% glycerol, 1 mM EDTA, 1 mM DTT, 1 mM PMSF]. The refolded protein was concentrated using an Amicon with PM10 membrane and loaded onto a superdex 200 gel filtration column. The fractions eluting from the column were run on 15% SDS-polyacrylamide minigels alongside the 20% $(NH_4)_2SO_4$ pellet and crude extract. The proteins were electroblotted using Novablot kit with MultiphorII electrophoresis system (LKB) onto a nitrocellulose membrane (.45μ) and the blots blocked and developed. The result presented on Fig 4.9 shows AmiR spread over fractions 10 to 22. A minor crossreacting product is seen above the AmiR band which was presumed to be MalE. The refolding process did not yield material that could be purified by gel filtration. Thus AmiR appears to be aggregating and forming multimers and refolding has not been successful.

4.5 N-terminal sequencing

Protein N-terminal sequencing was carried out to confirm the identity of the major band in the crude extracts. The 20% ammonium precipitate was resuspended in gel loading buffer, boiled for 5 mins and separated on 15%SDS-polyacrylamide gel (Laemmli). The proteins were then electroeluted onto protoblot membranes, stained with amidoblack for 2 mins and destained in 40% methanol until the bands were visible without background. The bands of the correct size were cut out and subjected to solid phase sequencing. The sequence of the first 20 amino acids was identical to that predicted from the DNA sequence.

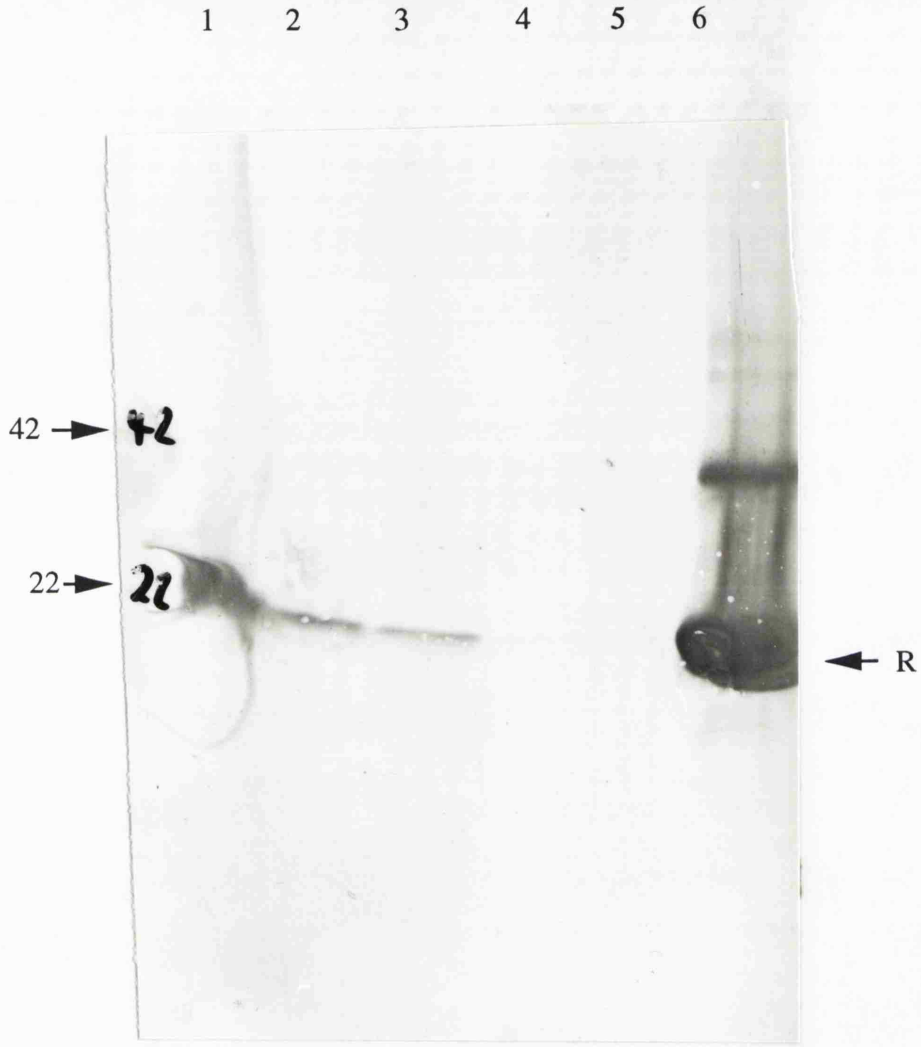


Fig 4.9 Western blot of refolded AmiR.
AmiR from the 20% ammonium fractionation step was denatured, refolded by dilution and fractionated on Superdex G100 column. The column fractions were analysed by western blotting. Lanes: 1) crude C600 (pGP1-2) (pMW51) extract; 2) column fractions 10-12; 3) column fractions 20-22; 4) column fractions 26-28 5) column fractions 38-40; 6) 20% ammonium sulphate pellet. Relative molecular weights are expressed in kDa and the arrow R indicates AmiR.

4.6 Summary

The main incentive for this piece of work was to obtain AmiR in sufficiently pure form for *in vitro* investigations of the transcription antitermination process (Chapter 5). Initial difficulties in overproducing the protein have been overcome using the T7 RNA polymerase/promoter system and the optimization of translation. Reports abound describing poor expression of *Pseudomonas* genes in *E.coli* (Buckel and Zehelem, 1981; Clarke and Laverack, 1983; Milton *et al*, 1983). The theories advanced for this phenomena are poor recognition of *Pseudomonas* promoters in *E.coli* due to significant variation in some promoter sequences (Rothmel *et al*, 1991) and the absence of activators in the heterologous host in case of positively regulated genes (Nakazawa and Inouye, 1986). Where large quantities of specific and stable mRNA have been produced, poor expression has been attributed to unfavourable distribution of rare codons caused by the high G+C content of *Pseudomonas* DNA (West and Iglewski, 1988).

In this study, the level of expression of AmiR has been shown, by deduction , to be limited only by the rate of initiation of translation. The AmiR levels achieved are apparently as high as those for *E.coli* bacteriophage proteins expressed with this system (Tabor and Richardson, 1985). This is in spite of a highly biased codon usage in the gene (Lowe *et al*, 1989) with only 41 of the possible 61 sense codons being used and only 8.7% having an A or T in the third position.

Whether the AmiR produced is denatured remains to be shown. A time course transcomplementation system set up by cotransforming pTM1 and pMW51 into JM109 (DE3) was carried out (data not shown). The non-induced overnight culture gives 5.6 units of activity which is about 10 fold less than obtained when the PCR^{d} *amiR* gene was cloned

into the pKK2.3 vector (Wilson S.A personal communication) to make pSW100. Only a modest increase in expression was seen after addition of IPTG to the JM109 (DE3) pTM1, pMW51 culture which shows that at least some the AmiR is active given that pTM1 on its own gives only 0.5 units of activity. Crude cell free extracts of pSW100 do not show a major new band on SDS-PAGE but AmiR can nevertheless be detected on Western blots indicating that only small amounts of AmiR are required for maximum amidase induction.

There is thus a possibility that AmiR requires very stringent conditions to remain in solution when expressed in above trace amounts and further purification attempts should address this problem first. There are reports of proteins that are only soluble during fractionation depending on the buffer, salt and glycerol concentrations. In fact during purification of the T7 gene 5 protein overexpressed using this system, as much as 2 M NaCl was initially needed to produce a soluble product (Tabor and Richardson, 1985). Preliminary threading of the AmiR sequence onto reported protein folds has shown structural resemblance of AmiR to the *E. coli* Endonuclease III with a hydrophobic N-terminal end and a hydrophilic C-terminus (Wilson personal communication) and presumably the hydrophobic domain is responsible for the observed aggregation.

CHAPTER 5

Investigation of the Leader mRNA/Protein Interaction

5.1 Introduction

Earlier studies of the amidase regulatory system showed that the *amiE* gene was subject to positive control by the *amiR* gene product (Farin and Clarke, 1978; Drew, 1984). However, these studies gave no insights into the mechanism of regulation. The better studied transcription factors (activators) in both eukaryotes and prokaryotes act by binding to DNA, as enhancer-like elements, and stimulating the isomerisation of the RNA polymerase at the promoter from a closed complex to an open complex. A second class of positive transcription regulators function by allowing readthrough of terminators or attenuators upstream of the gene. This occurs by either modifying the transcription complex, thus making it insensitive to terminators located at a distance from the point of action, or by preventing the formation of an RNA secondary structure which functions as a signal for termination in favour of a mutually exclusive antiterminator structure.

Sequencing of the *amiE* upstream region and construction of a 10 bp *Sma*I deletion (Drew and Lowe, 1989) led to the identification of a Rho-independenttranscription terminator whose disruption led to constitutive amidase expression which was indepedent of AmiR. The region upstream of the terminator shows some limited homology to the *bgl* operon of *E.coli* which was known to be regulated by transcription antitermination (Mahadevan and Wright, 1987). This led to the proposal of transcription antitermination by AmiR as the basic mechanism for the regulation of amidase expression.

Protein/ligand interactions can be demonstrated by specific titration of one by the other providing a method for assaying the unbound component exists. Initial experiments were designed to compete out the AmiR available for *trans* complementation by producing an excess of the leader RNA *in vivo* .

5.2 *In vivo* competition studies with the *amiE* leader region

5.2.1 Transcompetition

It was initially suspected that AmiR bound to the leader transcript during growth under inducing conditions and acted as the switch for the antitermination reaction. To test this assumption, the full leader region (*Hin*dIII/*Pst*I) (Fig 5.1) and the *Hin*dIII/*Sma*I fragment (Fig 5.2) which encompasses the leader with just the upward face of the transcription terminator were cloned into plasmid pGEM4Z and the effect of over expressing the leader transcripts tested with *amiE* and *amiR* being carried on separate plasmids. Plasmid pGEM4Z has the *lac* promoter upstream of the the *Hin*dIII site in addition to the T7 and SP6 promoters flanking the MCS. Expression of the leader in a strain lacking the T7 RNA polymerase could proceed from either the *amiE* promoter or from the *lac* promoter. These plasmids were then introduced into an *E. coli* host which carried two plasmids (*amiE* [pTM1] and *amiR* [pSW35]) that together produce high constitutive amidase expression. Titration out of AmiR caused by binding to the excess leader would be expected to reduce the amount of AmiR available for *trans* complementation and thus cause a decrease in amidase activity.

To make plasmid pMW42, the 450 bp *Hin*dIII/*Pst* I fragment of plasmid pAS20 (Fig 5.1) (Wison and Drew, 1991) was excised and purified from a 1.5 % agarose gel. It was inserted into *Hin*dIII/*Pst*I cut pGEM4Z, the recombinants transformed into *E.coli* JM109

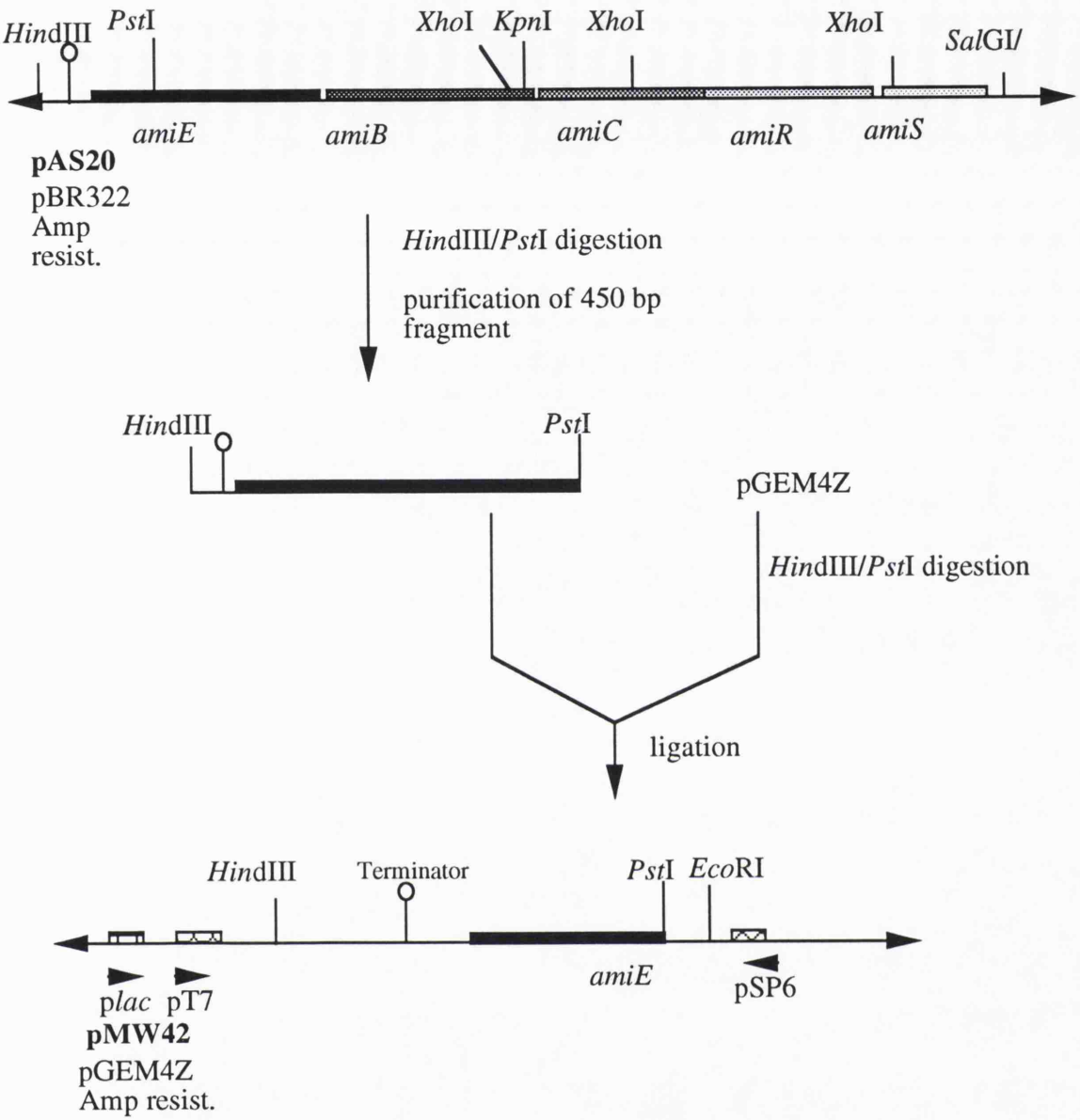


Fig 5.1 Construction of pMW42. The 450 bp *Hin*dIII/*Pst*I fragment of pAS20 carrying the full *amiE* leader and the first 63 codons of *amiE* has been inserted downstream of a T7 promoter in plasmid pGEM4Z.

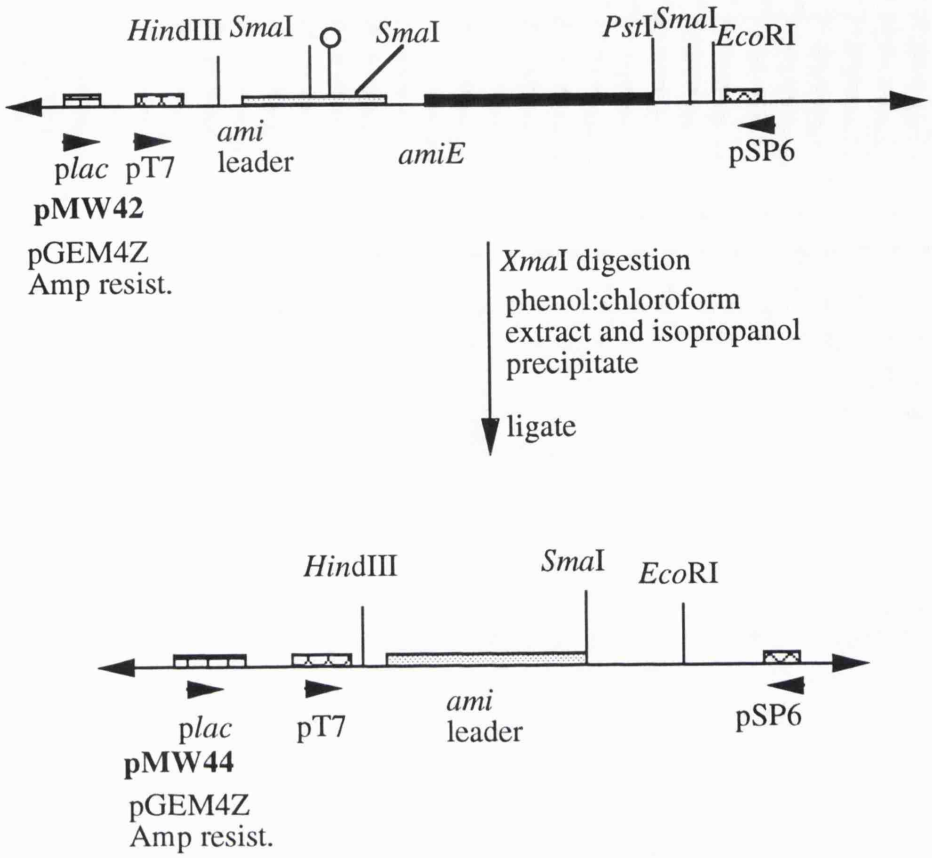


Fig 5.2 Construction of plasmid pMW44 from pMW42. Plasmid pMW42 was digested with the *Sma*I isoschizomer *Xma*I and ligated. A Plasmid that had lost the 250 bp *Sma*I fragment was identified by restriction enzyme mapping.

and the transformants selected on IPTG/X-gal/Amp plates. Plasmid DNA was isolated from white colonies and restriction enzyme mapped with *Hin*dIII and *Eco*RI. *E.coli* JM109 pMW42 was grown in a 100 ml culture for large scale plasmid DNA isolation and purification by the PEG precipitation. To make plasmid pMW44, pMW42 was digested extensively with *Xma*I, phenol:chloroform extracted twice , precipitated with isopropanol, washed with 70% ethanol and dried. The digest was resuspended in sterile water, ligated and transformed into *E. coli* JA221. A recombinant that had lost the 250bp *Xma*I/*Xma*I fragment (Fig 5.2) was named pMW44.

Competent JA221 pSW35 was cotransformed with plasmid pTM1 and either, plasmid pMW42 or pMW44. The transformants selected on Cm/Sm/Amp NA plates. The transformants were analysed by restriction enzyme mapping with *Hin*dIII. Colonies carrying all three plasmids were streaked out to single colonies for the hydroxamate transferase amidase assays and the results are presented in Table 5.1.

JA221 pTM1 (*amiE*) gives an average of 0.7 units of activity. JA221 (pTM1,pSW35) gives an average of 28 units of activity, which falls to 5 and 6 units -/+ IPTG respectively when plasmid pMW42 is present. This represents 17.8% and 21.4% of the original activity. With pMW44, averages of 9.6 and 11.6 units representing 34% and 41% activity respectively were obtained. There is no apparent effect of IPTG on the results which is probably due to good expression of the leader from the *amiE* promoter. These results clearly show competition by the two leader transcripts for AmiR available for *trans* complementation. The difference between the results with pMW42 and pMW44, 20% residual activity with pMW42 and 37% with pMW44 may indicate that part of the AmiR binding site is missing in pMW44. The conclusion here is that AmiR binds to the leader region upstream of the proximal *Sma* I site within the terminator.

Plasmids	IPTG +/-	Growth Conditions				
		Glucose	Glucose/ Lactamide	Succinate	Succinate/ Lactamide	Succinate Butyramide
pTM1,pSW35		31.3	28.0	29.2	27.5	25.7
pTM1, pSW35, pMW42	-	3.1	3.3	5.4	5.5	5.7
pTM1, pSW35, pMW42	+	3.2	3.8	9.6	8.3	5.8
pTM1, pSW35, pMW44	-	16.9	14.3	7.0	11.5	8.1
pTM1, pSW35, pMW44	+	11.7	19.1	7.4	4.8	5.4

Table 5.1 *In vivo* competition assays in *E. coli* JA221 with the *ami* leader provided in *trans* to *amiR*. The cultures were supplemented with tryptophan and leucine. The activities are presented as μ moles of acetylhydroxamate/minute/mg bacteria

Production of a regulator gene next to the point of action can amplify the response due to proximity effects. It was necessary, therefore, to look at this situation by cloning the leader region in *cis* to *amiR*. The *amiR* gene is normally located downstream of *amiE* and is expressed in the same direction (Cousens 1985). There is no evidence for *amiR* having its own promoter, however, DNA sequencing studies have identified two potential *rpoN* dependent promoters upstream of *amiC* (Cousens, 1985; Wilson, 1991). Deletion of the fragment carrying these promoters and part of *amiC* led to elevated amidase activity compared to wild type (Wilson, 1991). These findings suggest that *amiR* is normally expressed from the *amiE* promoter. However, a previously constructed plasmid (pCL54) carrying a deletion from the *Sma*I target at position 200 to the *Sma*I target at 1700 which

would have been predicted to be a good *amiR* expression vector did not *trans* complement *amiE* (Drew and Lowe, 1989).

5.2.2 Construction and analysis of *amiE* leader/*amiR* fusions

Several plasmids were constructed which had varying lengths of the *amiE* leader sequence stitched upstream of the 1.5 kb *XhoI*/*XhoI* *amiR* gene fragment. The AmiR produced from the *amiE* gene promoter would be expected to bind to the leader sequence (if the binding site was present) and residual AmiR could be assayed by transcomplementation with pTM1 (*amiE*).

Plasmid pMW40L (Fig 5.3) was constructed by ligating the *Hin*dIII/*Pst*I (450 bp) leader fragment into *Hin*dIII/*Pst*I cut pMW40 (Fig 4.1). The ligation mixture was used to transform competent *E. coli* JM109 to ampicillin resistance. Plasmid DNA was isolated from potential transformants and pMW40L identified by *Hin*dIII/*Eco*RI digestion compared to pMW40 digested with *Hin*dIII/*Eco*RI. Plasmid pMW40L thus contains the entire leader region and the first 63 codons of *amiE* joined to the 1.5 kb *amiR* gene fragment. The following constructs were then made to produce shorter 3' ends of the leader region. Plasmid pMW40L was digested with *Bst*EII and *Pst*I and the overhangs made blunt with the exonuclease activity of T4 DNA polymerase as described elsewhere. The digestion mixture was phenol:chloroform extracted twice, ethanol precipitated and ligated. The ligation mixture was transfomed into competent *E. coli* JM109 and plasmid DNA isolated from potential transformants. pMW46 (Fig 5.4) was identified by restriction enzyme mapping with *Hin*dIII and *Eco*RI. Plasmid pMW46 thus has the *amiE* leader region up to the middle of the terminator joined to the 1.5 kb *XhoI*/*XhoI* *amiR* gene fragment.

To make a shorter version of pMW46, plasmids pMW40 and pMW42 (Fig 5.1) were

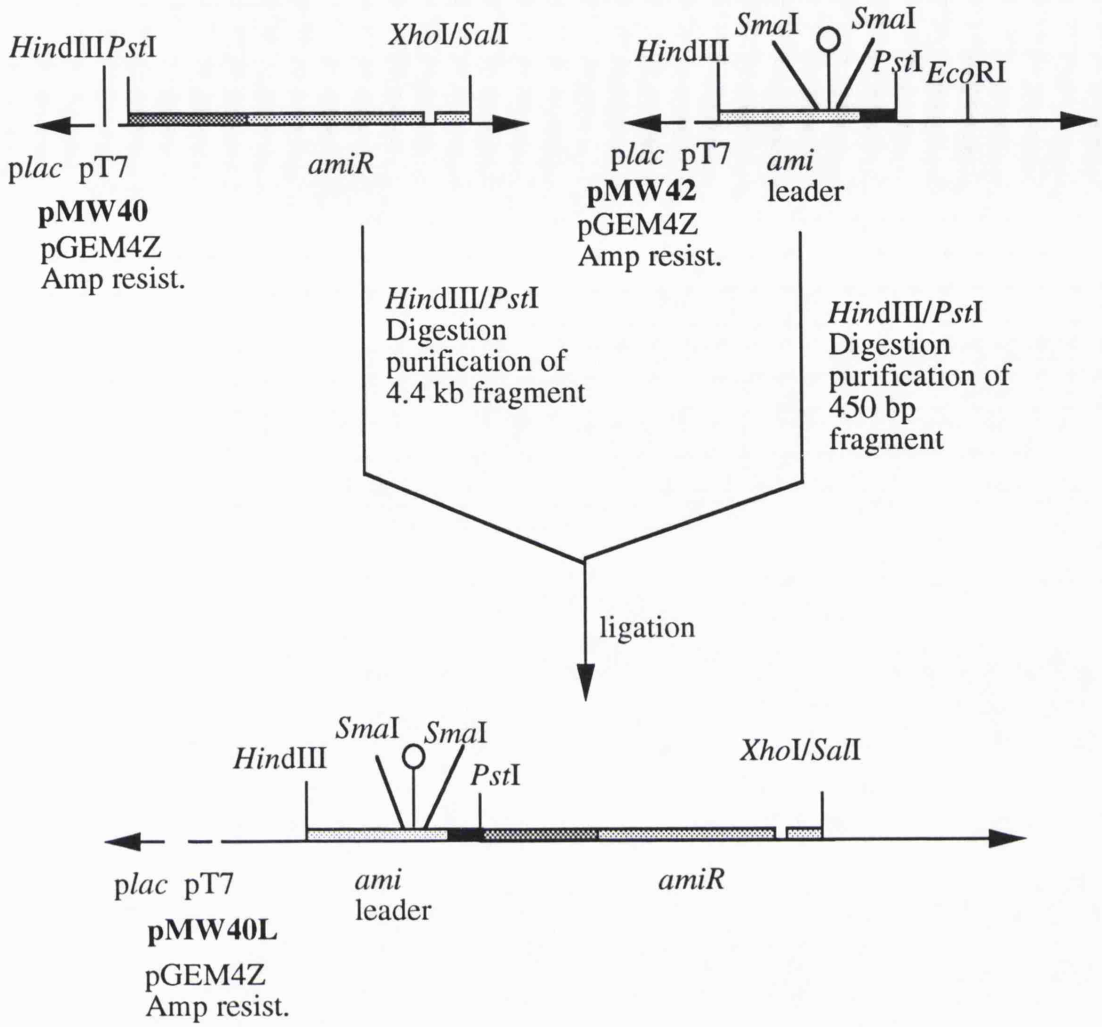


Fig 5.3 Construction of plasmid pMW40L
Plasmid pMW40 was digested with *Hin*dIII and *Pst*I and ligated to the *Hin*dIII/*Pst*I leader region excised from pMW42

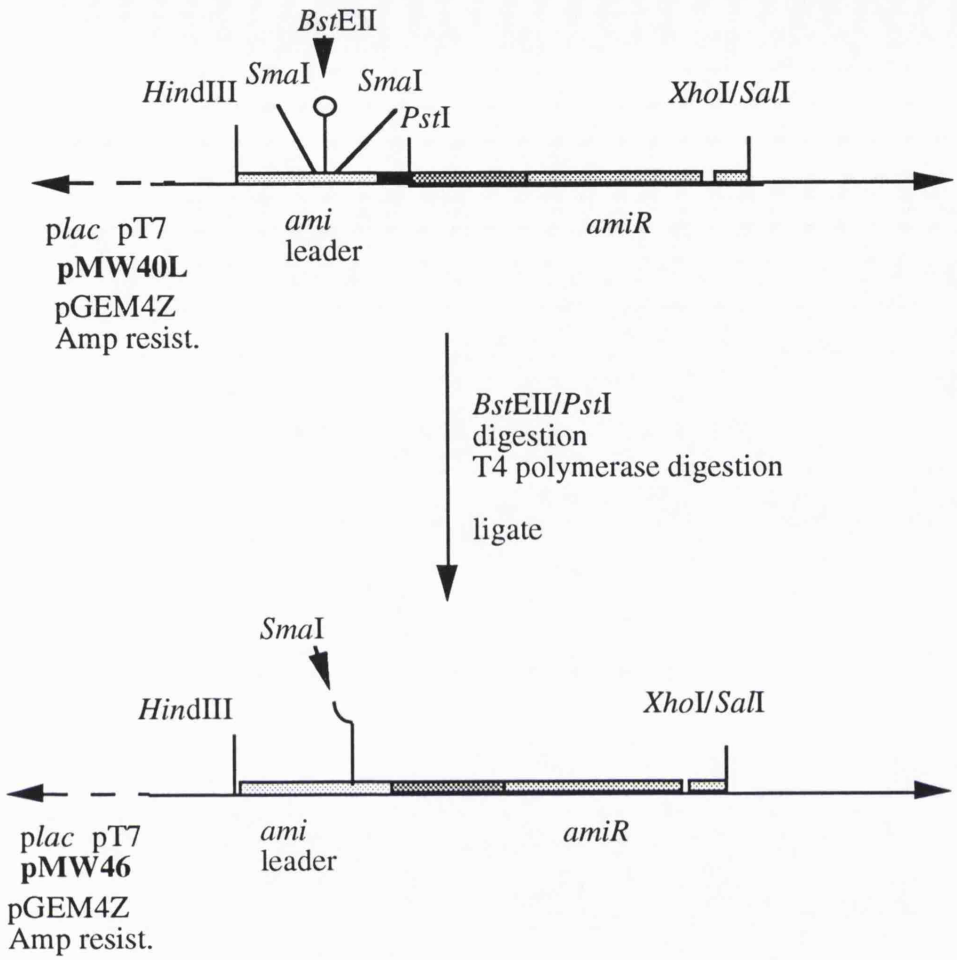


Fig 5.4 Construction of plasmid pMW46. Plasmid pMW40L was digested with *Bst*EII and *Pst*I, the ends made flush and ligated.

digested with *Pst*I and *Ban*II respectively and the overhangs made flush with T4 DNA polymerase. The digests were then phenol:chloroform extracted twice, ethanol precipitated and resuspended in x 1 buffer for digestion with *Eco*RI. The 1.5 bp pMW40 and 2.9 pMW42 fragments were purified by AGE, ligated and transformed into competent *E. coli* JM109. MPP DNA from ampicillin resistant colonies was restriction enzyme mapped with *Hin*dIII and *Eco*RI. Plasmid pMW47 (Fig 5.5) which is similar to pMW46 but has a further 9 bp deletion within the *amiE* leader region was thus identified.

The final deletion made went up to the *Fsp*I target in the leader region at position 140 and therefore the leader terminates only 7 bp downstream of the transcription start site (Wilson, 1991). This construct, pMW48 (Fig 5.6), was made by digesting pMW42 with *Fsp*I and purifying the 1.6 kp fragment from an agarose gel. The fragment was then extensively digested with *Hin*dIII and ligated into pMW40 which had been digested with *Pst*I to completion, the overhangs made blunt with T4 DNA polymerase and then digested with *Hin*dIII. Plasmid pMW48 was identified by restriction enzyme mapping with *Hin*dIII, *Eco*RI, *Hin*dIII/*Eco*RI, *Cla*I and *Hin*dIII/*Cla*I. Plasmids pMW40, pMW40L, pMW46, pMW47 and pMW48 were finally checked by digestion with *Hin*dIII and *Eco*RI to show the insert size (Fig 5.7).

All of these constructs together with pMW40 the positive control were transformed into *E. coli* JA221 pTM1 (*amiE*) (Wilson *et al.*, 1993). The cells were plated onto Cm/Amp/LA plates and resistant colonies patched on the same plates then grown in liquid medium for MPP and restriction enzyme analysis. Transformants were streaked out to single colonies on appropriate antibiotic plates and used for amidase assays. The results of the assays are presented in Table 5.2 are the averages of 5 duplicate assays carried out on different days. The full leader construct (pMW40L) gives 9.7 units of activity which is 28.9% of the

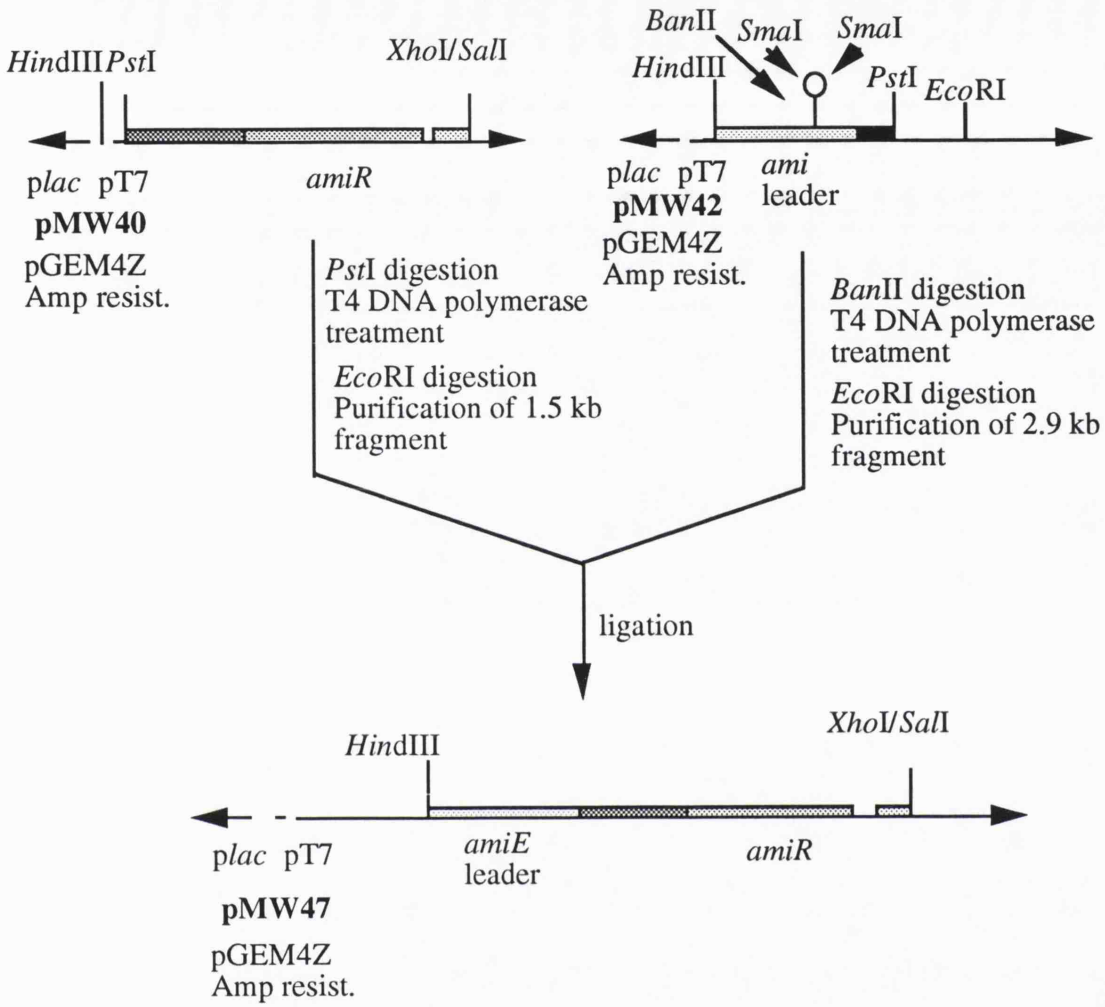


Fig 5.5 Construction of plasmid pMW47. The *amiR* gene was excised from pMW40 with *Pst*I (made blunt) and *Eco*RI and cloned into pMW42 cut with *Ban*II (made blunt) and *Eco*RI.

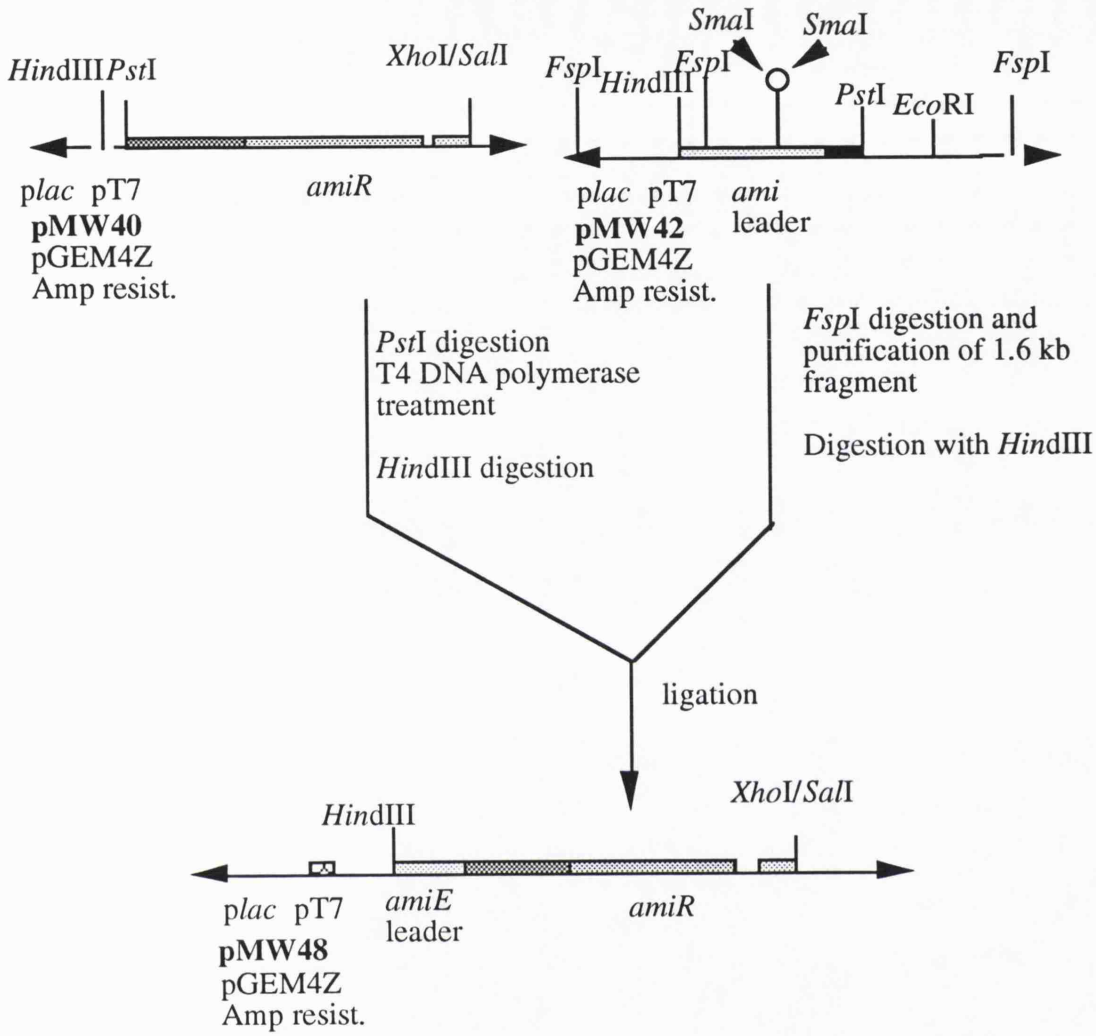


Fig 5.6 Construction of plasmid pMW48. The leader region fused to *amiR* carries only 6 bp from the transcription start site.

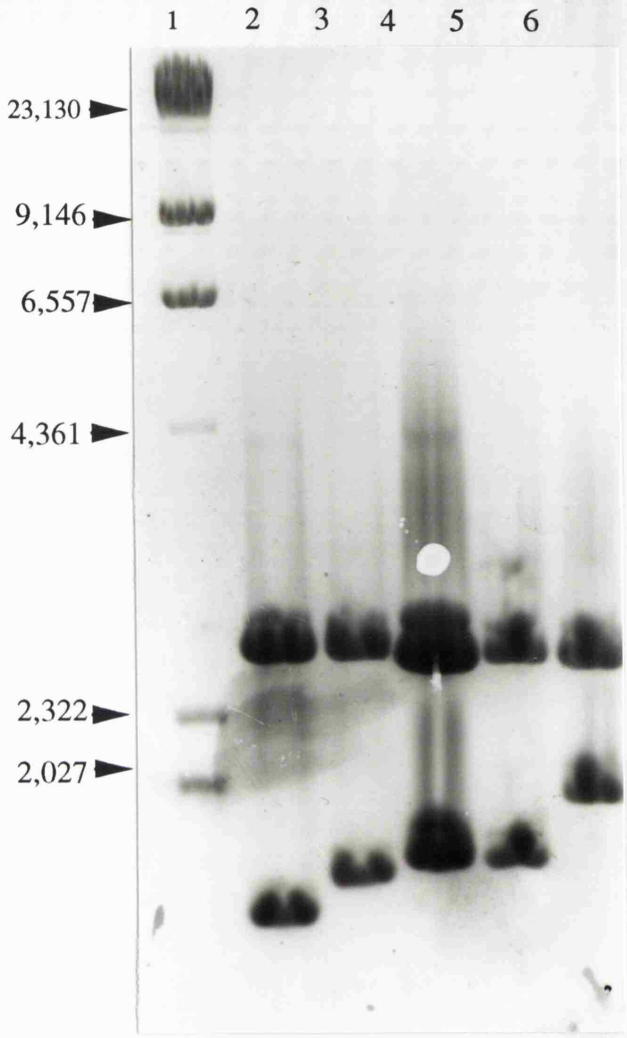


Fig 5.7 Ethidium bromide stained agarose gel of *ami* leader/ *amiR* fusion constructs.
Plasmid DNA was digested with *Eco*RI and *Hin*dIII and separated on a 0.8% agarose gel. Lanes: 1) lambda *Hin*dIII markers; 2) pMW40; 3) pMW48; 4) pMW47; 5) pMW46; 6) pMW40L

positive control (pMW40) indicating that the initial basal readthrough of the terminator produces enough AmiR to positively regulate the fused terminator and to *trans* complement *amiE*. The reduction from the pMW40 value (33.6 units) could result from, reduced transcription of *amiR* due to the presence of an upstream transcription terminator, an undefined reduction of transcription from the *lac* promoter caused by the presence of *amiE* promoter between it and *amiR* or titration of AmiR by the presence of a binding site(s) in the leader region. *E. coli* JA221 pTM1 pMW46, which has half of the terminator, gives 3.3 units of activity which is 10.0% of the pMW40 result. pMW46 has no functional terminator upstream of *amiR* and as such the reduction in activity could only be attributed to titration of AmiR by a binding site(s) located upstream of the *Bst*EII target in the leader region or the effect of *amiE* promoter. *E. coli* JA22, pTM1, pMW47 (with a further 9 bp deletion from pMW46) gives 12.4 units of activity which is 36.9% of the pMW40 result. *E. coli* JA221 pTM1 pMW48 gives an average of 15.8 units which represents 47% of the pMW40 (33.6 units).

Plasmid(s)	Specific Activity
pTM1	0.7
pTM1, pMW40	33.6
pTM1, pMW40L	9.7
pTM1, pMW46	3.3
pTM1, pMW47	12.4
pTM1, pMW48	15.8

Table 5.2. Amidase assays of *E.coli*, JA221, pTM1.
In vivo competition by the *ami* leader fused to *amiR*. The assays were done in nutrient broth overnight cultures and are expressed as as μ moles of acetylhydroxamate/minute/mg bacteria.

These results suggest that AmiR interacts directly with the leader RNA. A significant observation is that deletions upstream of the *Bst*EII (206) target led to a significant increase in amidase activity which suggests the presence of more free AmiR compared to that present with pMW46. A plausible explanation is that a binding site lies somewhere between the *Fsp*I target and the *Bst*EII target. The full leader situation appears to be inconsistent with the expected result, however, it could represent a mixed situation between limited termination and, if it were proven that AmiR binds to an alternate structure to the terminator, the presence of the majority of its binding site(s) trapped in the terminator structure. These experiments are however only pointers and clearly, *in vitro* evidence for the interaction was required. The following sections describe efforts towards that end.

5.3 *In vitro* transcription of the amidase leader region

5.3.1 Construction of Templates

The amidase leader region was cloned into pGEM4Z to make pMW42 with two objectives. Firstly, to use the construct for the *in vivo* titrations descibed above and secondly to use the flanking T7 RNA polymerase promoter to synthesize runoff transcripts *in vitro* . 500 ml *E.coli* JM109 pMW42 was grown overnight for a large scale plasmid DNA preparation which was then purified on a CsCl gradient.

Two 10 µg aliquots of this DNA were digested to completion with *Eco*RI and *Bst*EII respectively as determined by AGE. The digests were phenol:chloroform extracted twice, ethanol precipitated, dried, and then resuspended in sterile water at a concentration of 1 mg/ml. High specific activity RNA probe was synthesized with α-^{32}P CTP using the Riboprobe Gemini kit from Promega as described by the manufacturer. The transcripts were

purified on a 6% denaturing urea polyacryamide gel. To identify the location of the expected probes pBR322 *Hpa*II DNA fragments were dephosphorylated with CIP, end labelled with γ-^{32}P ATP using PNK and run alongside as markers.

The *Pst*I linearized template yielded two transcripts one corresponding to the full length leader and one shorter (Fig 5.8). By running a sequencing ladder alongside, the shorter transcript was estimated to correspond to termination at the run of T's immediately after the *amiE* terminator loop (data not shown). This is consistent with the distance migrated by the transcript from the *Bst*EII linearized template which runs slightly slower (Fig 5.8 Lane 2). This shows that the *amiE* leader terminator is recognized by the T7 RNA polymerase. This has been observed with some rho-independent transcription terminators including the *trp* attenuator but not others, although the reasons for discrimination are not clear (Jeng *et al.*, 1992).

5.3.2 Gel retardation analysis of AmiR/RNA complexes

5.3.2.1 Principles of gel retardation analysis

Gel retardation assays provide a convenient method for detecting specific RNA/protein interactions (Konarska and Sharp, 1986). The principle of the technique is based on the observation that nucleic acid/protein complexes exhibit reduced electrophoretic mobility in non-denaturing polyacrylamide or agarose gels (Lane *et al.*, 1992). The reasons for the phenomena are given as; increased mass of the nucleic acid due to bound protein, change in overall net charge caused by the protein contribution which is linked to the pH of electrophoresis buffer, conformational changes induced in the nucleic acid by the bound protein in addition to the gel concentration and composition. Elegant experiments have been performed to study such phenomena and used to map bends in DNA caused by regulator

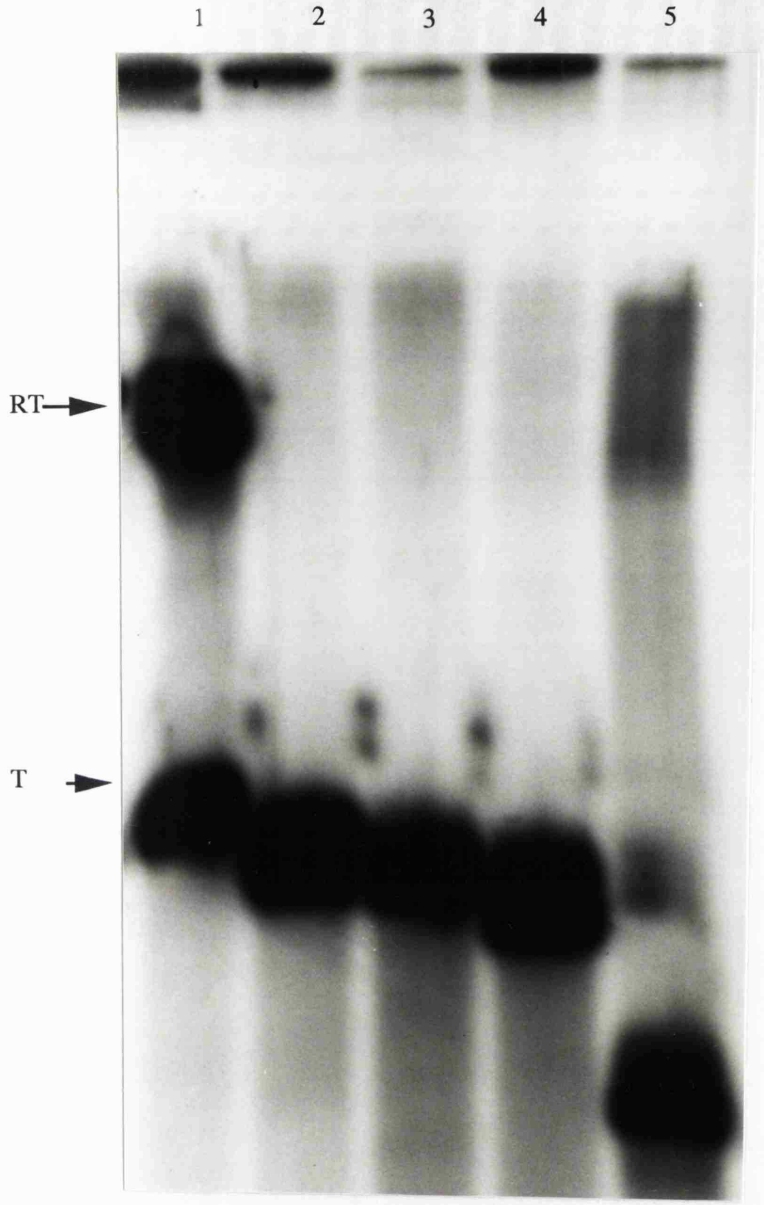


Fig 5.8 Purification of *in vitro* synthesised RNA transcripts.

Internally labelled runoff transcripts were separated on a denaturing 6% urea polyacrylamide gel. Lanes: 1) *Pst*I (450) linearized template transcript; 2) *Bst*EII (206) linearized template transcript; 3) *Ban*II (198) linearized template transcript; 4) *Ava*II (162) linearized template transcript 5) *Fsp*I (140) linearized template transcript. The terminated transcript in lane 1 is represented by the arrow T and readthrough transcript by the arrow RT

proteins. The absolute contribution of each of these factors is not known and is difficult to measure. The complexes are very stable during electrophoresis and in some circumstances they have been shown to have higher kinetic stability in the gel than in free solution (Fried and Crothers, 1984).

A potential pitfall however exists in that all proteins that recognize specific sites on nucleic acids also display some degree of non-specific affinity, the effect of which is to cause smearing and formation of multiple bands (Lane *et al.*, 1992). The techniques for eliminating this drawback are the use of high salt to dissociate non-specific interactions, having a high background of non-specific nucleic acid which include synthetic polymers for DNA gel retardations and the use of heparin (Lane *et al.*, 1992; Dall *et al.*, 1990). RNAase T1 has also been used in RNA bandshifts to digest away RNA nonspecifically bound to protein (Konarska and Sharp, 1986).

Gel retardation assays have been used to detect and identify both DNA and RNA binding proteins from crude cell extracts and have the advantage of requiring small quantities of material (picomole range) (Lane *et al.*, 1992). These factors were desirable because of the difficulties encountered in purification attempts of AmiR (Chapter 4).

5.3.2.2 Preparation of extract containing AmiR

The extracts used in the RNA/bandshift studies were prepared from *E.coli* pSW100 (S.A Wilson, personal communication). The cells were sonicated in cell lysis buffer and centrifuged as described elsewhere. The crude supernatant was eluted from a Q-sepharose ion exchange column with a 0-1 M NaCl gradient and fractions containing AmiR identified by western immunoblotting with the anti-AmiR rabbit antisera. The AmiR containing fractions were dialysed against the reaction buffer and concentrated by filtration in an

Amicon concentrator with a PM10 membrane. On some occasions the fractions containing AmiR from the Q-sepharose column were directly loaded onto a phosphate buffer preequilibrated heparin column and elutedwith a 0-1 M NaCl gradient. AmiR containing fractions were detected by Western blotting and concentrated for use in the gel retardation assays.

5.3.2.3 Leader transcript/AmiR binding studies

To investigate whether AmiR bound the leader transcript, 10,000 cpm of the high specific activity gel purified leader RNAs were incubated with the AmiR containing extracts as described in Materials and Methods and the complexes detected by separation on 5% native polyacrylamide gels followed by autoradiography. Each of the three probes, the full length and the terminated from the *Pst*I linearized pMW42 and the*Bst*EII run off transcript were incubated with a 2 x serial dilution of the AmiR containing extract and electrophoresed. A protein-free control was included as well an AmiR-free *E.coli* extract. Fig 5.9 shows the autoradiograph of the *Bst*EII transcript. No complexes were detected in lanes 1 and 2 which were loaded with a protein-free mixture and an AmiR-free *E.coli* extract incubation mixture respectively. Higher protein concentrations (towards the right) resulted in an increase in the amount of complex formed. The lagging RNA band shown by the arrow marked S is thought to be a result of secondary structure which is common when RNA is separated on a nondenaturing gel. Fig 5.10 and 5.11 show similar experiments carried out with the full length and terminated probes from the *Pst*I cut template. Similar results were obtained with no complex formation in lanes 1 and 2 loaded with protein free control and an AmiR free total *E. coli* extract. However, there appears to be an affinity difference between the *Pst*I linearized terminated probe and the *Bst*EII runoff transcript with the former having a higher affinity. The reason for this is not known but could simply be due to minor pippeting discrepances given the sensitivity of these assays.

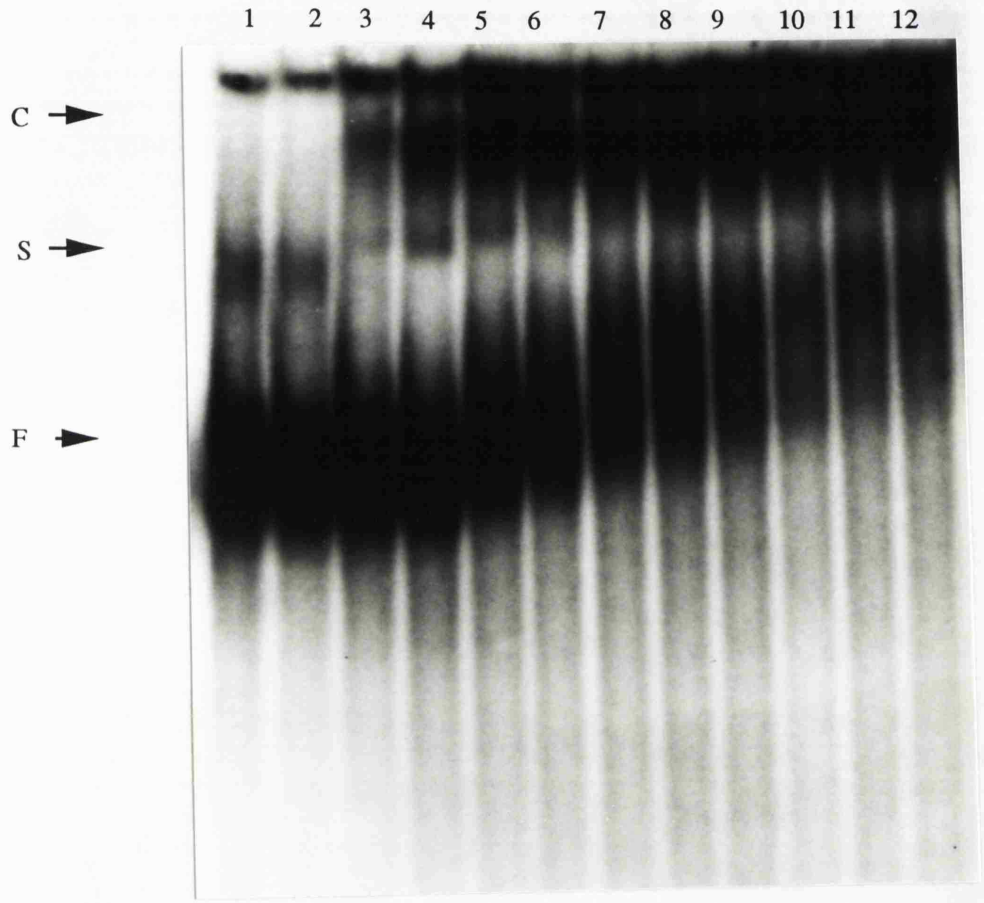


Fig 5.9 Bandshift assay for complex formation between AmiR and transcripts from *Bst*EII (206) linearized template.

Serially diluted Q-sepharose purified extracts were incubated with internally labelled transcript and separated on a nondenaturing 5% polyacrylamide gel. Lanes 1) protein free transcript; 2) transcript with AmiR free total *E. coli* extract; 3) transcript with 1/512 diluted AmiR extract; 4) transcript with 1/256 diluted AmiR extract; 5) transcript with 1/128 diluted AmiR extract; 6) transcript with 1/64 diluted AmiR extract; 7) transcript with 1/32 diluted AmiR extract; 8) transcript with 1/16 diluted AmiR extract; 9) transcript with 1/8 diluted AmiR extract; 9) transcript with 1/4 diluted AmiR extract; 10) transcript with 1/2 diluted AmiR extract; 11) transcript with 1 ul undiluted AmiR extract; 12) transcript with 2 ul undiluted AmiR extract. Free transcript is indicated by the arrow F and complex by the arrow C. Arrow S is thought to be secondary structure

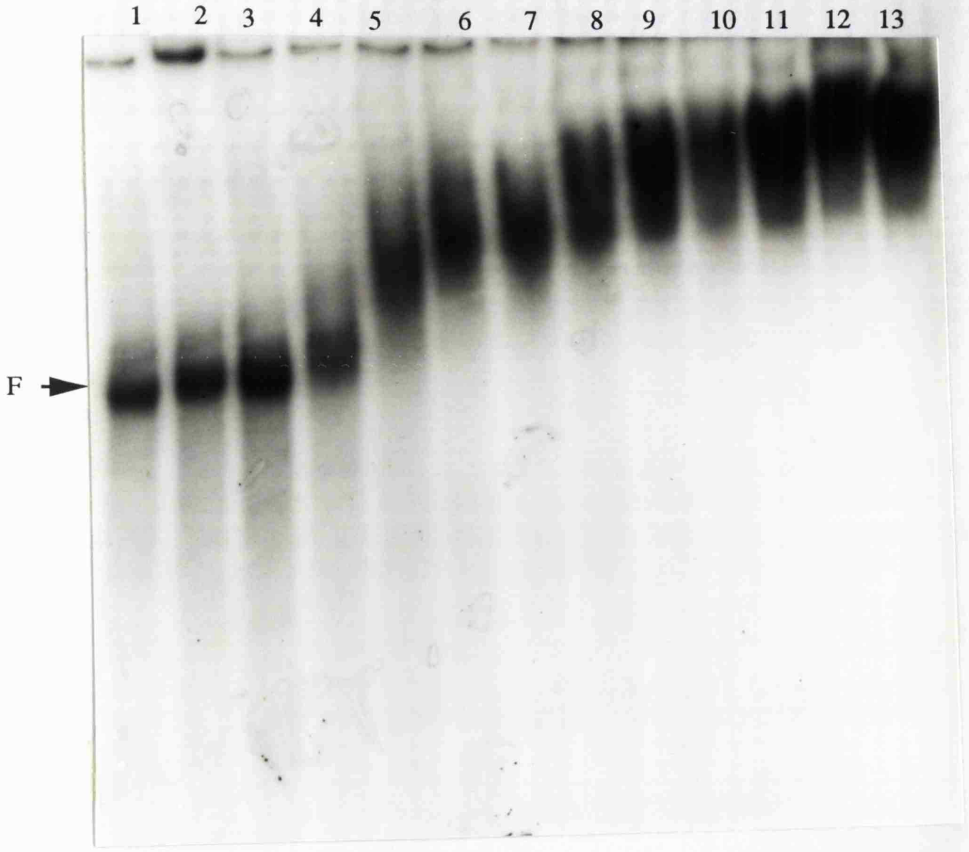


Fig 5.10 Bandshift assay for complex formation between AmiR and the full length transcript from *Pst*I (450) linearized template.

Serially diluted Q-sepharose purified extracts were incubated with internally labelled transcript and separated on a nondenaturing 5% polyacrylamide gel. Lanes 1) protein free transcript; 2) transcript with AmiR free total *E. coli* extract; 3) transcript with 1/512 diluted AmiR extract; 4) transcript with 1/256 diluted AmiR extract; 5) transcript with 1/128 diluted AmiR extract; 6) transcript with 1/64 diluted AmiR extract; 7) transcript with 1/32 diluted AmiR extract; 8) transcript with 1/16 diluted AmiR extract; 9) transcript with 1/8 diluted AmiR extract; 9) transcript with 1/4 diluted AmiR extract; 10) transcript with 1/2 diluted AmiR extract; 11) transcript with 1 ul undiluted AmiR extract; 12) transcript with 2 ul undiluted AmiR extract; 13) transcript with 3 ul undiluted AmiR extract. Free transcript is indicated by arrow F

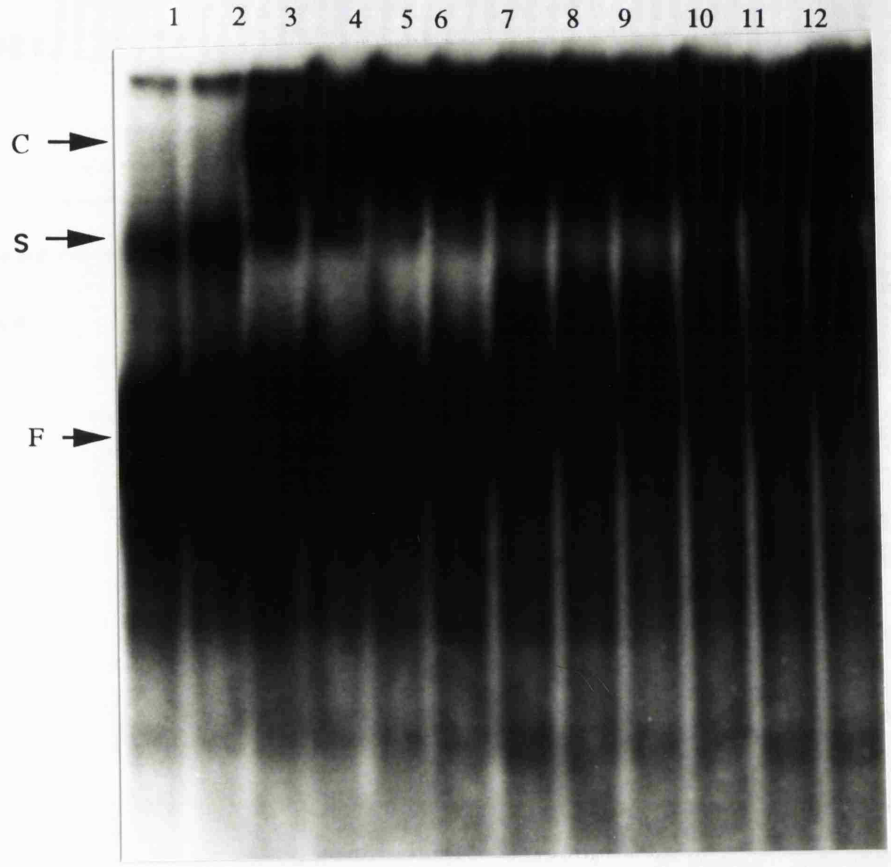


Fig 5.11 Bandshift assay for complex formation between AmiR and the terminated transcript from *Pst*I (450) linearized template.

Serially diluted Q-sepharose purified extracts were incubated with internally labelled transcript and separated on a nondenaturing 5% polyacrylamide gel. Lanes 1) protein free transcript; 2) transcript with AmiR free total *E. coli* extract; 3) transcript with 1/512 diluted AmiR extract; 4) transcript with 1/256 diluted AmiR extract; 5) transcript with 1/128 diluted AmiR extract; 6) transcript with 1/64 diluted AmiR extract; 7) transcript with 1/32 diluted AmiR extract; 8) transcript with 1/16 diluted AmiR extract; 9) transcript with 1/8 diluted AmiR extract; 9) transcript with 1/4 diluted AmiR extract; 10) transcript with 1/2 diluted AmiR extract; 11) transcript with 1 ul undiluted AmiR extract; 12) transcript with 2 ul undiluted AmiR extract. Free transcript is indicated by arrow F and complex by arrow C. Band shown by the arrow S is thought to be due to secondary structure.

The experimental conditions used to eliminate non-specific binding are discussed above. These band shift experiments were carried out against a large excess of total yeast RNA (40 μg) with respect to the labelled RNA (about 4 ng). This represents about 3200 molar excess of nonspecific RNA, the same order of magnitude used in other studies (Otridge and Golnick, 1993; Dall *et al.*,1990). This and the fact that the total extract from *E.coli* lacking AmiR did not form a complex shows that the binding is specific.

5.3.2.4 Studies of the binding Specificity

To study the specificity of the binding reaction an attempt was made to compete out the AmiR with specific cold leader mRNA. The reactions were set up with a constant amount of the AmiR extract and 10,000 cpm of ^{32}P labelled terminated probe from the *Pst*I linearized template. Fig 5.12 shows the result of the experiment. No shifts are apparent in lanes 1 and 2 which were loaded with free probe and the AmiR free *E.coli* control extract. Lane 3 shows a typical RNA protein complex indicated by arrow F. Lanes 4, 5, 6, 7 and 8 had in addition, .35 ng, 3.5 ng, 35 ng, 140 ng and 210 ng of cold specific RNA which had been purified on a denaturing polyacrylamide gel as described previously added at the same time as the labelled probe. In lanes 4, 5 and 6 just as in lane 3, two bands were observed, one migrating with the free probe and a lagging RNA/protein complex. Lane 7 which had in addition to the probe, 140 ng of RNA representing a 35 fold molar excess of cold competitor has only one band representing the free probe. Similarly one band representing protein free RNA was observed in lane 8. Thus a 35 fold molar excess of cold competitor RNA totally competed out the probe. This is a large excess but of course the reaction mixture contains a high level of non-specific RNAwhich can be increased to 120 ug with complex formation still occuring. Other workers have reported a requirement of up to 450 fold excess to obtain

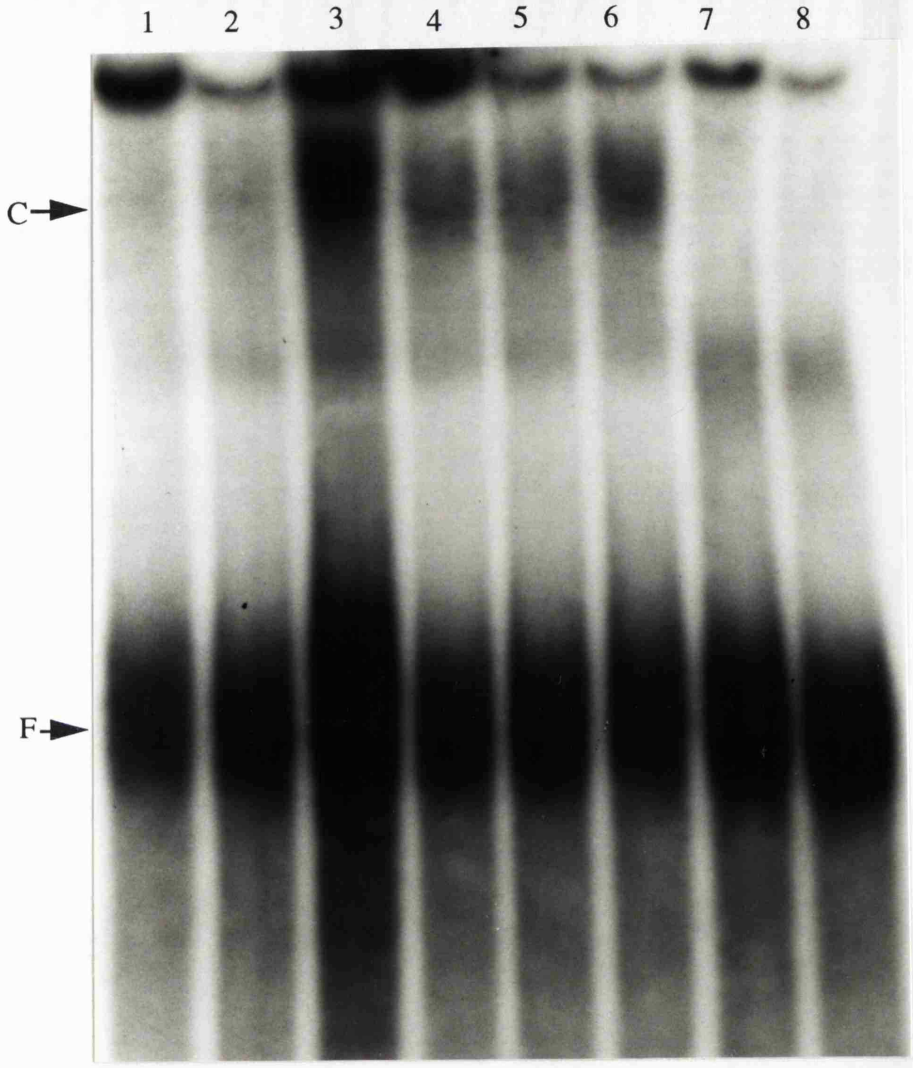


Fig 5.12 Cold specific competition of 'AmiR/*Bst*EII template runoff transcript 'complexes.

The AmiR containing extract was incubated with the internally labelled transcript in the presence of increasing amounts of cold specific RNA and separated on nondenaturing 5% polyacrylamide gel. Lanes: 1) free transcript; 2) transcript incubated with AmiR free protein extract; 3) transcript incubated with nonsaturating amounts of AmiR; 4) transcript incubated with nonsaturating amounts of AmiR in the presence of 0.35 ng specific cold RNA; 5) transcript incubated with nonsaturating amounts of AmiR in the presence of 3.5 ng specific cold RNA; 6) transcript incubated with nonsaturating amounts of AmiR in the presence of 35 ng specific cold RNA; 7) transcript incubated with nonsaturating amounts of AmiR in the presence of 140 ng specific cold RNA; 8)transcript incubated with nonsaturating amounts of AmiR in the presence of 210 ng specific cold RNA. Complexes are indicated by arrow C and free transcript by arrow F.

complete specific cold competition for unclear reasons (Otridge and Gollnick, 1993). Further evidence for the specific nature of the complexes will presented in the following sections.

5.3.2.5 Determining the minimum size of leader RNA bound by AmiR

The Riboprobe transcription system offers the ability to produce desired lengths of RNA runoff transcripts by linearizing the template with appropriate restriction enzymes and the production of complementary transcripts by use of the opposing T7 and SP6 promoters. Different lengths of ^{32}P labelled RNA transcripts were made from pMW42 linearized with *Ban*II, *Ava*II, and *Fsp*I whose recognition sites are highlighted on Fig 5.13. The probes were eluted from a denaturing polyacrylamide gel as described earlier. *Ban*II cleavage leaves protruding 3' termini and such templates are known to result in the synthesis of extraneous transcripts in addition to the expected ones (Promega technical manual). The overhangs were therefore trimmed with T4 DNA polymerase prior to phenol:chroloform extraction, ethanol precipitation and probe synthesis.

RNA bandshifts were performed with these probes with an AmiR extract as described. The result of this experiment is shown in Fig 5.14. For the *Fsp*I transcript 0, 1, and 6 µl of AmiR containing extract were used and 0, 1, 2, 4 and 6 µl samples were used for both the *Ban*II and *Ava*II transcripts. A retarded complex is seen in lane 13 of the *Ban*II transcript (Fig 5.14) but not in any of the rest. The *Ban*II target at position 198 therefore appears to delimit the 3' end of the AmiR binding site.

To delineate the 5' boundary of the AmiR/leader interaction extraneous DNA sequences including the *amiE* promoter were removed and transcripts were synthesized from plasmid

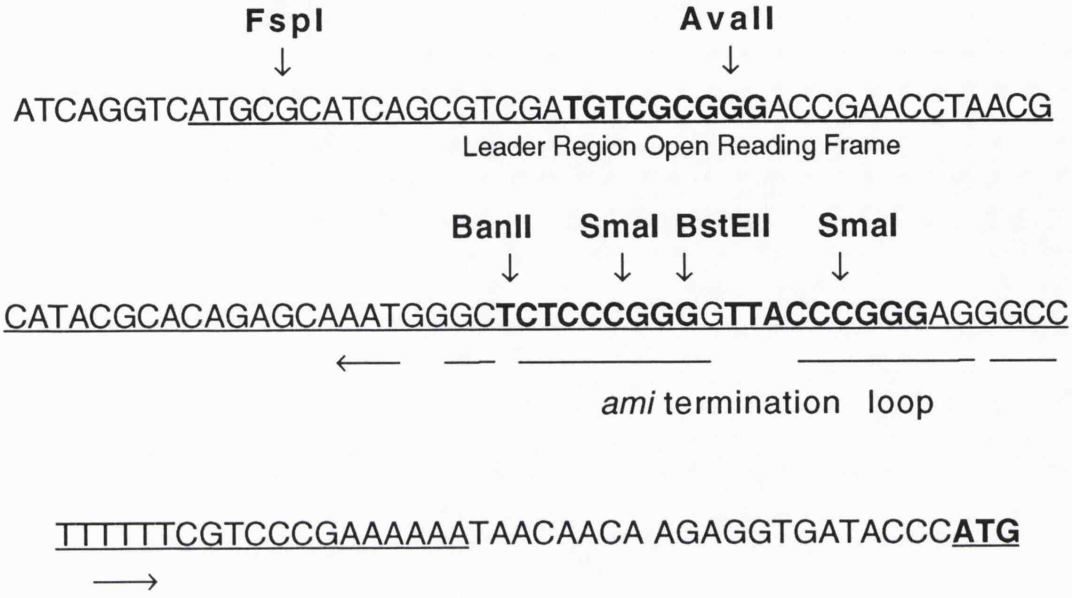


Fig. 5.13 DNA sequence of the amidase leader transcript.
The cutting sites for various restriction enzymes are shown, the leader ORF is underlined, the extent of the termination loop is shown by the arrows under the sequence and the *amiE* gene initiation codon is shown in bold.

Fig 5.14 Bandshift assays for complex formation between AmiR and transcripts made from:
*Fsp*I (140); *Ava*II (162) and *Ban*II linearized templates.
Q-Sepharose purified AmiR containing extract was incubated with internally labelled transcripts and separated on a nondenaturing 5% polyacrylamide gel. Lanes: 1) free *Fsp*I transcript; 2) *Fsp*I transcript incubated 1 ul AmiR extract; 3) *Fsp*I transcript incubated 6ul AmiR extract; 4) free *Ava*II transcript; 5)*Ava*II transcript incubated 1 ul AmiR extract; 6)*Ava*II transcript incubated 2ul AmiR extract; 7)*Ava*II transcript incubated 4ul AmiR extract; 8)*Ava*II transcript incubated 6ul AmiR extract; 9) free *Ban*II transcript; 10)*Ban*II transcript incubated 1 ul AmiR extract; 11)*Ban*II transcript incubated 2ul AmiR extract;12)*Ban*II transcript incubated 4ul AmiR extract;13)*Ban*II transcript incubated 6ul AmiR extract. Free transcripts are represented with arrows F, and complexes with arrow C

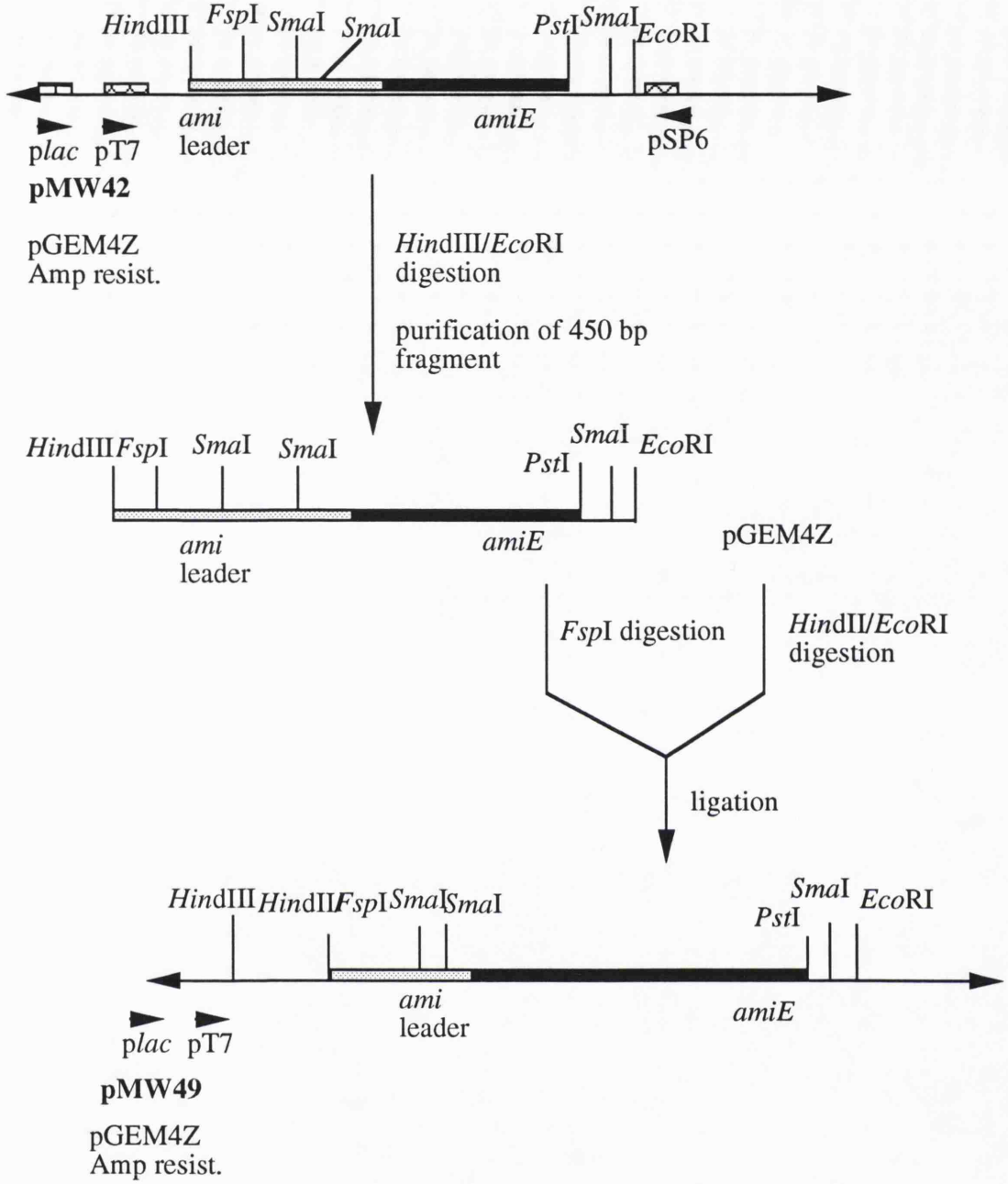


Fig 5.15 Construction of plasmid pMW49. The *Fsp*I/*Pst*I fragment from pMW42 was subcloned into pGEM4Z.

pMW49 which was constructed as follows (Fig 5.15) : plasmid pMW42 (Fig 5.1) was digested with *Hin*dIII/*Eco*RI and the 450bp fragment purified by AGE and elution from glass milk using the Geneclean kit; the eluted DNA was digested with *Fsp*I and stitched into *Hin*dII/*Eco*RI cut pGEM4Z. The recombinant plasmid was transformed into *E.coli* JM109 and selected histochemically on X-gal/IPTG/Amp L agar plates. A plasmid with the required insert was confirmed by restriction enzyme mapping with *Hin*dIII/*Eco*RI. For transcript synthesis, plasmid pMW49 was linearized with *Eco*RI, phenol;chloroform extracted twice, ethanol precipitated and resuspended at a concentration of 1mg/ml in sterile water after drying. The linear template was used in the *in vitro* RNA synthesis and once again the full length and terminated probes purified on a denaturing polyacrylamide gel.

Protein extract used in the following bandshifts was purified on a Q-sepharose column and a similarly treated extract of *E. coli* lacking AmiR was included as a control. The protein concentration of the extracts was equalized after estimation using the Biorad Miniprotein assay. In the bandshifts presented on Fig 5.16, the full length probe was used. The AmiR extract was diluted in the control extract such that all incubations had the same protein concentration with lane 1 having 5 µl of the control extract with no AmiR and lane 5, 5 µl of AmiR alone. The figure shows complex formation in lanes 4 and 5 which represent 3 µl and 5 µl of AmiR.

Thus the AmiR binding site lies between the *Fsp*I target just downstream of the transcription start site and the *Ban*II target (possibly encompassing it) on the upward face of the transcription terminator. These experiments however, do not determine the precise boundaries of the binding site.

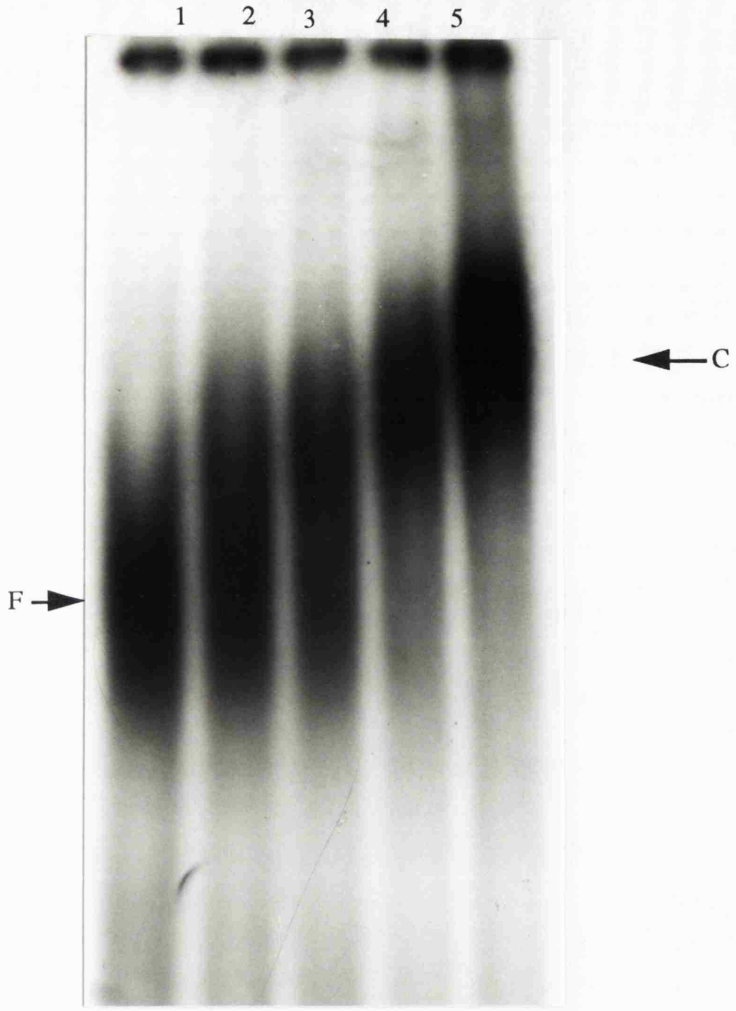


Fig 5.16 Bandshift assay for complex formation between AmiR and *Fsp*I (140)/*Pst*I (450) transcript.

Internally labelled transcript was incubated with an AmiR containing extract diluted in control extract to maintain equal protein concentrations. The incubation mixtures were separated on a nondenaturing 5% polyacrylamide gel. Lanes:1) transcript incubated with 5 ul control extract; 2) transcript incubated with 4 ul control extract and 1 ul AmiR; 3) transcript incubated with 3 ul control extract and 2 ul AmiR; 4) transcript incubated with 1 ul control extract and 4 ul AmiR; 5) transcript incubated with 5 ul AmiR extract. The arrow F represents free probe and arrow C the RNA/protein complex.

The 3 µl AmiR 'concentration' was chosen for a cold specific competition. Large scale cold RNA was synthesized as described and the DNA template removed by digestion with RNAse free DNAse (Promega). The synthesis mixture was then phenol:chloroform extracted, ethanol precipitated dried and resuspended in DEPC treated sterile water. RNA bandshift reactions were set up with increasing amounts of cold specific RNA with the total yeast RNA concentration being maintained as above. The result of the cold competition is shown in Fig 5.17. Lane 1 was loaded with the incubation mixture without protein, and represents the distance of migration of free probe. Lane 2 incubation mixture contained 3 µl AmiR extract and complex formation is evident (top arrow). 520 ng, 2.6 µg and 5.2 µg of cold specific RNA were added to the bandshift incubations loaded in lanes 3, 4, and 5 at the same time as the probe. Total competition was observed with 2.6 µg (about x 500) of specific RNA against a background of 80 µg of total yeast RNA. Panel 2 of the same figure shows the same experiment with the AmiR extract replaced by the control extract. There is no evidence of complex formation that could be competed out and all the lanes are identical with all bands running at the same place as the free probe (lane 1).

A protein titration carried out with probe made from *Fsp*I linearized pMW42 did not show any evidence of complex formation even at the maximum protein concentration used in the above experiments. These results together with those described previously show conclusive evidence for the physical interaction of leader RNA and AmiR protein. It is proposed that a binding site(s) for AmiR lies between the *Fsp*I (140) and *Bst*EII (206) targets. Some sequence features in nucleic acids that provide tell tale signals for regulatory proteins include, palindromes, repeats and RNA secondary structures. The sequences downstream of the *Fsp*I target including the terminator was inspected for these features. No obvious RNA

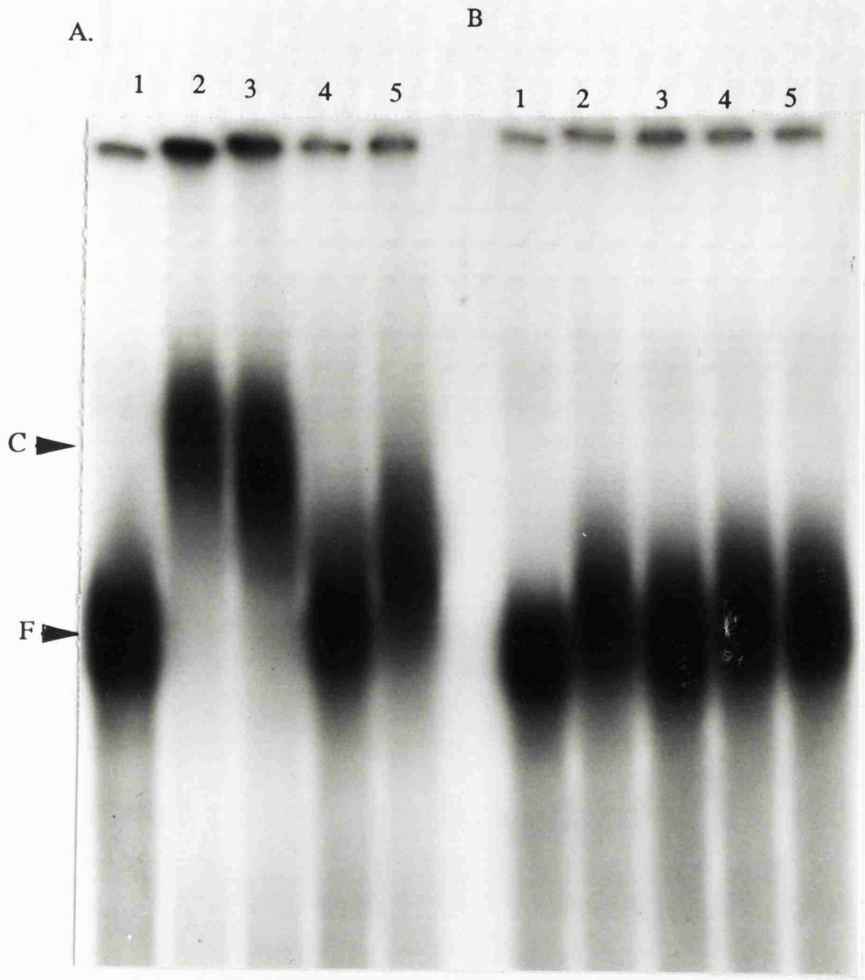


Fig 5.17 Bandshift assay for cold specific competition of 'AmiR/*Fsp*I (140)-*Pst*I (450) complex.

A) Labelled transcript was incubated with AmiR in the presence of increasing amounts of cold specific RNA and separated on a nondenaturing 5% polyacrylamide gel.Lanes: 1) free transcript; 2) transcript incubated with AmiR extract; 3) transcript incubated with AmiR extract in the presence of 520 ng cold specific RNA; 4) transcript incubated with AmiR extract in the presence of 2.6 ug cold specific RNA; 5) transcript incubated with AmiR extract in the presence of 5.2 ug cold specific RNA; B) In panel B, the AmiR containing extract was substituted with an AmiR free *E. coli* control extract. The free transcript is indicated by F and the complex by C.

secondary structures emerged when the RNA Fold program was used, however three homologous sequences (bold in Fig 5.13) were identified and it was initially considered that they might be AmiR binding sites. In the next Chapter two of these sites were changed by *in vitro* mutagenesis to investigate this possibility.

An experiment was designed to try and map the AmiR binding site by RNA footprinting. A 100 µl reaction was set up using the pMW42 *Pst*I linearized template. The total RNA preparation was dephosphorylated using CIP as described elsewhere then purified on a denaturing polyacrylamide gel with HpaII digested pBR322 being used as markers. The RNA band and markers were visualized under UV light by staining with ethidium bromide. The RNA was eluted as described elsewhere and end labelled with γ- ^{32}P ATP using PNK. The end labelled probe was checked for integrity by running on denaturing polyacrylamide gel and some minor degradation was observed. The probe was titrated with RNAse I (10 u/µl) diluted in the range .5 to .001 u/µl to establish optimal concentration of the enzyme required to obtain a single hit ladder. Optimization was never achieved (data not shown) which could have been caused by extensive secondary structure denying RNAseI equal access to all parts of the sequence and lack of time prevented further investigations.

5.4 Summary

Results of the *in vivo* competition show that production of the leader both in *cis*- and *trans*- to *amiE* in a constitutive system reduces the amount of AmiR available for expression of amidase. This experiment is analogous to one reported by Houman and others (1990) working with the *bgl* operon of *E. coli* where production in *trans*- of the BglG binding site reduced the expression of β–galactosidase reporter gene by as much as 15 fold. In the

amidase system this titration of transcription antiterminator was not as great which may reflect differences in affinities of BglG and AmiR for their respective binding sites.

The *amiE* leader/*amiR* fusion deletion series shows a large increase in amidase expression in pMW47 when compared to pMW46 (Table 5.2). A further deletion (pMW48) does not show such a profound effect indicating that the potential binding site lies 5' to the *Bst*EII target (206) and probably overlaps the *Ban*II target (198).

A gel retardation assay has been established with labelled RNA runoff transcripts and semipure AmiR. Cold competition studies showed that this was a specific binding and use of the leader sequence deletions show that AmiR binding has been delineated between the *Ava*II (161) target at the 5' end and the *Ban*II (198) target at the 3' end. Unfortunately a lack of time prevented the footprinting studies from being carried out to further define the AmiR binding site. The possibility that AmiR recognizes an RNA secondary structure similar to one recognized by the proposed functional homologs, the BglG family of transcription antiterminators, was considered and the potential amidase leader RNA folds generated. This and inspection by eye ruled out such a possibility. Information on the structure of RNA/protein complexes and recognition patterns is as of now limited despite the importance of these biomolecular structures in the metabolism of nucleic acids.

Gel retardation experiments were carried out with the addition of purified AmiC and inducing and repressing amides (data not presented) and these results are discussed in Chapter seven.

CHAPTER 6

In vitro mutagenesis studies of the leader region

6.1 Introduction

Wild type amidase expression requires activation by a diffusible positive regulator, AmiR (Farin and Clarke, 1978; Drew and Lowe, 1989; Cousens *et al.*, 1987) and an AmiR binding site has now been identified in the *amiE* leader region (Chapter 5). Transcription antitermination as it is now understood is effected by two distinct mechanisms (see Introduction). The first is exemplified by the N-antitermination system of λ phage in which case the N protein captures RNA polymerase at a specific RNA hairpin, the *nut* site, by looping the intervening RNA sequence between its binding site and the paused RNA polymerase. This leads to a cascade of events where several host proteins are recruited and the RNA polymerase is rendered insensitive to transcription terminators located some distance away (Das, 1993). The *E.coli* BglG protein is the archetype of the alternative mechanism. It binds to the leader region of the *bgl* operon just upstream of the transcription terminator and stabilizes an alternative secondary structure which preempts formation of the terminator hairpin loop (Houman *et al.*, 1990). There is no evidence that AmiR functions in an analogous manner to λ N protein. However, the amidase leader region shows limited homology to the BglG recognition sequence and in addition is involved in control of a catabolic operon. It was thus considered that the mechanism of *bgl* operon regulation would serve as the best model for further investigations of the amidase regulatory system.

Within the amidase leader region is a short open reading frame of 35 amino acids (Drew and Lowe, 1989) which is reminiscent of the amino acid biosynthetic operons of enteric

organisms regulated by transcription attenuation (Yanofsky, 1988). The aims of the following experiments were to determine the function, if any, of the open reading frame and to identify sequences required for AmiR mediated transcription antitermination using *in vitro* mutagenesis.

6.2 Principles of oligodeoxyribonucleotide directed mutagenesis

In vitro mutagenesis has become an invaluable technique in molecular biology because it allows the investigator to effect precise changes within genes and their regulatory elements in a predetermined way. Several methods have been developed for this and they include, total synthesis of the gene fragment in the form of oligonucleotides with the necessary changes followed by annealing and cloning (cassette mutagenesis), use of the PCR technique to synthesize the gene of interest from selected oligonucleotides and the oligonucleotide heteroduplex mutagenesis used here.

The biochemical procedure of the heteroduplex technique involves: annealing of the mutagenic oligonucleotide to a complementary region on a single stranded template to form a heteroduplex; use of the oligonucleotide to prime synthesis of the second strand *in vitro* and selection of the mutant phenotype *in vivo* (Hutchison *et al.*, 1978; Razin *et al.*, 1978). Genotypic selection can be carried out by: restriction analysis in cases where a restriction site has been created or destroyed; southern hybridization preferably using short oligonucleotides based on the difference in melting temperatures between the wild type and heteroduplexes; and by direct sequencing when the mutation efficiency is high enough to make it feasible.

The success of the technique is dependent on the choice of the DNA polymerase used for

second strand synthesis for three reasons (Smith, 1985). Some DNA polymerases are unable to synthesize long stretches of DNA without pausing at regions of strong secondary structure. This can be solved by addition of single stranded binding protein but has the limitation of displacement of the mutagenic oligonucleotide as well. The enzyme should have high fidelity to avoid introducing unwanted mutations and should not have strand displacement activity which can be a problem whether one or two oligonucleotides are used. The strand displacement activity can be controlled by using higher ligase to DNA polymerase activity which ensures ligation of the growing 3' terminus to the 5' end of the oligonucleotide before displacement can occur. The bacteriophage T4 DNA polymerase has been found to be superior to Klenow fragment of *E.coli* DNA polymerase I for its lack of strand displacement function (Masamune and Richardson, 1971; Wassal, 1974).

Single strand templates for *in vitro* mutagenesis are prepared from recombinant filamentous phages. Initially ϕ X174 was used (Hutchison *et al.*, 1978; Razin *et al.*, 1978) but currently M13 derivatives are preferred because of the histochemical selection of clones, presence of an engineered multiple cloning site and possibility of using the universal sequencing primers. Chimeric plasmids with a filamentous phage origin of replication (Phagemids) (Dente *et al.*,1983; Geider *et al.*, 1985) have proved even more valuable because cloned genes can be maintained as plasmids for expression and other studies and then be induced to secrete one strand by infection of the *E.coli* F' carrying host with the appropriate helper bacteriophage.

The theoretical 50% mutant yield based on the semiconservative DNA replication mechanism is normally not achieved because of selection against the unmethylated *in vitro* synthesized strand *in vivo*. This directed 6-N-methyladenine methylation by the *dam*

methylase is to ensure that the parental strand is always methylated during replication as a marker during excision repair of mismatches (Friedberg, 1985). Mutant yields can therefore be increased by transformation of the newly synthesized heteroduplex into *dam-E. coli* hosts, or a host defective in the excision repair mediated by the *mutL, mutS* or *mutH* loci. Low frequency mutant yield is also sometimes caused by a mechanism that biases replication completely in favour of one or the other of two strands depending on the distance of the mismatch from the origin of replication (Smith, 1985).

Refinement of the technique has been achieved by production of the template in a uracil DNA glycosylase defective (*dut-, ung-*) *E.coli* (Kunkel, 1985) host which incorporates 1% of the T residues positions with uracil (Friedberg, 1985). Subsequent transformation of the *in vitro* synthesized duplex into repair competent *E.coli* results in the preferential degradation of the uracil containing strand. An alternative is to use a template in which a selectable marker such as an antibiotic resistance has been mutated by introduction of nonsense codons. A repair oligonucleotide is then annealed together with the mutagenic oligonucleotide during *in vitro* synthesis.

The site directed mutagenesis in the following sections was carried out using the 'Altered sites™ *in vitro* mutagenesis system of Promega' . The steps involved are summarised in the flow diagram Fig 6.1.

6.3 Effect of mutations that alter the leader region open reading frame

The role of the *amiE* leader region open reading frame was investigated by firstly introducing a stop codon after 8 amino acids (161) and secondly by insertion of a T between bases 163 and 164 which extended the ORF into the *amiE* gene out of frame. To

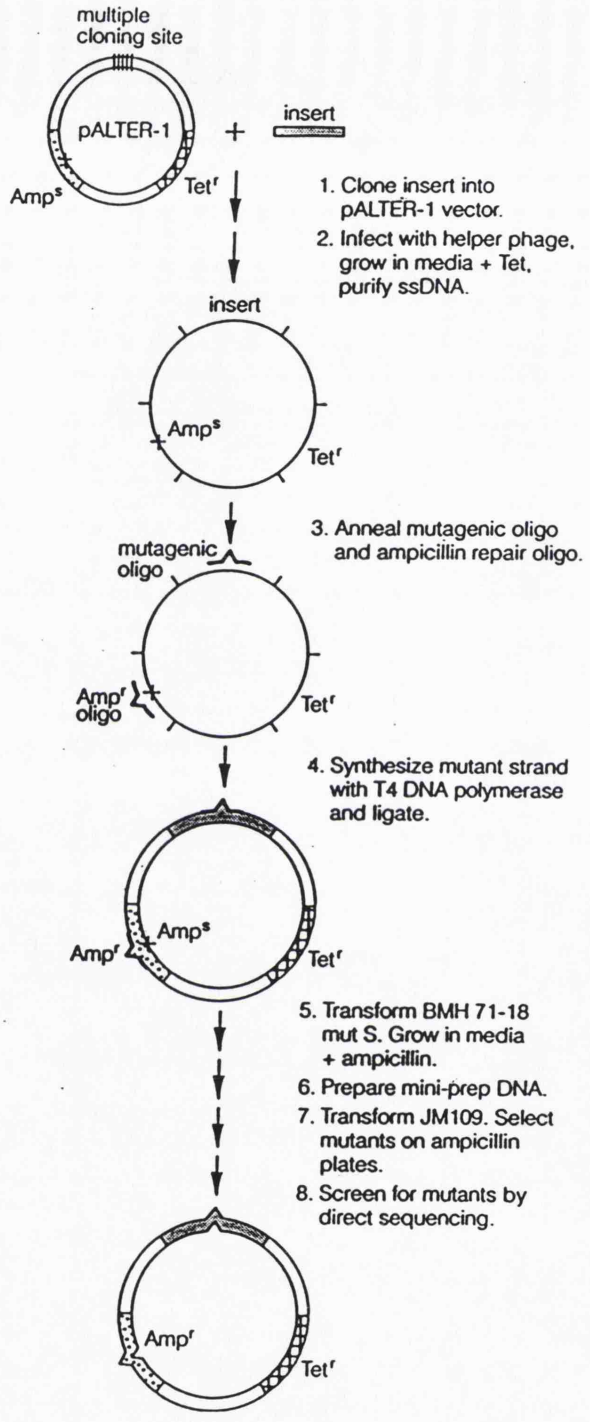


Fig 6.1 Schematic diagram of the site directed mutagenic steps. Phagemid pALTER-1 has an f1 origin of replication.

make the mutagenesis template, the 450bp *Hin*dIII/*Pst*I fragment from plasmid pJB950 was excised from an agarose gel and purified by elution from glass milk using the Geneclean kit. It was then cloned into the phagemid pAlter™ to create plasmid pSB100 (Fig 6.2). The ligation mixture was used to transform competent *E.coli* JM109 to tetracycline resistance. A plasmid with the correct insert was identified from resistant colonies by restriction enzyme analysis of isolated plasmid DNA with *Hin*dIII and *Pst*I. Phagemid pAlter™ has ColE and f1 origins of replication and can, therefore, be induced to produce one packaged DNA strand by infection of the *E.coli* host with either R408 or M13KO7 bacteriophages. To make single stranded DNA, JM109 pSB100 was infected with the helper phage R408 (Dotto *et al.*, 1984) as described earlier and growth continued for a further 8 hours (Promega, technical manual). The culture was centrifuged and phages recovered from the supernatant by PEG precipitation, lysed with chloroform:isoamyl alcohol and SS DNA purified as described earlier before resuspension in 20 µl sterile water. The phage DNA concentration was estimated by AGE of a sample as 0.1 mg/ml.

To introduce a stop codon at position 161, an *in vitro* synthesized oligonucleotide (5'AGGTTCGGTCACGCGACATCG) was phosphorylated using T4 PNK and 10 ng of this reaction mixture was annealed together with 10ng of the ampicillin repair oligonucleotide provided (5'-d[pGTTGCCATTGCTGCAGGCATCGTGGTG]) to 100 ng of the SS pSB100 DNA. The *in vitro* second strand synthesis reaction was carried out as recommended by the manufacturers and the entire synthesis reaction transformed into competent *E.coli* BMH 71-18 *mutS* . This strain carries a *mutS* mutation which suppresses mismatch repairs hence increasing the mutant yield. The transformation protocol was essentially identical to that described elsewhere with two variations, 4 ml of prewarmed LB was added to the cells after the heat shock and cooling step and ampicillin (final

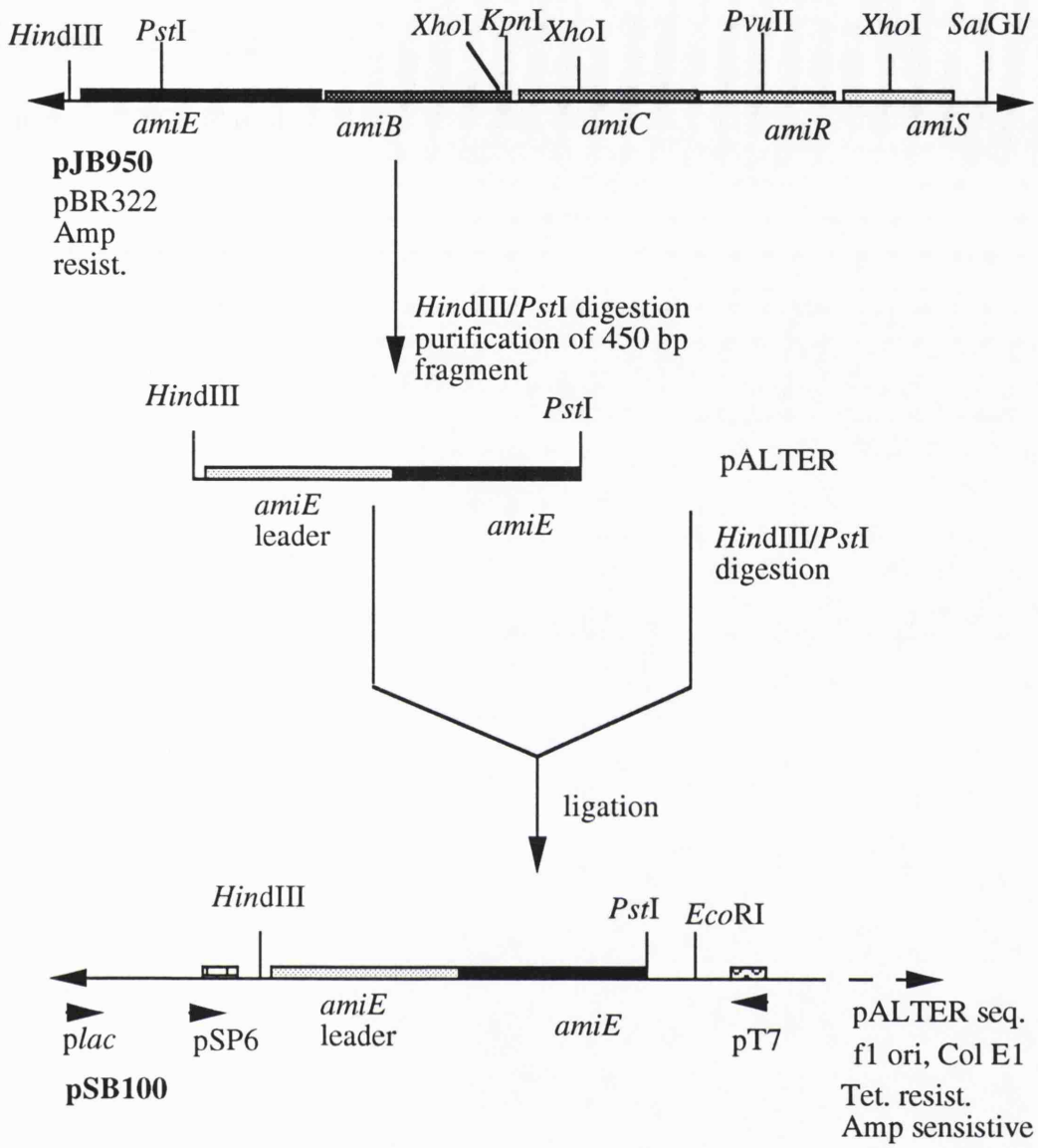


Fig 6.2 Construction of pSB100. The amidase leader region carrying the first 63 *amiE* codons was cloned into pALTER for making SS *in vitro* mutagenesis templates

concentration of 125 μg/ml) was added directly to the cells after recovery prior to aeration overnight at 37°C. Plasmid DNA was isolated from the overnight cultures and transformed into *E.coli* JM109 with selection for tetracyline/ampicillin resistant transformants. Plasmid DNA isolated from the resistant colonies was subjected to DS sequencing with the reverse primer and the plasmid, pMW6, with a stop codon in the short ORF identified. No other mutations within this region of the leader were observed.

Carrying out mutagenesis on a short gene fragment has the advantage of reducing the chance of unwanted mutations occurring elsewhere. Plasmid pSB100 and consequently pMW6 has only the first 63 amino acid coding region of the *amiE* gene. To reconstruct the rest of the gene, the pJB950 1.9 kb *Pst*I/*Kpn*I DNA fragment which carries the rest of the *amiE* gene and part of the downstream *amiB* gene was purified from an agarose gel by elution from glass milk as described. The fragment was cloned into pMW6 which had been digested with *Kpn*I and *Pst*I. The ligation mixture was transformed into *E.coli* JM109 and plasmid DNA was isolated from ampicillin resistant colonies and restriction enzyme mapped with *Hin*dIII and *Eco*RI. Plasmid pMW11 (Fig 6.3) was identified.

Plasmid pSB100/1/2 was made from pSB100 by *in vitro* mutagenesis in an identical way to pMW6 except that the mutagenic oligonucleotide (5'-GTTAGGTTCGGATCCCGCGACAT-3') had a T inserted between bases 163 and 164. This creates a unique *Bam*HI site in addition to extending the leader ORF into the *amiE* gene. Potential pSB100/1/2 mutants were initially identified by restriction enzyme analysis with a) *Pst*I which gave a 4.45 kb and a 1.67 kb fragment confirming the presence of the repair oligonucleotide which restores the *Pst* I site within the *bla* gene and b) *Bsc*I/*Bam*HI double digests. Single strands were made from JM109 pSB100/1/2 by infection with

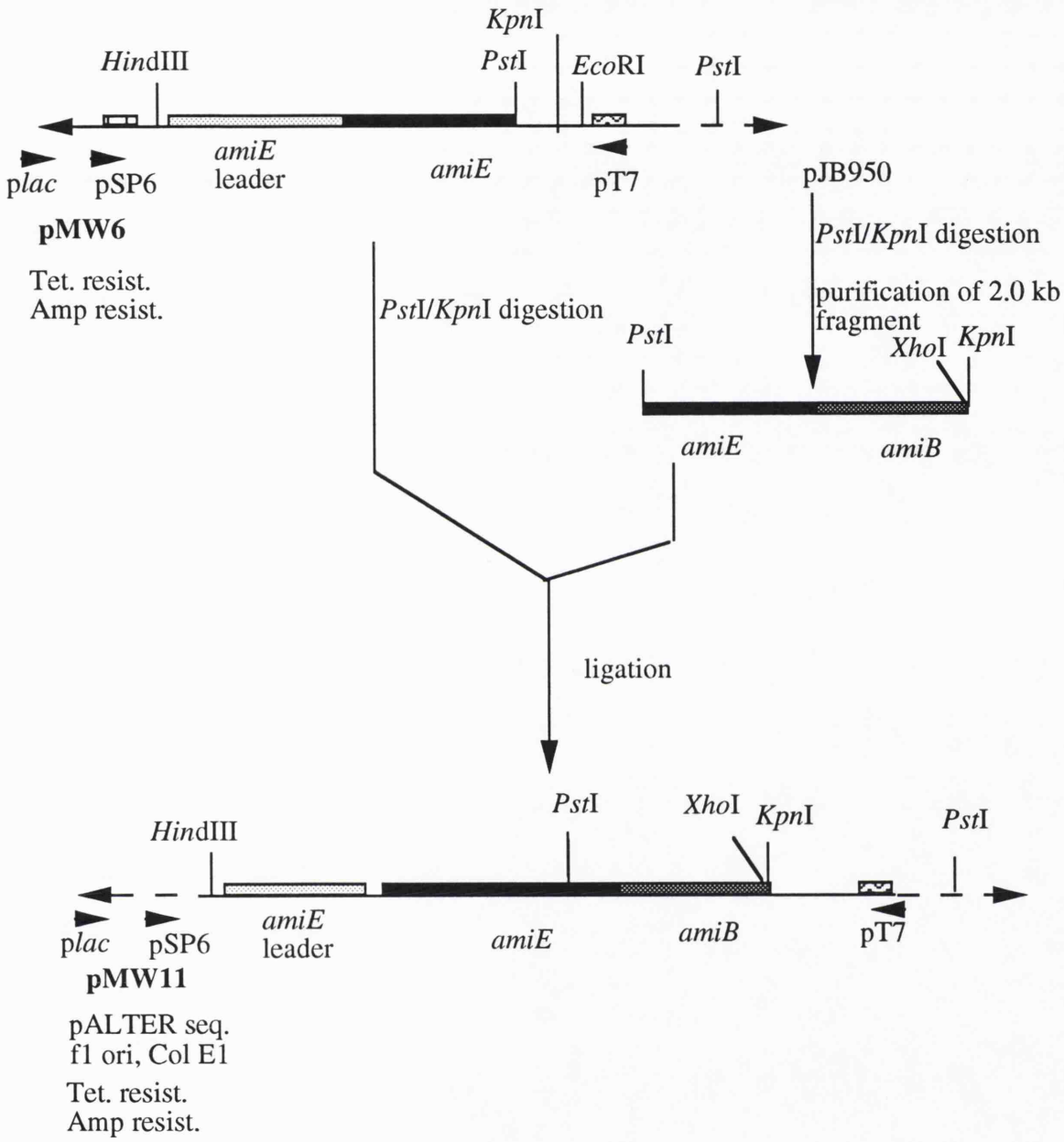


Fig 6.3 Construction of plasmid pMW11. *amiE* gene was reconstructed by cloning the 2.0 *Pst*I/*Kpn*I fragment into mutagenised pSB100.

phage R408 and purified as described previously for sequencing using the M13 universal sequencing primer. The mutation was confirmed free of other unwanted changes within the leader region around the oligonucleotide complementary site. Plasmid pSB102 was made from pSB100/1/2 by reconstructing the *amiE* gene in an identical way to pMW11.

A control plasmid pSB101 was constructed by cloning the 2.4 kb *Hin*dIII/*Xho*I fragment from plasmid pJB950 into the *Hin*dIII/*SalGI* site p-Alter™. Plasmid pSB101 was characterised by restriction analysis with *Hin*dIII and *Eco*RI.

Plasmids pSB101, pSB102 and pMW11 were transformed into *E.coli* JA221 and JA221 pSW35 (*amiR*) (Wilson, 1991) and amidase assays carried out using the hydroxamate transferase assay (Table 6.1 A). Plasmid pSB101 the wild type gave an average of 0.8 units by itself and 44.3 units in the presence of pSW35 showing the expected *trans*-complementation by pSW35. The stop codon leader mutant pMW11 gave 4.6 units by itself suggesting considerable readthrough of the terminator which is located 30 bp downstream of the mutation point. Plasmid pSW35 is able to *trans*-complement pMW11 producing 65.3 units of activity. A similar result was observed with pSB102 which gave 2.2 units by itself and 59 units when *trans*-complemented by *amiR* . These results show that the leader ORF plays no part in the antitermination reaction. The increased readthrough transcription found with pMW11 and pSB102 in the absence of pSW35 cannot be explained at this stage.

6.4 Mutations within the leader terminator sequence

The amidase leader terminator resembles a consensus *E.coli* ρ–independent transcription terminator. These are the most well characterised terminators and they consist of a G+C

rich dyad symmetrical sequence which encodes an RNA hairpin structure followed by a continuous tract of Us where the RNA polymerase terminates. These features are recognised as important for efficient termination (Farnham and Platt, 1982; Bremmer and Trifonov, 1984; Platt, 1986) and a mechanism has been proposed whereby formation of the RNA hairpin causes the RNA polymerase to pause with transcript release being facilitated by the unstable nature of the U:A transcript/template interactions (Farnham and Platt, 1980).

To check whether the terminator sequence itself was involved in the transcription antitermination reaction, a mutation was designed which introduced two *Sst*II targets within the terminator stem regions. These changes replacing the two *Sma*I targets (CCCGGG) with *Sst*II targets (CCGCGG) introduced two changes in each of two of three potential AmiR recognition sequences (TNt/aCg/cCGGGR). These sequences (highlighted Fig 6.4) were identified as potential binding sites by sequence inspection simply because they occurred three times in the same region. The *Sst*II changes preserve the G:C base pairing of the terminator loop by swopping a G/C pair on the upward face with one from the downward face to avoid changes in the terminator stem stability. The mutagenesis was carried out on plasmid pSB100 following the procedure described above. Plasmid pMW9 was identified after the second round of transformation in *E.coli* JM109 by restriction enzyme analysis with *Sst*II/*Bsc*I and *Sst*II/*Eco*RV. DS sequencing using reverse primer was performed to cover the mutagenesis region and confirmed the sequence. The amidase gene was reconstructed in pMW9 in an identical way to pMW11 (above) by cloning in, the 2.4 kb *Pst*I/*Kpn*I fragment from pJB950. 3/10 ampicillin resistant colonies were found to have the correct insert by restriction analysis with *Hin*dIII and *Eco*RI and one pMW10 chosen for amidase assays in *E.coli* JA221 (Table 6.1 C).

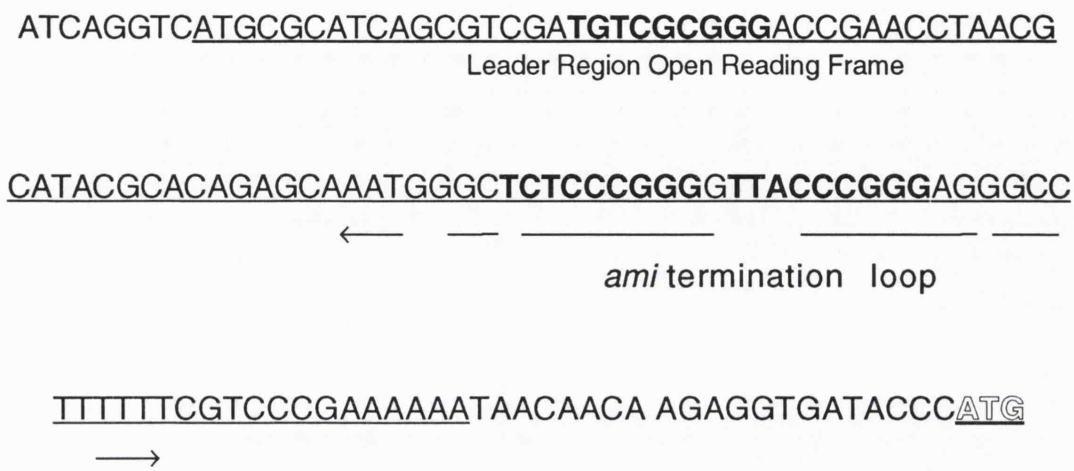


Fig. 6.4 DNA sequence of the amidase leader transcript showing the three potential AmiR recognition sequences.
The leader ORF is underlined, the *amiE* start codon is underlined and shown in outline, the termination loop is shown by the broken arrows and the potential AmiR recognition sequences are shown in bold.

Plasmid pMW10 gave a low constitutive AmiR independent amidase activity of 2.9 units compared to the parent plasmid pSB101 (0.8 units). This could mean a small reduction in the overall stability has occurred but it seems unlikely since the calculated free energy of formation for both termination loops is the same. In the presence of pSW35, (*amiR*) an activity of 67.6 units were obtained about 1.5 x the pSB101, pSW35 activity (44.3 units). Again this increase was attributed to the contribution of the AmiR independent terminator readthrough. These mutations were expected to be sufficient to affect AmiR binding if it recognized a sequence encompassing the stem however, recognition of the secondary structure *per se* was not ruled out. However, at this stage it appears that the three repeat sequences (Fig 6.4) are not involved in AmiR mediated antitermination.

6.5 Mutagenesis of the BglG leader homology region

The transcriptional antitermination mechanism regulating amidase expression was proposed after DNA sequencing, analysis of the *amiE* upstream leader region and construction of a deletion mutation within the proposed terminator (Drew and Lowe, 1989). This study also identified a region upstream of the terminator which showed homology to the BglG leader region where a 6 bp insertion eliminated BglG mediated antitermination (Mahadevan and Wright, 1987). More recent studies have identified other regulated rho-independent terminators with homologies, both in sequence and position in upstream regions: STB1 and STB2 in the *arb* operon of *E. chrysanthemi* (Hassouni *et al.*, 1992) T1 and T2 in the *bgl* operon of *E. coli* (Houman *et al.*, 1990); the *B. subtilis* upstream terminators of the *bgl* operon, endoglucanase gene, and levansucrase gene (Aymerich and Steinmetz, 1987; Shimotsu and Henner, 1986) and the sucrose metabolic operon *sacPA* of *B. subtilis* (Debarbouille *et al.*, 1990). The proposed antiterminator motifs (RAT sequences) are of approximately 30 nucleotides and consist of two broken imperfect palindromes termed

boxA and boxB (Hassouni *et al.*, 1992) with boxB partially overlapping the terminator stem. A mechanism has been proposed whereby the RNA transcripts of these sequences fold into similar secondary structures stabilised by the respective protein factor to preempt terminator formation (Houman *et al.*, 1990; Aymerich and Steinmetz, 1992). It was of great interest to investigate the equivalent amidase region not least because the homologous region lay within the AmiR binding region as determined by the gel shift assays (Chapter 5).

The creation of a terminator codon in plasmid pMW11 changed nucleotide 161 from G to T which was identified as a *Bgl* homologous nucleotide (Drew and Lowe, 1989) without dampening the antiterminator activity of AmiR. To further investigate the antitermination reaction a series of changes were made to the leader region by construction of *Bam*HI restriction target sites at various locations by site directed mutagenesis. Initially the region immediately downstream of the GGA triplet (CCGAACCTA) was mutated to CGGATCCTA. This had the advantage of creating a unique *Bam*HI site which could be used for initial mutant screening. All the experimental details for mutagenising plasmid pSB100 to plasmid pSB100/1/3 and the ensuing *amiE* reconstruction to plasmid pSB103 were identical to those of plasmid pSB102. The mutagenic oligonucleotide effected two substitutions in pSB103 which are shown in bold and had no effect on the leader ORF.

Plasmid pSB103 was transformed into *E. coli* JA221 and JA221 pSW35 for amidase assays. The results (Table 6.1 B) showed a 40% reduction (25.8 units of activity) in amidase expression in the presence of AmiR compared to the wild type pSB101, pSW35 (44.3 units) value without an effect on the basal level terminator readthrough. This result provided the impetus for introducing mutations along the sequence up to the terminator start site in order to further investigate the AmiR binding site. Mutations leading to the creation

	Plasmid(s)	Specific Activity
A.	pSB101	0.8
	pSB101, pSW35	44.3
	pMW11	4.6
	pMW11, pSW35	65.3
	pSB102	2.2
	pSB102, pSW35	59.1
B.	pSB103	0.9
	pSB103, pSW35	25.8
	pMW12	0.2
	pMW12, pSW35	3.8
	pMW13	0.4
	pMW13, pSW35	48.0
	pMW25	0.2
	pMW25, pSW35	32.1
	pMW26	0.2
	pMW26, pSW35	21.2
	pMW27	0.1
	pMW27, pSW35	1.7
	pMW28	0.1
	pMW28, pSW35	0.7
C.	pMW10	2.9
	pMW10, pSW35	67.6

Table 6.1. Amidase assays in *E. coli* carrying recombinant plasmids.

Effect on AmiR dependent transcription antitermination by changes that: A) disrupt the leader ORF; B) change the *bgl* leader homologous sequences and C) conservative changes within the terminator. The amidase assay results are presented as µl mole of acetyl hydroxamate/min/mg of bacteria

of *Bam*HI sites were once again chosen not only to simplify the initial mutant screening but to also as a basis for creation of novel deletions and insertions by shuffling fragments from the newly made mutants.

Four oligonucleotides ,

[5'-GCTCTGTGCGTATGGATCCGGTTCGGTCCCGC-3']

[5'-CATTTGCTCTGTGCGGATCCGTTAGGTTCGGT-3']

[5'-AGAGCCCATTTGCTGGATCCGTATGCGTTAGG]

[5'-CGGGAGAGCCCATTGGATCCGTGCGTATGCGTTA]

were used in site directed metagenesis of plasmid pSB100 as described above and plasmids pMW17, 14, 15, and 16 initially identified by restriction enzyme mapping plasmid DNA from ampicillin resistant colonies with *Bsc*I and *Bam*HI after the second round transformation into *E.coli* JM109. The correct sequences were then confirmed using DS sequencing with reverse primer. The *amiE* gene was reconstructed by cloning in the 2.4 kb *PstI/Kpn*I pJB950 fragment as described above and the new plasmids pMW25, pMW26, pMW27, and pMW28 identified by *Hin*dIII/*Eco*RI restriction enzyme mapping.The plasmids were then transformed into *E.coli* JA221 and JA221 pSW35 for amidase assays as described previously (Table 6.1 B).

E. coli pMW25 gave 0.22 units of activity by itself and 32.1 units with pSW35 showing a 30% reduction in AmiR antitermination activity compared to pSB101 pSW35 although this particular change involved 4 base substitutions. Plasmid pMW26 averaged 0.21 u on its own and 21.22 u with plasmid pSW35 a 48% reduction. The two changes in pMW26 are within the GCAT BglG leader homologous sequence (Fig 5.13) which corresponds to the terminal (hairpin) loop of the BglG RAT (Amster-Choder and Wright, 1993).

Interestingly, changes within these bases have been shown to lead to increased BglG binding resulting in higher activity of the β-galactosidase reporter gene (Aymerich and Steinmetz, 1992).

Plasmid pMW27 gave 0.11 units of activity on its own and an average of only 1.66 units with pSW35 which represents 4% of the pSB101 pSW35 (44.3 u) value. A similar situation was observed with plasmid pMW28 which gave 0.11 units on its own and 0.72 units with pSW35. These mutations almost totally eliminate AmiR mediated antitermination without interfering with the terminator structure.

6.6 Generation of novel deletions/insertions

Two other mutations constructed within this region were a deletion and an insertion. To make plasmid plasmids pMW12 and 13, plasmids pSB102 and 103 were both digested with *Bam*HI and *Eco*RI and the fragments separated by AGE. All the four fragment bands were excised under long wave UV lamp and purified by elution from 'Glass milk'. The 2.4 kb fragment of pSB102 was ligated to the 5 kb fragment of plasmid pSB103 for pMW13 construction and the reverse situation generated pMW12. This in effect created a 4 bp deletion between positions 162 and 166 plus an A to T substitution at position 167 in plasmid pMW12 and a 5 bp insertion (-CGGAT-) between positions 162 and 163 for pMW13. The two constructs were confirmed by DS sequencing using the reverse primer and transformed into *E.coli* JA221 and JA221 pSW35 for amidase assays (Table 6.1 B).

JA221 pMW12 gave 0.2 units of activity less than the pSB101 activity (0.8 u) and 3.8 units with plasmid pSW35 a greater than 10 fold reduction compared to the 44.3 units of JA221 pSB101 pSW35. This mutation therefore drastically reduces transcription

antitermination and to lesser extent increases termination efficiency in the absence of AmiR despite being 22 bp upstream of the terminator. JA221 pMW13 gave 0.4 u of activity and 48 u in the presence of plasmid pSW35 a value which is very similar to plasmid pSB101 despite the 5 bp insertion. This would indicate that the potential AmiR binding site lies 3' to the BglG leader homologous triplet GGA.

Unfortunately, a lack of time prevented further reconstructions from being made and analysed.

6.7 Gel Retardation Experiments with Mutant Leader Sequences

In the previous section, *in vitro* mutagenesis was used to isolate mutants in which the AmiR dependent antitermination reaction was disrupted. Gel retardation experiments were carried out with *in vitro* synthesized transcripts from two of these mutants. Plasmids pMW12 (the 4 bp deletion mutant) and pMW15 (parent of pMW27) were linearized with *Bst*EII and used as templates for *in vitro* synthesis of labelled ([α-^{32}P]CTP) transcripts using SP6 RNA polymerase as described in Chapter 5. The transcripts were isolated from a urea denaturing 6% polyacrylamide gel and used for gel retardation with an AmiR extract eluted from a Q-Sepharose column as described in Chapter 5. The resulting bandshifts are shown in Fig 6.5. Lanes 1-5 show tracks loaded with reactions from the deletion mutant (pMW12). Two bands are seen in lanes 1 and 2 which were loaded with protein free control reaction and an AmiR free *E. coli* protein extract control. It is likely that these two bands represent two structures of RNA. Lanes 3,4, and 5 were loaded with reactions containing 1 µl, 2 µl, and 4 µl of AmiR. There is no evidence of bandshifts with the 1 µl AmiR reaction, some complex formation is evident with the 2 µl AmiR reaction and about

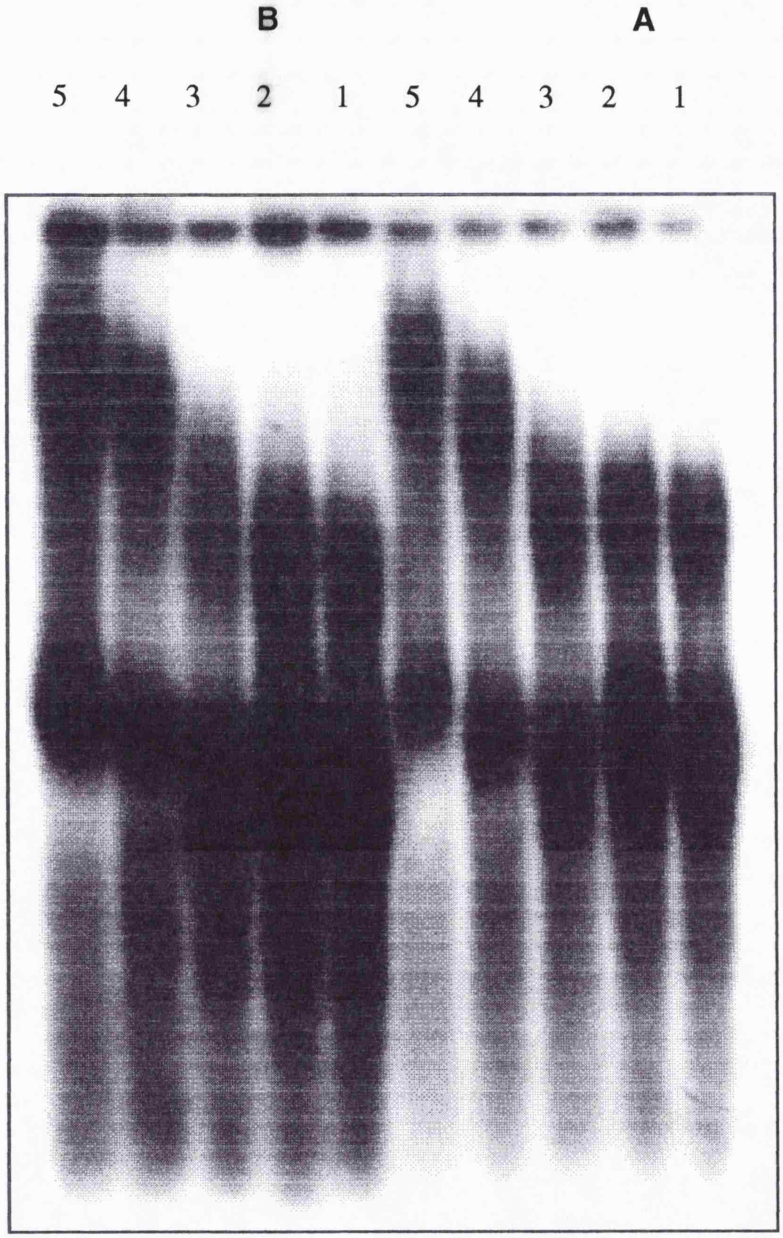


Fig 6.5 RNA bandshift assay of leader mutants/AmiR complex formation.
In vitro synthesized RNA was incubated with AmiR extract and separated on a nondenaturing 5% polyacrylamide gel. Panel A was loaded with deletion mutant (pMW12) reactions and panel B with identical reactions of the substitution mutant (pMW15). A Lanes: 1) RNA in the absence of protein; 2) RNA incubated with AmiR free protein control extract; lane 3) RNA incubated with 1 µl AmiR extract; lane 4) RNA incubated with 2 µl AmiR extract and lane 5) RNA incubated with 4 µl AmiR extract

75% of the RNA is complexed with 4 µl of AmiR. Lanes 6-10 were loaded with identical reactions carried out with RNA from the base substitution mutation (pMW15). Again in the absence of AmiR (lanes 6 and 7) two bands are seen and complex formation is evident at 2 µl AmiR (lane 8) and about 50% of the transcript is complexed with 4 µl AmiR (lane 10). This, a single one-off experiment was carried out under pressure of time and the relative amounts (affinities) of AmiR required for complex formation between these and the wild type leader sequence were not determined. It is however apparent that despite great reduction in AmiR dependent antitermination in the mutants there is still sufficient recognition between AmiR and the mutant leader RNAs to cause bandshifting.

6.8 Summary

Attempts to disrupt the short leader open reading frame by creating a translation stop codon (TGA) (plasmid pMW11) unexpectedly led to an increase in the background constitutive amidase activity with no apparent effect on transcription antitermination (Table 6.2). A similar observation was made with the 1bp insertion variant, plasmid pSB102, that extends the leader reading frame into *amiE*. These two mutations are unlikely to have any direct effect on the hairpin terminator stability, a basic tenet for the rho-independent termination (Yager and Von Hippel, 1987). This observation is however, not without precedent. Telesnitsky and Chamberlin (1989) have shown that the efficiency of transcription termination at well defined Rho-independent terminators is dependent on the specific promoter and the readthrough effects map to within the first 25 bp of the 'initial transcribed region.' Using the P*tac* promoter, the distance between the promoter and the terminator could be greatly varied without any difference in the 'factor independent transcription antitermination. In *E.coli* bacteriophages λ and 83 late promoters, promoter associated

<<<<<<<<<<<<<<<<<<<<<<----->>>>>

GGA CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G

bgl homology region

Plasmid	Changes to Wild Type (bold)	Amidase Activity	
		-AmiR	+AmiR
pSB100	GGA CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	0.8	44.3
pMW11	**T**GA CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	4.6	65.3
pSB102	GGA **T** CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	2.2	59.1
pSB103	GGA **CGG** ATC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	0.9	25.8
pMW12	GGA **TC** CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	0.2	3.8
pMW13	GGA **CGGAT** CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	0.4	48.0
pMW25	GGA CCG AAC C**GG ATC** CAT ACG CAC AGA GCA AAT GGG CTC TCC CGG G	0.2	32.1
pMW26	GGA CCG AAC CTA ACG **GAT C**CG CAC AGA GCA AAT GGG CTC TCC CGG G	0.2	21.2
pMW27	GGA CCG AAC CTA ACG CAT ACG **GAT CCA** GCA AAT GGG CTC TCC CGG G	0.1	1.7
pMW28	GGA CCG AAC CTA ACG CAT ACG CAC **GGA TCC** AAT GGG CTC TCC CGG G	0.1	0.7
pMW10	GGA CCG AAC CTA ACG CAT ACG CAC AGA GCA AAT GGG CTC TCC **GC**G G	2.9	67.6

Table 6.2 Summary of mutations on the *ami* leader and their effect on antitermination by AmiR.

The changes are shown in bold face. Arrow heads mark the transcription terminator

transcription antitermination that works with different terminators has been mapped to within the first 30 bp transcribed region (Goliger *et al.*, 1989). In both these reports no obvious alternate secondary structures, the paradigm of transcription attenuation (Yanofsky, 1988), could be generated.

The results described in this section were all performed in *E. coli.* However , it seems unlikely that any major differences in antitermination reaction will be found in *Ps. aeruginosa.* One point that should be noted is that the mutations in plasmid pSB102 and pMW11 are located respectively 33 and 35 bp downstream of the transcription start site (Wilson, 1991) which could explain the modest terminator disruption if they were at the 3' limit of a regulatory region. In all the other constructs, the changes appeared to accentuate transcription termination.

This study has investigated the *amiE* leader region by mutagenesis to look for changes that disrupt transcription antitermination with respect to the *bgl* homology/RAT sequences. Many of the mutants constructed involved multiple changes and as such the results are difficult to interpret unambiguously. Further investigations will probably be necessary to define the bases that interact directly with AmiR in the antitermination reaction.

Plasmids pMW11, pSB102, and pMW13 all appear to antiterminate correctly and can therefore be used to define the transcription antitermination signal (TAS). This finding rules out a relationship between the *ami* TAS and the *bgl* homology region which includes residues 161-163. Changes downstream of residue 163 start to affect the antitermination reaction. The two changes in pSB103 CCGAAC to CGGATC cause a 40% reduction in antitermination and the 4 bp deletion in pMW12 reduces antitermination to less than 10% of the Wild Type level defining important interactions between this region of the leader with

AmiR in the antitermination reaction. Moving towards the termination loop the changes in pMW25 (CTA ACG to CGG ATC) and pMW26 (CAT ACG to GAT CCG) causes respectively 30% and 48% decreases in antitermination activity. This indicates that these sequences have some involvement in the antitermnation reaction. The final two changes in pMW27 (CAC AGA to GAT CCA) and pMW28 (AGA GCA to GGA TCC) reduce the antitermination reaction to almost zero. This indicates a major role of these sequences in the antitermination reaction.

The results presented in this chapter thus show that the changes to the *ami* leader in the RAT equivalent region disrupt the antitermination reaction. Recent studies of the *E. coli bgl* and *B. subtilis sacB* and *sacPA* systems have shown extensive homologies in their RAT sequences and suggested that the antitermination reaction may involve factor dependent low ΔG leader secondary structure formation which could preempt the formation of the terminator loop. It would appear unlikely that the amidase system functions in this way since no alternative secondary structure is possible with the *ami* leader region.

The single experiment in which AmiR was shown to bind to the mutant leader mRNAs which were disrupted in the antitermination reaction indicates that the reaction is a multistage process. It would appear that AmiR binding to the leader sequence is the primary event and there is at least one subsequent stage in the process for the antitermination reaction to proceed.

CHAPTER 7

Discussion

The work presented in this thesis has led to an increased understanding of the way in which the regulation of the amidase operon occurs in the following ways.

a) Only the two previously identified regulator genes, *amiR* (Farin and Clarke, 1975; Cousens *et al.*, 1987; Lowe and Drew, 1989) and *amiC* (Wilson, 1991; Wilson and Drew, 1991) are required for the wild type amidase induction phenotype although involvement of general host factors conserved between *E. coli* and *Ps. aeruginosa* cannot be ruled out.

b) The low expression of the *amiR* gene *in vivo*, as determined by Coomassie staining of SDS-polyacrylamide gels of cell free extracts, of both *Ps. aeruginosa* and *E. coli* is possibly a result of poor recognition of the translation initiation signal and may be related to the overlap of the *amiC* terminator and *amiR* initiator codons in the wild type configuration.

c) The AmiR protein interacts with the leader RNA at a region upstream and possibly overlapping the Rho-independent terminator sequence. This situation appears to be analogous to the mechanism of gene regulation observed with the highly homologous group of transcription antiterminators of the BglG family, and other regulatory systems that are mediated via similar RNA/protein interactions in both prokaryotes and eukaryotes (Das, 1993).

d) Transcription termination and antitermination events are possibly mutually exclusive since mutations have been generated in the proposed AmiR binding site that reduce antitermination with no apparent effect on the termination reaction. Conversely, base substitution mutations within the terminator have made it somewhat leaky to transcription without any effect on transcriptional antitermination.

Earlier studies of the amidase system of the closely related serotype PAO with mutator phage D3112 (Rehmant and Shapiro, 1983) indicated that *amiB* might be involved in amidase induction from the fact that insertion mutations within the locus led to low amidase expression phenotypes. The in-frame 423 bp *amiB* deletion mutation (plasmid pMW22) (Chapter 3) appears to rule out this possibility in that it gave the wild type inducible phenotype in both *E. coli* and *Ps. aeruginosa.* The effect of the mutator phage insertion at this position is therefore most likely to cause its effect by interrupting expression of the downstream *amiC,R,S* genes as a result of the introduction of phage terminators or other transcriptional polarity elements. A corollary of this interpretation would be that the σ^{54} promoter sequences upstream of *amiC* in the wild type operon are non-functional (Wilson, 1991; Wilson and Drew, 1991). Previous studies have shown that the amidase genes are transcribed on a 5 kb mRNA and that amidase expression is much higher in *Ps. aeruginosa rpoN-* strains raising the possibility of a negative regulatory role for the above elements (Wilson, 1991; Wilson and Drew, 1991).

Previous work using cloned *amiR*, *amiC* and *amiE* genes on separate plasmids in *E. coli* confirmed the negative regulatory role of *amiC* but left open the question of how the inducible phenotype would be generated (Wilson, 1991). Lack of inducibility in this three plasmid complementation system could be due to presence of a truncated AmiR N-terminal fragment since the *amiC* construct also carries the coding sequence for the first 30 AmiR N-terminal residues. This interpretation would of course be valid only if AmiR functioned as a dimer/oligomer in which case the truncated forms of AmiR would exert a dominant negative effect. It has now been shown that coordinate expression of the *amiC/amiR* genes is the important factor for inducible expression which is consistent with the fact that the *amiC* termination codon overlaps *amiR* initiation codon (Chapter 3). This feature is

sometimes observed with coordinately expressed genes in which it is important that expression of the downstream gene occurs only in the presence of the upstream gene product and involves the proportional expression of the genes (Studier and Studier, 1983).

The precise mechanism by which AmiC regulates AmiR activity is not known, however, there is evidence that it is possibly by direct protein/protein interaction. *Ps. aeruginosa* PAC1(Wild-Type) carrying an *amiR* expression plasmid (pDC35) showed a mixed regulatory phenotype (constitutive but further inducible) suggesting saturation of a regulatory component by the excess AmiR (Cousens, 1985). AmiC has been shown to be the inducer (amide) binding component and possibly functions by a mechanism clearly distinct from other well studied two component sensor/regulator regulatory circuits (Wilson *et al.*, 1993). It is distinct from the sensor-kinase/response regulators of the NtrB family in that AmiC does not have kinase activity.

The mechanism of regulation of the amidase operon is also different from the two component regulatory systems of the PEP-PTS sensor (HPK)/transcription antiterminator family described for aromatic β–glucosides catabolism in some bacteria and sucrose utilisation in *B. subtilis* (Amster-Choder and Wright, 1993). AmiR shows no homology to the highly homologous transcription antiterminators of this group, and its recognition sequence, despite having some conserved blocks of nucleotides, has no possibility of adopting a consensus type antiterminator secondary structure (Aymerich and Steinmetz, 1992). A true palindromic sequence has been identified within the proposed AmiR binding region (Fig 7.1) but changes within this sequence do not correlate with changes in the antitermination reaction. Transcription antitermination in the *bgl* homologous systems is linked to transport of the substrate. The sensor protein phosphorylates the substrate and

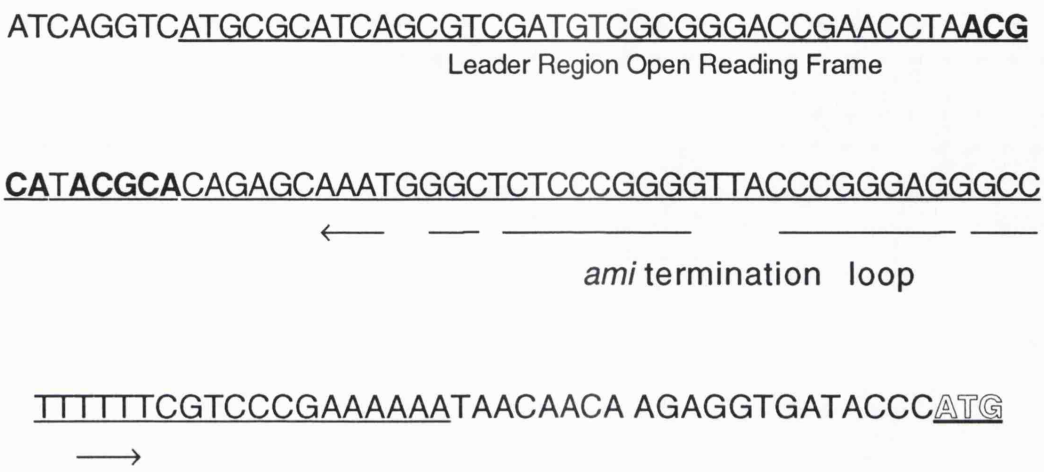


Fig. 7.1 DNA sequence of the amidase leader transcript showing the true palindrome. The leader ORF is underlined, the *amiE* start codon is underlined and shown in outline, the termination loop is shown by the broken arrows and the true palindome is shown in bold.

dephosphorylates (and activates) the antiterminator protein during substrate uptake and inactivates the transcription antiterminator by phosphorylation in the absence of the substrate (Amster-Choder and Wright, 1993). This has been shown in the *sacPA* operon, whereby mutations within the PTS uptake system impair inducibility of the operon (Arnaud *et al.*, 1992). In the amidase operon, induction is apparently independent of active transport of the substrate. In *E. coli*, inducible amidase expression is seen with just *amiE, amiC* and *amiR* and in addition the *amiB* mutation has no effect on the process.

With regard to AmiC repression no catalytic activity has been detected in AmiC and it is proposed that it functions by interacting directly with AmiR possibly sequestering the latter to stop interaction with leader RNA (Wilson *et al.*, 1993). Experiments designed to test this hypothesis were inconclusive. AmiC has no RNA binding activity, as would be expected, but it also does not appear to affect RNA/AmiR complex formation in either the presence of an inducing amide, a repressing amide or by itself (data not presented). Further investigations are clearly required especially since preliminary data shows that the two proteins possibly form a tightly bound complex *in vitro* irrespective of the presence of amides (Wilson, personal communication). These experiments were, however, carried out with impure AmiR extracts and their significance is not yet known. In the absence of a mode of action for AmiC, investigation of possible interactions with the paused elongation complex which will clearly involve a large number of proteins may be a worthwhile exercise when AmiR is finally purified.

It is now apparent that mutations leading to constitutive amidase synthesis could occur in the *amiE* upstream transcription terminator (*amiR* independent), the *amiC* gene (*amiR* dependent) and possibly the *amiR* gene (resistant to *amiC* negative control). Most of the large number of previously isolated constitutive mutants appear to fall into the second class

(R. E. Drew, personal communication).

The results of *in vivo* titrations with sequential deletions of the leader sequence both in *cis*- and *trans*- to *amiR* showed that the AmiR binding site lay 5' to the *Bst*II target at position 206 and encompasses the *Ban*II target at position 198 at least in part (Table 5.2). *In vitro* gel shift assays showed that the AmiR binding site lay between the *Fsp*I target at position 140 and the *Ban*II target at position 198 (Chapter 5). These findings were confirmed by *in vitro* mutagenesis which identified important nucleotides in the transcription antitermination reaction at position 164 to 168 less important ones at position 171 to 180 and critical ones in juxtaposition to the upward face of the hairpin loop (position 183-191) (Chapter 6). However, bandshift experiments with two of the mutant leaders (pMW12 and pMW27), both of which are severely disrupted in the antitermination reaction, show normal gel retardation and presumably normal AmiR binding. This is the first evidence that binding and antitermination are two separate presumably sequential events. Although the mutagenesis studies are incomplete, it is expected that further investigations will define the residues involved in RNA binding and the antitermination reaction. However, the results so far suggest that AmiR functions by direct steric interference of the formation of the terminator hairpin loop in its function as a transcription antiterminator. This would be similar in principle, but not mechanistically, to the BglG family of proteins and would represent convergent evolution of these regulatory systems.

Preliminary structure threading analysis (Jones *et al.*, 1992) of AmiR by using folds in protein data bases has indicated structural but not direct homology to the *E. coli* Endonuclease III, a nucleic acid binding protein, with a hydrophobic N-terminal domain and a hydrophilic C-terminal domain (S. A. Wilson, personal communication). The presence of the hydrophobic domain may be the main cause of the propensity of AmiR to

aggregate as observed in the purification attempts described in Chapter 4.

It is possible to envisage an AmiR mechanism of action analogous to the binding of the Rho protein to *rut* sequences. In these cases initial Rho binding promotes folding of RNA, thus making more contacts with the RNA to form the terminator structure. The analogy stops here however, because Rho is an ATPase unlike AmiR. In the inchworm model of transcription elongation (Chapter 1), a segment of the newly synthesized transcript of up to 9 nucleotides does not leave the product site until the translocation step. Simple terminators then function by the leading face of the terminator stem interacting with the newly synthesized lagging face within the product site in the elongation complex (Fig 1.2). The function of antiterminators in such a scenario would therefore involve direct contact between the antiterminator and the elongation complex at least sterically if not actively. That AmiR indeed interacts with the nascent transcript upstream of the terminator during its function was shown by the production of mutations that interfered with the antitermination step without affecting normal terminator function in the absence of AmiR. Two of these mutant transcripts were shown to bind AmiR in *in vitro* gel shift assays although the affinity was not determined. This would appear to suggest that AmiR binds to the leader and this initial recognition leads to folding of the RNA or other conformational changes that signal antitermination in an at least two step process. The two mutations tested were on either side of the proposed binding site which is at least 33 nucleotides long and thus must be folded in some way to be bound by a sequence specific protein.

RNA recognition by proteins appears to be very varied and the complexity of the situation is not helped by the large number of possible intramolecular RNA interactions to form higher order structures (pseudoknots) (Mattaj, 1993). Protein recognition of RNA secondary structure often involves RNA loops and bulges since these structures expose the

RNA backbone and bases to interaction with the protein groups (Mattaj, 1993). The BglG/SacY family of proteins appear to recognize bulges in the secondary structure stem with the bases in double helical stems playing a minor role (Aymerich and Steinmetz, 1992). The HIV proteins Tat, an antiterminator, and Rev, also recognize bulges and loops in the RNA during their reactions although the precise interactions have not been fully characterized (Mattaj, 1993). By eye it is not possible to fold the AmiR recognition site into structures similar to the BglG group RAT elements (Aymerich and Steinmetz, 1992). Sequence analysis of AmiR is likely to be important in trying to predict the RNA binding motif, however, such a study is unlikely to shed any light on the structure of the RNA sequence it recognizes. The current families of RNA binding proteins do not appear to recognize predictable secondary structures and members of the same family recognize a diverse range of secondary structures (Mattaj, 1993).

Previous studies and the work presented in this thesis can be summarised in a model (Fig 7.2). Under noninducing conditions, the vast majority of transcripts initiating at the *amiE* σ^{70} promoter terminate at the leader sequence transcription terminator. However, a basal level readthrough of the terminator allows expression of low amounts of all downstream genes with AmiE being expressed in relatively higher amounts because of the downstream terminator. Under these conditions AmiC inactivates AmiR most probably by the formation of a protein/protein complex. Under inducing conditions, the binding of amides to AmiC either causes a conformational change within the protein/protein complex or causes the complex to dissociate allowing AmiR to function as a transcription antiterminator.

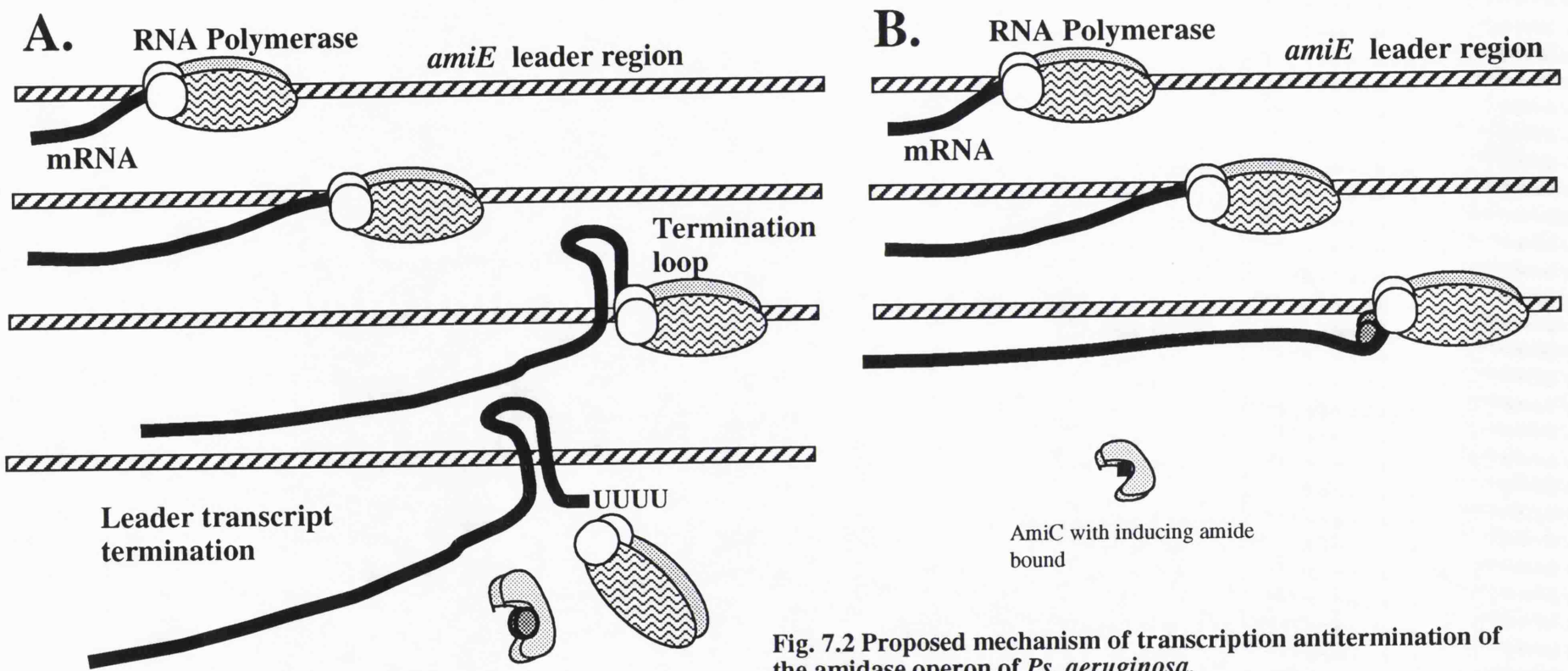


Fig. 7.2 Proposed mechanism of transcription antitermination of the amidase operon of *Ps. aeruginosa*.
A. In the absence of inducing amides AmiC sequesters AmiR and the termination loop forms. **B.** In the presence of inducing amides, which bind to AmiC, AmiR is free to interact with the recognition sequence and prevent formation of the termination loop.

Notes

Parts of this work have been published and presented at meetings

1. Wilson, S. A., Wachira, S. J., Drew, R. E., Jones, D., and Pearl, L. H. 1993. Antitermination of amidase expression in *Pseudomonas aeruginosa* is controlled by a novel cytoplasmic amide binding protein. EMBO J. 12: 3637-3642

2. Wilson, S.A., Wacheria, S.J., Drew, R.E., Jones, D., and Pearl, L.H. (submitted to Microbiology). Identification of a putative ABC transporter in the *Pseudomonas aeruginosa* amidase operon.

3. The *Ps. aeruginosa* transcription antiterminator, AmiR, binds to the leader mRNA in the amidase operon. Wachira, S. J., Drew, R. E., Pearl, L. H., and Wilson, S. A. Fourth International symposium on *Pseudomonas*: Biotechnology and Molecular Biology. Vancouver Canada. 1993.

4. Thanks to Sharon Banin who carried out the original *in vitro* mutagenesis studies as part of her third year undergraduate project.

Arnaud, M., Vary, P., Zagorek, M., Klier, A., Debarbouille, M., Postma, P., Rapaport. 1992. Regulation of *sacPA* operon of *Bacillus subtilis* : identification of phosphotransferase system components involved in SacT activity. J. Bacteriol. 174: 3161-3170

References

Adhya, S., and Miller, W. 1979. Modulation of two promoters of the galactose operon of *Escherichia coli.* Nature 279: 492-494

Aiba, H., Nakamura, T., Mitani, T., and Mori, H. 1985: Mutations that alter the allosteric nature of cAMP receptor protein of *Escherichia coli.* EMBO J. 4: 3329-3332

Ambler, R. P., Auffret, A. D., and Clarke, P. H. 1987. The amino acid sequence of the aliphatic amidase from *Pseudomonas aeruginosa.* FEBS Lett. 215: 285-290

Ambler, R. P., and Wynn, M. 1973. The amino acid sequences of cytochrome c-551 from three species of *Pseudomonas.* Biochem. J. 131: 485-498

Amster-Choder, O., and Wright, A. 1993. Transcriptional regulation of the bgl operon of *Escherichia coli* involves phosphotransferase system-mediated phosphorylation of a transcriptional antiterminator. J. Cellular Biochemistry 51: 83-90

Arndt, K., and Chamberlin, M. J. 1988. Transcription termination in *Escherichia coli.* Measurement of rate of release from Rho-independent terminators. J. Mol. Biol. 202: 271-285

Arndt, K., and Chamberlin, M. J. 1990. RNA chain elongation by *Escherichia coli* RNA polymerase. Factors affecting the stability of elongating ternary complexes. J. Mol. Biol. 211: 79-108

Asano, Y., Tachibana, M., Yani, Y., and Yamada, H. 1982. Purification and characterization of amidase which participates in nitrile degradation. Agric. Biol. Chem. 46: 1175-1181

Aymerich, S., and Steinmetz, M. 1992. Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. Proc. Natl. Acad. Sci. USA 89: 10410-10414

Aymerich, S., and Steinmetz, M. 1987. Cloning and preliminary characterization of the *sacS* locus from *Bacillus subtilis* which controls the regulation of the exoenzyme levansucrase. Mol. Gen. Genet. 208: 114-120

Babitzke, P., Gollnick, P., and Yanofsky, C. 1992.The *mtrAB* operon of *Bacillus subtilis*

Bear, D.G., Hicks, P.S., Escudero, K. W., Andrews, C. L., McSwiggen, J.A., and Von Hippel, P.H. 1988. J. Mol. Biol. 199: 623-635

encodes GTP cyclohydrolase I (MtrA), an enzyme involved in folic acid biosynthesis, and MtrB a regulator of tryptophan biosynthesis. J. Bacteriol. 174: 2059-2064

Babitzke, P., and Yanofsky, C. 1993. Reconstitution of *Bacillus subtilis trp* attenuation *in vitro* with TRAP, the *trp* RNA binding attenuation protein. Proc. Natl. Acad. Sci. USA 90: 133-137

Bagdasarian, M., Franklin, F. C. H., Lurz, R., Ruckert, B., Bagdasarian, M. M., and Timmis, K. N. 1981. Specific purpose cloning vectors. II Broad host range, high copy number RSF1010-derived vectors and a host-vector system for cloning in *Pseudomonas.* Gene 16: 237-247

Barik, S., Ghosh, B., Whalen, W., Lazinski, D., and Das, A. 1987. An antitermination protein engages the elongating transcription apparatus at a promoter-proximal recognition site. Cell. 50: 885-889

Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. 1991. Role of mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II. Mol. Cell. Biol. 11: 1195-1206

Bissonnette, L., and Roy, P. H. 1992. Characterization of InO of *Pseudomonas aeruginosa* plasmid pVS1, an ancestor of integrons of multiresistance plasmids and transposons of gram-negative bacteria. J. Bacteriol. 174: 1248-1257

Boyer, H. W., and Roullard-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli.* J. Mol. Biol. 41: 459-472

Bramley, H. F., and Kornberg, H. L. 1987. Sequence homologies between proteins of bacterial phosphoenolpyruvate-dependent sugar phosphotransferase systems: identification of possible phosphate-carrying histidine residues. Proc. Natl. Acad. Sci. USA 84: 4777-4780

Bramley, H. F., and Kornberg, H. L. 1987. Nucleotide sequence of *bglC*, the gene specifying enzyme^{IIbgl} of the PEP: sugar phosphotransferase system in *Escherichia coli* K12, and overexpression of the gene product . J. Gen. Microbiol. 133: 563-573

Brammar, W. J., and Clarke, P. H. 1964. Induction and repression of *Pseudomonas aeruginosa* amidase. J. Gen. Microbiol. 37: 307-319

Brendel, V., and Trifonov, E. N. 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res.12: 4411-4427

Briat, J. -F., and Chamberlin, M. J. 1984. Identification and characterization of a new transcriptional termination factor from *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 81: 7373-7377

Burnette, W. N. 1981. "Western blotting": electrophoretic transfer of proteins fron sodium dodecyl sulphate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated proteinA. Anal. Biochem. 112: 195-304

Brammar, W. J., Clarke, P. H., and Skinner, A. J. 1967. Biochemical and genetic studies with regulator mutants of *Pseudomonas aeruginosa* amidase. J. Gen. Microbiol. 47: 87-102

Brammar, W. J., Charles, I. G., Matfield, M., Cheng-Pin, L., Drew, R. E., and Clarke, P. H. 1987. The nucleotide sequence of the *amiE* gene of *Pseudomonas aeruginosa.* FEBS Lett. 215: 291-294

Brown, P. R. 1969. Studies on wild type and mutant amidase proteins from *Pseudomonas aeruginosa.* PhD thesis, University of London

Brown, P. R., Smyth, M. J., Clarke, P. H., and Rosemeyer M. A. 1973. The subunit structure of the aliphatic amidase from *Pseudomonas aeruginosa.* Eur. J. Biochem. 34: 177-187

Brunner, M., and Bujard, H. 1987. Promoter recognition and promoter strength in the *Escherichia coli* system. EMBO J 6: 3139-3144

Buckel, P., and Zehelein, E. 1981. Expression of *Pseudomonas fluorescens* D-galactose dehydrogenase in *E. coli.* Gene 16: 149-159

Burgess, R. R. 1976.Purification and physical properties of *E. coli* RNA polymerase. *In* RNA polymerase (Losick, R., and Chamberlin, M. ed.) pp 69-106 Cold Spring Harbor Laboratory. Cold spring Harbor, N.Y.

Burokhov, S., Sagitov, V., and Goldfarb, A. 1993. Transcript cleavage factors from *E. coli.* Cell 72: 459-466

Byng, G. S., Whitaker, R. J., Gherna, R. L., and Jensen, R. A. 1980. Variable enzymological patterning in tyrosine biosynthesis as a means of determining natural relatedness among the pseudomonadaceae. J. Bacteriol. 144: 247-251

Cane, P., and Williams, P. A. 1986. A restriction map of naphthalene catabolic plasmid pWW60-1 and location of some of its catabolic genes. J. Gen. Microbiol. 132: 2919-2929.

Carpousis, A. J., and Gralla, J. D. 1980.Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation *in vitro* at the *lacUV5* promoter. Biochemistry 19: 3245-3253

Chamberlin, M. J. 1982. Bacterial DNA-dependent RNA polymerases. *In* The enzymes. vol XV pp 61-86. Academic Press, Inc.

Chatterjee, A. K., and Vivader, A. K. (ed.) Advances in plant pathology, vol 4, pp 1-218. Academic Press, Inc, N.Y.

Chitinis, C. E., and Ohman, D. E. 1993. Genetic analysis of the alginate biosynthetic gene cluster of *Pseudomonas aeruginosa* shows evidence of an operonic structure. Mol. Microbiol. 8: 583-590

Clarke, L., and Carbon, J. 1978. Functional expression of cloned yeast DNA in *Escherichia coli*: specific complementation of argininosuccinate lyase (*argH*) mutations. J. Mol. Biol. 120: 517-532

Clarke, P. 1972. Biochemical and immunological comparision of aliphatic amidases produced by *Pseudomonas* species. J. Gen. Microbiol. 71: 241-257

Clarke, P. H. 1984. Amidases of *Pseudomonas aeruginosa. In* Microorganisms as model systems for studying evolution (Mortlock, R. P. ed.) pp 187-231. Plenum Press, N.Y.

Clarke, P. H., and Drew, R. 1988. An experiment in enzyme evolution. Studies with *Pseudomonas aeruginosa* amidase. Bioscience Reports. 8: 103-120

Clarke, P. H., Drew, R. E., Turbeville, C., Brammar, W. J., Ambler, R. P., and Auffret, A. D. 1981. Alignment of cloned *amiE* gene of *Pseudomonas aeruginosa* with the N-terminal sequence of amidase. Biosci. Rep. 1: 299-307

Clarke, P. H., and Laverack, P. D. 1983. Expression of the *argF* gene of *Pseudomonas aeruginosa* in *Pseudomonas aeruginosa, Pseudomonas putida, and Escherichia coli.* J. Bacteriol. 154: 508-512

Clarke, P. H., and Ornston, N. 1975. Metabolic pathways and regulation: II. *In* Genetics and biochemistry of *Pseudomonas* (Clarke, P. H., and Richmond, M. H. eds). pp263-340. John Wiley and Sons

Clarke, P. H., and Slater, J. H. 1986. Evolution of enzyme structure and function in *Pseudomonas. In* The bacteria A treatise on structure and function (Sokatch, J. R. ed.), pp 71-144. Academic Press

Cousens, D. J. 1985. Studies on the amidase regulator gene of *Pseudomonas aeruginosa.* PhD thesis University of London

Dambroski, A. J. and Platt, T. 1988. Structure of ρ factor: An RNA-binding domain and a separate region with strong similarity to protein ATP-binding domains. Proc. Natl. Acad. Sci. USA. 82: 2538-2542

Darlix, J. L., Sentenac, A., and Fromageot, P. 1971. Binding of termination factor Rho to RNA polymerase and DNA. FEBS Lett. 13: 165-168

Cousens, D. J., Clarke, P. H., and Drew, R. E. 1987. The amidase regulatory gene (*amiR*) of *Pseudomonas aeruginosa.* J. Gen. Microbiol. 133: 2041-2052

Crothers, D. M., and Steitz, T. A. 1992. Transcriptional activation by *Escherichia coli* CAP protein. *In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.) pp 501-534. Cold Spring Harbor Laboratory Press

Crutz, A. M., Steinmetz, M., Aymerich, S., Richter, R., and Le Coq, D. 1990. Induction of levansucrase in *Bacillus subtilis*: an antitermination mechanism negatively controlled by the phosphotransferase system. J. Bacteriol. 172: 1043-1050

Currier, T. C., and Morgan, M. K. 1982. Direct DNA repeat in plasmid R68.45 is associated with deletion formation and concomitant loss of chromosome mobilization ability. J. Bacteriol. 150: 251-259

Dall, D. J., Anzola, J. V. Xu, Z., and Nuss, D. L. 1990. Structure-specific binding of wound tumor virus transcripts by a host factor: Involvement of both terminal nucleotide domains. Virology 179: 599-608

Darst, S. A., Kubalet, E. W., and Kornberg, R. D. 1989. Three dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. Nature 340: 730-732

Das, A. 1993. Control of transcription termination by RNA-binding proteins. Annu. Rev. Biochem. 62: 893-930

Das, A. 1992. How the phage lambda N gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in the nascent RNA. J. Bacteriol. 174: 6711-6716

De'barbouille, M., Arnaud, M., and Fouet, A., Klier, A., and Rapoport, G. 1990. The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. J. Bacteriol. 172: 3966-3973

de Lorenzo, V., Herrero, M., Metzke, M., and Timmis, K. N. 1991. An upstream XylR- and IHF- induced nucleoprotein complex regulates the σ^{54}-dependent Pu promoter of TOL plasmid. EMBO J 10: 1159-1167

Dean, H. F., and Morgan, H. F. 1983. Integration of R 91-5::Tn501 into the

Dotto, G. P., Horiuchi, K., and Zinder, N. D. 1984. The functional origin of bacteriophage f1 DNA replication. Its signals and domains. J. Mol. Biol. 172: 507

Pseudomonas putida PPN chromosome and genetic circularity of the chromosome map. J. Bacteriol. 153: 485-487

Dente, L., Cesareni, G., Cortese, R. 1983. pEMBL: a new family of single stranded plasmids. Nucleic Acids Res. 11: 1645-1655

Deretic, V., Konyecsni, W. M. Mohr, C. D., Maritin, D. W., and Hibbler, N. S. 1989. Common denominators of promoter control in *Pseudomonas* and other bacteria. Bio/technology 7: 1249-1254

Deretic, V., Tomasek, P., Darzins, A., and Chakrabarty, A. M. 1986. Gene amplification induces mucoid phenotype in *rec-2 Pseudomonas aeruginosa* exposed to kanamycin. J. Bacteriol. 165: 387-395

Dilauro, R., Taniguchi, T., Musso, R., and deCrombrugghe, B. 1979.Unusual location and function of the operator in the *Escherichia coli* galactose operon. Nature 279: 494-500

Dixon, R. C. 1986.The *xylABC* promoter from *Pseudomonas putida* TOL plasmid is activated by nitrogen regulatory genes in *Escherichia coli.* Mol. Gen. Genet. 203: 129-136

van der Drift, C., and Janssen, D. B. 1985. Regulation of enzymes under nitrogen control in *Pseudomonas aeruginosa.* Current Topics in Cellular regulation 26: 485-490

Drew, R. E. 1984. Complementation analysis of the aliphatic amidase genes of *Pseudomonas.* J. Gen. Microbiol. 130: 3101-3111

Drew, R. E., Clarke, P. H., and Brammar, W. J. 1980. The construction *in vitro* of derivatives of bacteriophage lambda carrying the amidase genes of *Pseudomonas aeruginosa.* Mol. Gen. Genet. 177: 311-320

Drew, R. E., and Lowe, N. 1989. Positive control of *Pseudomonas aeruginosa* amidase synthesis is mediated by a transcription antitermination mechanism. J. Gen. Microbiol. 135: 817-823

Drew, R., and Wilson, S. 1992. Regulation of amidase expression in *Pseudomonas aeruginosa.In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.), pp 207-213. American Society For Microbiology, Washington, D.C.

Dunn, J. J., and Studier, F. W. 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. 166: 477-535

Friedberg, E. C. 1985. *DNA repair* New York: W. H. Freeman. 614pp

El Hassouni, M., Chipaux, M., and Barras, F. 1990. Analysis of the *Erwinia chrysanthemi arb* genes, which mediate metabolism of aromatic β-glucosides. J. Bacteriol. 172: 6261-6267

El Hassouni, M., Henrissat, B., Chippaux, M., and Barras, F. 1992. Nucleotide sequences of the arb genes, which control ß-glucoside utilization in *Erwinia chrysanthemi*: comparision with the *Escherichia coli bgl* operon and evidence for a new β-glycohydrolase family including enzymes from eubacteria, archeabacteria, and humans. J. Bacteriol. 174: 765-777

Farin, F. 1976. Studies on the mechanism of regulation of amidase synthesis in *Pseudomonas aeruginosa.* PhD thesis, University of London

Farin, F., and Clarke, P. H. 1978. Positive regulation of amidase synthesis in *Pseudomonas aeruginosa.* J. Bacteriol. 135: 379-392

Farnham, P.J., and Platt, T. 1982. Effects of DNA base analogs on transcription termination at the tryptophan operon attenuator of Escherichia coli. Proc. Natl. Acad. Sci. USA 79: 998-1002

Franklin, F. C. H., Bagdasarian, M. Bagdasarian, M. M., Timmis, K. N. 1981. Molecular and functional analysis of TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring *meta* cleavage pathway. Proc. Natl. Acad. Sci. USA. 78: 7458-7462

Fried, M. G., and Crothers, D. M. 1984. Kinetics and mechanisms in the reaction of gene regulatory proteins with DNA. J. Mol. Biol. 172: 263-282

Friedman, D. I. 1988. Integration host factor: A protein for all reasons. Cell 55: 545-554

Friedman, D. I., Imperiale, M. J., and Adhya, S. L. 1987. RNA 3' end formation in control of gene expression. Ann. Rev. Genet. 21: 453-488

Friedman, D. I., Olson, E. R., Johnson, L. L., Alessi, D., and Craven, M. G. 1990. Transcription dependent competition for a host factor: the function and optimal sequence of the phage lambda boxA transcription antitermination signal. Genes Dev. 4: 2210-2222

Gamper, H. B., and Hearst, J. E. 1982. A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation, and ternary complexes. Cell 29: 81-90

Geider, K., Hohmeyer, C., Haas, R., Meyer, T. F. 1985. A plasmid cloning system utilizing the replication and packaging functions of the filamentous bacteriophage fd. Gene : 341-344

Goldberg, J. B. 1992. Regulation of alginate volume in *Pseudomonas aeruginosa.* transcription
complex possibly via the Nus proteins (Amster-Choder and Wright, 1993).

Goliger, J. A., Yang, X., Guo, H-C., and Roberts, J. W. 1989. Early transcribed sequences affect termination efficiency of *Escherichia coli* RNA polymerase. J. Mol. Biol. 205: 331-341

Gomada, M., Inouye, S., Imaishi, H., Nakazawa, A., and Nakazawa, T. 1992. Analysis of an upstream regulatory sequence required for activation of the regulatory gene *xylS* in xylene metabolism directed by the TOL plasmid of *Pseudomonas putida.* Mol. Gen. Genet. 233: 419-426

Gordon, A. J. E., Burns, P. A., Fix, D. F., Yatagai, F., Allen, F. L., Horsfall, M. J., Halliday, J. A., Gray, J., Bernelot-Moens, C., and Glickman, B. W. 1988. Missense mutation in the *lacI* gene of *Escherichia coli.* Inferences on the structure of the repressor protein. J. Mol. Biol. 200: 239-251

Gottesman, S., Squires, C., Pichersky, E., Carrington, M., Hobbs, M., Mattick, J. S., Dalrymple, B., Kuramitsu, H., Shiroza, T., Forster, T., Clark, W. P., Ross, B., Squires, C. L., and Maurizi, M. R. 1990. Conservation of the regulatory subunit for the *clp* ATP-dependent protease in prokaryotes and eukaryotes. Proc. Natl. Acad. Sci. USA 87: 3513-3517

Gralla, J. D. 1992. *lac* repressor. *In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.) pp 629-642. Cold Spring Harbor Laboratory Press

Gralla, J. D. 1990. Promoter recognition and mRNA initiation by *E. coli* Eσ70.*In* Methods Enzymol (Goeddel, D. V. ed) . 185: 37-54

Grantham, R., Gautier, C., Gouy, M., Mercier, R., and Pave, A. 1980.Codon catalog usage and the genome hypothesis. Nucl. Acids Res. 8: r49-r62

Grantham, R., Gautier, C., and Gouy, M. 1980.Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type. Nucl.

Acids Res. 8: 1893-1912

Greenblatt, J. 1992. Protein-protein interactions as critical determinants of regulated initiation and termination of transcription. *In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.), pp203-226. Cold Spring Harbor Laboratory Press

Gribskov, M., Devereux, J., and Burgess, R. R. 1984. The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucl. acids Res. 12: 539-549

Groisman, E. A., and Casabadan, M. J. 1987. Cloning of genes from members of the family *Enterobacteriaceae* with mini-Mu bacteriphage containing plasmid replicons. J. Bacteriol. 169: 687-693

Gross, C. A., Lonetto, M., and Losick, R. 1992. Bacterial sigma factors. *In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.), pp129-176. Cold Spring Harbor Laboratory Press

Haas, D. 1983. Genetic aspects of biodegradation by pseudomonads. Experientia. 39: 1199-1213

Haber, R., and Adhya, S. 1988. Interaction of spartially separated protein-DNA complexes for control of gene expression: operator conversion. Proc. Natl. Acad. Sci. USA 85: 9683-9687

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-580

Harayama, S., Lehrbach, P. R., Timmis, K. N. 1984. Transposon mutagenesis analysis of meta-cleavage pathway operon genes of the TOL plasmid of *Pseudomonas putida mt-2*. J. Bacteriol. 160: 251-255

Harayama, S., Leppick, R. A., Rekik, M., Mermond, N., Lehrbach, P. R., Reineke, W., and Timmis, K. N. 1986. Gene order of the TOL catabolic plasmid upper pathway operon and oxidation of both toluene and benzylalcohol by the *xylA* product. J. Bacteriol. 167: 455-460

Hayes, W. 1968. The Genetics of Bacteria and their Viruses: studies in basic genetics and molecular biology. Blackwell scientific publications

Howard, B. H. and De Crombrugghe, B. 1976. ATPase activity required for termination of transcription by the *Escherichia coli* protein factor ρ J. Biol. Chem. 251: 2520-2524

Henikoff, S., Haughn, G. W., Calvo, J. M., and Wallace, J. C. 1988. A large number of bacteria activator proteins. Proc. Natl. Acad. Sci. USA 85: 6602-6606

Henrissat, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280: 309-316

Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W., and Hermodson, M. A. 1986. A family of related ATP-binding subunits coupled to many distinct biological process in bacteria. Nature 323: 448-450

von Hippel, P. A., Bear, D. G., Morgan, W., and Mcswiggen, J. 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. 53: 389-446

Holloway, B. W., and Krishnapillai, V. 1975. Bacteriophages and bacteriocins. *In* Genetics and Biochemistry of Pseudomonas (Clarke, P. H., and Richmond, M. A. ed.), pp 99-135. Wiley. London

Holloway, B. W., and Morgan, A. F. 1986. Genome organization in *Pseudomonas*. Ann. Rev. Microbiol. 40: 79-105

Horwitz, R. J., Li, J., and Greenblat, J. 1987. An elongation control particle containing the N gene transcriptional antitermination protein of bacteriophage lambda. Cell 51: 631-641

Houman, F., Diaz-Torres, M. R., and Wright, A. 1990. Transcriptional antitermination in the *bgl* operon of E. coli is modulated by a specific RNA binding protein. Cell 62: 1153-1163

Hutchison, C. A. III, Phillips, S., Edgell, M. A., Gillam, S., Jahuke, P., Smith, M. 1978. Mutagenesis at a specific position in a DNA sequence. J. Biol. Chem. 253: 6551-6560

Inouye, S., Gomada, M., Sangodkar, U. M. X., and Nakazawa, A. 1990. Upstream regulatory sequence for transcriptional activator XylR in the first operon of xylene metabolism on the TOL plasmid. J. Mol. Biol. 216: 251-260

Izban, M. G., and Luse, D. S. 1992. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'-5' direction in the presence of elongation factor S II. Genes Dev. 6: 1342-1356

Jeng, S. T., Gardner, J. F., and Gumport, R. I. 1990. Transcription termination by bacteriophage T7 RNA polymerase at Rho-independent terminators. J. Biol. Chem. 265: 3823-3830

Kassavetis, G. A., and Chamberlin, M. J. 1981. Pausing and termination of transcription within the early region of bacteriophage T7 DNA *in vitro.* J. Biol. Chem.256: 2777-2786

Janssen, D. B., op den Camp, H. J. M., Leenan, P. J. M., and van der Drift, C. 1980. The enymes of ammonia assimilation in *Pseudomonas aeruginosa.* Arch. Microbiol. 124: 197-203

Johnston, H. M. and Roth, J. R. 1981. DNA sequence changes of mutations altering attenuation control of histidine operon of *Salmonella typhimurium.* J. Mol. Biol. 145: 713-734

Jones, D.T., Taylor, W.R., Thornton, J.M. 1992. A new approach to protein fold recognition. Nature. 358: 86-89

Kammerer, W., Deuschle, U., Gentz, R., and Bujard, H. 1986. Functional dissection of *Escherichia coli* promoters: information in the initial transcribed region is involved in late steps of the overall process. EMBO J 5: 2995-3000

Kelly, H. and Clarke, P. H. 1962. An inducible amidase produced by a strain of *Ps. aeruginosa.* J. Gen. Microbiol. 27: 305-316

Kerpolla, T. K., and Kane, C. M. 1990. Analysis for the signals for transcription termination by purified RNA polymerase II. Biochemistry 29: 269-278

Kleina, L. G., and Miller, J. H. 1990. Genetic studies of lac repressor XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. J. Mol. Biol. 212: 295-318

Kornaska, M. M., and Sharp, P. A. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell 46: 845-855

Kramer, B., Kramer, W., and Fritz, H.J. 1984. Different base/base mismatches are corrected with different efficiencies by methyl-directed DNA mismatch-repair system of *E. coli.* Cell 38: 879-887

Krawiec, S., and Riley, M. 1990. Organization of the bacterial chromosome. Microbiol. Rev. 54: 502-539

Krummel, B., and Chamberlin, M. J. 1989. RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary complexes. Biochemistry 28: 7829-7842

Krummel, B., and Chamberlin, M. J. 1989. RNA initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary comlexes. Biochemistry 28: 7829-7842

Krummel, B., and Chamberlin, M. J. 1992 a. Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Deoxyribonuclease I footprint of defined complexes. J. Mol. Biol. 225: 239-250

Krummel, B., and Chamberlin, M. J. 1992 b. Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Individual complexes halted along different transcription units have distinct and unexpected biochemical properties. J. Mol. Biol. 225: 221-237

Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA 82: 488-492

Kuroda, M. I., Shimotsu, H., Henner, D. J. and Yanofsky, C. 1986. Regulatory elements common to the *Bacillus pumilus* and *Bacillus subtilis trp* operons. J. Bacteriol. 167: 792-798

Kustu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. 1989. Expression of σ^{54}-dependent genes is probably united by a common mechanism. Microbiol. Rev. 53: 367-376

Laemli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

Landick, R., Carey, J., and Yanofsky, C. 1985. Regulation of aspartate transcarbamoylase synthese in *Escherichia coli*: Analysis of deletion mutations in the promoter region of the *pyrBI* operon. Proc. Natl. Acad. Sci. USA 82: 4663-4667

Lane, D. Prentki, P., and Chandler, M. 1992. Use of gel retardation to analyze protein-nucleic acid interactions. Microbiol. Rev. 56: 509-528

Lee, D. N. Phung, L. Stewart, J., and Landick, R. 1990. Transcription pausing by *Escherichia coli* RNA polymerase modulated by downstream DNA sequences. J. Biol. Chem.265: 15145-15153

Lowery-Goldhammer, C. and Richardson, J. P. 1974. An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with the Rho termination factor. Proc. Natl. Acad. Sci. USA. 71: 2003-2007

Lehming, N., Sartorius, J., Kisters-Woike, B., von Wilcken-Bergmann, B., and Muller-Hill, B. 1990. Mutant lac repressors with new specificities hint at rules for protein-DNA recognition. EMBO J 9: 615-621

Leong, J., Bitter, W., Koster, M., Venturi, V., and Weisbeek, P.J. 1992. Molecular analysis of iron assimilation in plant growth-promoting *Pseudomonas putida* WCS358. *In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.), pp 30-36. American Society For Microbiology, Washington, D.C.

Liu-Johnson, H. -N., Gartenberg, M. R., and Crothers, D. M. 1986. The DNA binding domain and bending angle of the *E. coli* CAP protein. Cell 47: 995-1005

Lobell, R. B., and Schleif, R. F. 1990. DNA looping and unlooping by AraC protein. Science 50: 528-532

Loper, J. E., Orser, C. S. Panopoulos, N. J., and Schroth, M. N. 1984. Genetic analysis of fluorescent pigment production in *Pseudomonas syringae pv. syringae*. J. Gen. Microbiol. 130: 1507-1515

Lowe, N., Rice, P. M., and Drew, R. E. 1989. Nucleotide sequence of the aliphatic amidase regulator gene (*amiR*) of *Pseudomonas aeruginosa.* FEBS Lett. 246: 39-43

Lukat, G. S., and Stock, J. B. 1993. Response regulation in bacterial chemotaxis. J. Cellular Biochemistry 51: 41-46

Lukat, G. S., MMcCleary, W. R., Stock, A. M., and Stock, J. B. 1992. Phosphorylation of bacterial response regulators by low molecular weight phospho-donors. Proc. Natl. Acad. Sci. USA 89: 718-722

Lynn, S. P. Bauer, C. E., Chapman, K., and Gardner, J. F. 1985. Identification and characterization of mutants affecting transcription termination at threonine operon attenuator. J. Mol. Biol. 183: 529-541

MacGregor, C. H., Wolff, J. A., Arora, S. K., Hylemon, P. B., and Phibbs, P. V., Jr. 1992. Catabolite repression control in *Pseudomonas aeruginosa. In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.), pp 198-206. American Society For Microbiology, Washington, D.C.

Magasanik, B. 1993. The regulation of nitrogen utilization in enteric bacteria. J. Cellular

Masasume, Y., and Richardson, C. C. 1971. Strand displacement during deoxyribonucleic acid synthesis at single strand breaks. J. Biol. Chem. 246: 2692-2698

Mayaux, J. -F., Cerbelaud, E., Soubrier, F., Faucher, D., and Petre, D. 1990. Purification, cloning, and primary structure of an enantiomer selective amidase from *Brevibacterium* sp. strain R312: structural evidence for genetic coupling with nitrile hydratase. J. Bacteriol. 172: 6764-6773

Biochemistry 51: 34-40

Magasanik, B., and Neidhardt, F. C. 1987. Regulation of carbon and nitrogen utilization.In '*Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology (Neidhardt, F. C. ed.) pp 1318-1325. American Society for Microbiology, Washington, D. C.

Mahadevan, S., Reynolds, A. E., and Wright, A. 1987. Positive and negative regulation of the *bgl* operon of *Escherichia coli.* J. Bacteriol. 169: 2570-2578

Mahadevan, S., and Wright, A. 1987. A bacterial gene involved in transcription antitermination: regulation at a Rho-independent terminator in the *bgl* operon of *E. coli.* Cell 50: 485-494

Marrs, C. F., Schoolnik, G., Koomey, J. M., Hardy, J., Rothbard, J., and Falkow, S. 1985. Cloning and sequencing of a *Moraxella bovis* pilin gene. J. Bacteriol. 163: 132-139

Martinez, E., and de la Cruz, F. 1990. Genetic elements involved in Tn21 site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO. J. 9: 1275-1281

Mason, S. W., and Greenblatt, J. 1991. Assembly of transcription elongation complexes containing the N protein of phage λ and *Escherichia coli* elongation factors NusA, NusB, NusG, and S10. Genes Dev. 5: 1504-1512

Mattaj, I. W. RNA recognition: A family matter? Cell 73: 837-840

Mayaux, J. -F., Cerbeland, E., Soubrier, F., Faucher, D., Teh, P., Blauche, F., and Petre, D. 1991. Purification, cloning, and primary structure of a new enantiomer-selective amidase from a *rhodococcus sp* strain: structural evidence for a conserved genetic coupling with nitrile hydratase. J. Bacteriol. 173: 6694-6704

McClure, W. 1980. Rate-limiting steps in RNA chain initiation. Proc. Natl. Acad. Sci. USA 77: 5634-5638

McClure, W. R. 1985. Mechanism and control of transcription initiation in prokaryotes. Ann. Rev. Biochem. 54: 171-204

Mercier, J., Lachapelle, J., Couture, F., Lafond, M., Vezina, G., Boissinot, M., and Levesque, R. C. 1990. Structural characterization of *tnpI*, a recombinase locus in Tn21 and

Mizusawa, S., and Ward, D. F. 1982. Bacteriophage Lambda vector for cloning with *Bam*HI and *Sau*3A. Gene 20:317-322

related ß-lactamase transposons. J. Bacteriol. 172: 3745-3757

Messing, J. 1983. New M13 vectors for cloning. *In* Methods Enzymol. 101: 20-78

Miller, J. H. 1978. The *lacI* gene: its role in *lac* operon control and its use as a genetic system. *In* The operon (Miller, J. H., and Reznikoff, W. S. ed.), pp 31-88. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

Miller, J. H., Coulondre, C., Hofer, M., Schmeissner, U., Sommer, H., Schimtz, A.,and Lu, P. 1979. Genetic studies of *lac* repressor. IX. Generation of altered proteins by suppression of nonsense mutations. J. Mol. Biol. 131: 191-222

Minton, N. P., Atkinson, T., and Sherwood, R. F. 1983. Molecular cloning of the *Pseudomonas* carboxypeptidase G2 gene and its expression in *Escherichia coli* and *Pseudomonas putida.* J. Bacteriol. 156: 1222-1227

Morales, V., Bagdasarian, M. M., and Bagdasarian, M. 1990. Promiscuous plasmids of the IncQ group: mode of replication and use for gene cloning in gram-negative bacteria. In *Pseudomonas*: Biotransformations, Pathogenesis, and evolving Biotechnology (Silver, S., Chakrabarty, A. M., Iglewski, B., and Kaplan, S. ed.), pp 229-241. American Society for Microbiology

Morgan, E. A. 1986. Antitermination mechanisms in rRNA operons in *E. coli.* J. Bacteriol. 108: 1-5

Munson, L. M., and Reznikoff, W. S. 1981. Abortive initiation and long ribonucleic acid synthesis. Biochemistry 20: 2081-2085

Mustaev, A., Kashlev, M., Zaychikov, E., Grachev, M., and Goldfarb, A. 1993. Active centre rearrangement in RNA polymerase initiation complex. J. Biol. Chem. 268: 19185-19187

Mylroie, J. R., Friello, D. A., Siemens, T. V., and Chakrabarty, A. M. 1977. Mapping of *Pseudomonas* putida chromosomal genes with recombinant sex-factor plasmid. Mol. Gen. Genet. 157: 231-237

Nakamura, Y., and Uchida, H. 1983. Isolation of conditionally lethal mutations affecting synthesis of *nusA* protein of *Escherichia coli.* Mol. Gen. Genet. 190: 196-203

Neff, N., and Chamberlin.W. J. 1980. Termination of transcription by *Escherichia coli* ribonucleic acid polymerase *in vitro.* Effect of altered reaction conditions and mutations in the enzyme protein on termination with T7 and T3 deoxyribonucleic acids. Biochemistry 19: 3005-3015.

Oda, T. and Takanami, M. 1973. Observations on the structure of the termination factor Rho and its attachment to DNA. J. Mol. Biol. 71: 799-802

Nakazawa, T., and Inouye, S. 1986. Cloning of *Pseudomonas* genes in *E. coli. In* The Bacteria: A Treatise on Structure and Function (Sokatch, J. R. ed.), vol X pp357-378

Nakazawa, A., Inouye, S., Gomada, M., Imaishi, H., and Nakazawa, T. 1992. Upstream regulatory sequence for transcriptional activation of catabolic genes on the TOL plasmid.*In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.) pp 353-357. American Society For Microbiology, Washington, D.C.

Nakazawa, T., Inouye, S., and Nakazawa, A. 1990. Regulatory systems for expression of *xyl* genes on the TOL plasmid. *In Pseudomonas*: Biotransformations, Pathogenesis, and evolving Biotechnology (Silver, S., Chakrabarty, A. M., Iglewski, B., and Kaplan, S. ed.), pp 133-140. American Society for Microbiology.

Nodwell, J. R., and Greenblatt, J. 1991. The *nut* site of bacteriophage σ is made of RNA and is bound by transcription antitermination factors on the surface of RNA polymerase. Genes Dev. 5: 2141-2151

Novick, R. P., Clowes, R. C., Cohen, S. N., Curtiss III, R., Datta, N., and Falkow, S. 1976. Uniform nomenclature for bacterial plasmids. Bacteriol. Rev. 40: 168-189

Ogata, R. T., and Gilbert, W. 1978. An amino-terminal fragment of *lac* repressor binds specifically to *lac* operator. Proc. Natl. Acad. Sci. USA 75: 5851-5854

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. 1985. Structure of large fragments of *Escherichia coli* DNA Polymerase I complexed with dTMP. Nature 313: 762-766

Olson, E. R., Flamm, E. L., and Friedman, D. I. 1982. analysis of *nutR*: a region of phage lambda required for antitermination of transcription. Cell 31: 61-70

Otridge, J., and Gollnick, P. 1993. MtrB from *Bacillus subtilis* binds specifically to *trp* leader RNA in a tryptophan-dependent manner. Proc. Natl. Acad. Sci. USA 90: 128-132

Palleroni, N. J. 1992. Present situation of the taxonomy of aerobic pseudomonads. *In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.), pp 105-115. American Society For Microbiology, Washington, D.C.

Palleroni, N. J. 1984. Taxonomy of pseudomonads. *In* The bacteria: a treatise on structure and function (Sokatch, J. R. ed.) vol. X pp 3-25. Academic Press, Inc.

Platt, T. 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. 55: 339-372

Query, C. C., Bently, R. C., and Keene, J. D. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70 K U1 snRNP protein. Cell. 57: 89-101

Palleroni, N. J., Ballard, R. W., Ralston, E., and Duodoroff, M. 1972. Deoxyribonucleic acid homologies among some *Pseudomonas* species. J. Bacteriol. 110: 1-11

Rehmat, S., and Shapiro, J. A. 1983. Insertion and replication of of *Pseudomonas aeruginosa* mutator phage D3112. Mol. Gen. Genet. 192: 416-423

Popham, D. L., Szeto, D., Keener, J., and Kustu, S. 1989. Function of bacterial activator protein that binds to transcriptional enhancers. Science 243: 629-635

Potts, J. R., and Clarke, P. H. 1976. Effect of nitrogen limitation on catabolite repression of amidase, histidase and urocanase in *Pseudomonas aeruginosa*. J. Gen. Microbiol. 116: 381-389

Razin, A., Hirose, T., Itakura, T., Riggs, A. D.1978. Efficient correction of a mutation by use of chemically synthesized DNA. Proc. Natl. Acad. Sci. USA 75: 4268-4270

Reines, D. 1992. Elongation factor-dependent transcript shortening by the template-engaged RNA polymerase II. J. Biol. Chem. 267: 3795-3800

Reines, D., Chamberlin, M. J., and Kane, C. M. 1989. Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block of transcription in a human gene *in vitro*. J. Biol. Chem. 264: 10799-10809

Reynolds, R., and Chamberlin, M. J. 1992. Parameters affecting transcription termination by *Escherichia coli* RNA polymerase II. Construction and analysis of hybrid terminators. J. Mol. Biol. 224: 53-63

Reynolds, R., and Chamberlin, M. J. 1992. Parameters affecting transcription termination by *Escherichia coli* RNA polymerase I. Analysis of 13 Rho-independent terminators. J. Mol. Biol. 224: 31-51

Reznikoff, W. S. 1983. Some bacterial transposable elements: their organization, mechanism of transposition, and role in genome evolution. *In* Gene function in prokaryotes (Beckwith, J., Davies, J. A., and Galiant, J. A. ed.), pp 229-252. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

Reznikoff, W. S. 1992. Catabolite gene activator protein activation of *lac* transcription. J. Bacteriol. 174: 655-658

Richardson, J. P. 1990. Rho-dependent transcription termination. *Biochim. Biophys. Acta.* 1048: 127-138

Roberts, J. W. 1993. RNA and protein elements of *E. coli* and λ transcription antitermination complexes. Cell. 653-655

Rothmel, R. K., Chakrabarty, A. B., and Darzins, A. 1991. Genetic systems in *Pseudomonas. In* Methods in Enzymology (Miller, J.H. ed), 204: 485-540

Saier, M. H., Jr., Yamada, M., Erni, B., Suda, K., Lengder, J., Ebner, R., Argos, P., Rak, B., Schnetz, K., Lee, C. A., Stewart, G. G., Breidt, F., Waygood, E. B., Peri, K. G., and Doolittle, R. F. 1988. Sugar permeases of the bacterial phosphoenolpyruvate-dependent phosphotransferase system: sequence comparisions. FASEB J. 2: 199-208

Sambrook, J., Fritsch, C. F., and Maniatis, T. 1989. Molecular cloning - *a laboratory manual manual*. Second edition. Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467

Sasse-Dwight, S., and Gralla, J. D. 1988. Probing of the *E. coli glnALG* upstream activation mechanism *in vivo*. Proc. Natl. Acad. Sci. USA 85: 8934-8938

Schell, M. A., and Sukordhaman, M. 1989. Evidence that the transcription activator encoded by the *Pseudomonas putida nahR* gene is evolutionary related to the transcription activators encoded by the *Rhizobium nodD* genes. J. Bacteriol. 171: 1952-1959

Schleif, R. 1992. Regulation of the L-arabinose catabolic operon araBAD. *In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.), pp 643-665. Cold Spring Harbor Laboratory Press

Schnetz, K., Toloczyki, C., and Rak, B. 1987. β-glucoside (*bgl*) operon of *Escherichia coli* K12: nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. J. Bacteriol. 169: 2579-2790

Sentenac, A., Riva, M., Thuriaux, P., Buhler, J., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., Mann, C., Chiannilkulchai, N., Stettler, S., and Mariotte, S. 1992. Yeast RNA polymerase subunits and genes. *In* Transcriptional Regulation (McKnight, S.

L. and Yamamoto, K. R. ed.), pp 27-54. Cold Spring Harbor Laboratory Press

Shimotsu, H., and Henner, D. J. 1986. Modulation of *Bacillus subtilis* levansucrase gene expression by sucrose and regulation of the steady-state mRNA level by *sacU* and *sacQ* genes. J. Bacteriol. 168: 380-388

Shimotsu, H., Kuroda, M. I., Yanofsky, C., and Henner, D. J. 1986. A novel form of transcription attenuation regulates expression of *Bacillus subtilis* tryptophan operon. J. Bacteriol. 166: 461-471

Shine, J. and Dalgarno, L. 1975. Determinant of cistron specificity in bacterial ribosomes. Nature. 254: 34-38

SivaRaman, L., Reines, B., and Kane, C. M. 1990. Purified elongation factor S II is sufficient to promote readthrough by purified RNA polymerase II at specific termination sites in the human histone H3.3 gene. J. Biol. Chem. 265: 14554-14560

Sluder, A.E., Greenleaf, A. L., and Price, D. H. 1989. Properties of a *Drosophila* RNA polymerase II elongation factor. J. Biol. Chem. 264: 8963-8969

Smith, M. 1985. *In vitro* mutagenesis. Annu. Rev. Genet. 19: 425-462

Smyth, P. F., and Clarke, P. H. 1975. Catabolite repression of *Pseudomonas aeruginosa* amidase: isolation of promoter mutants. J. Gen. Microbiol. 90: 91-99

Soubrier, F., Levy-Schil, S., Mayaux, J. -F., Petre, D., Arnaud, A., and Crouzet, J. 1992. Cloning and primary structure of the wide-spectrum amidase from *Brevibacterium sp.* R312: high homology to the *amiE* product from *Pseudomonas aeruginosa.* Gene 116: 99-104

Sousa, R., Patra, D., Lafer, E.M. 1992. Model for the mechanism of bacteriophage T7 RNAP transcription initiation and termination. J. Mol. Biol. 224: 319-334

Stanier, R. Y., Watcher, D., Gasser, C., and Wilson, A. C. 1970. Comparative immunological studies of two *Pseudomonas* enzymes. J. Bacteriol. 102: 351-362

Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Tre'boul, G., and Gay, P. 1985. The DNA sequence of the gene for secreted *Bacillus subtilis* enzyme levansucrase and its genetic control sites. Mol. Gen. Genet. 200: 220-228

Stock, J. B., Stock, A. M., and Mottonen, A. M. 1990. Signal transduction in bacteria. Nature (London) 344: 395-400

Stokes, H. W., and Hall, R. M. 1989. A novel family of potentially mobile DNA elements encoding site-specific gene integration functions: Integrons. Mol. Microbiol. 3: 1227-1233

Straney, S. B., and Crothers, D. M. 1987. *lac* repressor is a transient gene-activating protein. Cell 51: 699-707

Studier, W. F., and Moffat, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189: 113-130

Su, W., Porter, S., Kustu, S., and Echols, H. 1990. DNA looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. Proc. Natl. Acad. Sci. USA 87: 5504-5508

Surrat, C. K., Milan, S., and Chamberlin, M. J. 1991. Spontaneous cleavage of RNA in ternary complexes of *E. coli* RNA polymerase, and its significance for the mechanism of transcription. Proc. Natl. Acad. Sci., USA. 88: 7983-7987

Tabor, S. 1990. Expression using the T7 RNA polymerase/promoter system. *In* Current Protocals in Molecular Biology (Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. eds.), pp 16.2.1-16.2.11. Green Publishing and Wiley-Interscience, New York.

Tabor, S. and Richardson, C.C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA 82: 1074-1078

Tam, R., and Saier, M. H. 1993. Structural, functional, and evolutionary relationship among extracellular solute-binding receptors of bacteria. Microbiol. Rev. 57: 320-346

Tautz, D., and Renz, M. 1983. An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. Anal. Biochem.132: 14-19

Telesnitsky, A., and Chamberlin, M. J. 1989. Terminator-distal sequences determine the *in vitro* effeciency of the early terminators of bacteriophages T3 and T7. Biochemistry 28: 5210-5218

Thony, B., and Hennecke, H. 1989. The -24/-12 promoter comes of age. FEMS Microbiol. Rev. 63: 341-358

Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA . 76: 4350-4354

Tsuda, M., and Lino, T. 1987. Genetic analysis of a transposon carrying toluene degrading genes on a TOL plamid pWWO. Mol. Gen. Genet. 210: 270-276

Tummler, B., Rohmling, U., Ratnaningsih, E., Morgan, A. F., Krishnapillai, V., and Holloway, B. W. 1992. *In Pseudomonas* : Molecular Biology and Biotechnology (Galli, E. Silver, S., and Witholt, B. ed.), pp 9-11. American Society For Microbiology, Washington, D.C.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1: 945-991

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. and Weiner, A. M. 1987. Molecular Biology of the Gene. The Benjamin/Cummings publishing company, Inc.

Watson, J. M., and Holloway, B. W. 1978. Chromosome mapping in *Pseudomonas aeruginosa* PAT. J. Bacteriol. 133: 1113-1125

Weber, I. T., and Steitz, T. A. 1984. Model of specific complex between catabolite gene activator protein and B-DNA suggested by electrostatic complementarity. Proc. Natl. Acad. Sci. USA 81: 3973-3977

Weiss, D. S., Klose, K. E., Hoover, T. R., North, A. K., Porter, S. C., Wedel, A. B., and Kustu, S. 1992. Prokaryotic transcription enhancers.*In* Transcriptional Regulation (McKnight, S. L. and Yamamoto, K. R. ed.), pp 667-694. Cold Spring Harbor Laboratory Press

West, S. E. H., and Iglewski, B. H. 1988. Codon usage in *Pseudomonas aeruginosa.* Nucl. Acids Res. 16: 9323-9335

Whitaker, R. J., Byng. G. S., Gherna, R. L., and Jensen, R. A. 1981. Diverse enzymological patterns of phenylalanine biosynthesis in pseudomonad bacteria are

conserved in parallel with DNA/DNA homology groupings. J. Bacteriol. 147: 526-534

Wilson, S. A. W. 1991. Regulation of *Pseudomonas aeruginosa* amidase expression by AmiC. PhD thesis University of London

Wilson, S. A., Chayen, N. E., Hemmings, A. M., Drew, R. E., and Pearl, L. H. 1991. Crystallization and preliminary X-ray data for the negative regulator (AmiC) of the amidase operon of *Pseudomonas aeruginosa.* J. Mol. Biol. 222: 869-871

Wilson, S., and Drew, R. 1991. Cloning and DNA sequence of *amiC*, a new gene regulating expression of the *Pseudomonas aeruginosa* aliphatic amidase, and purification of the *amiC* product. J. Bacteriol. 173: 4914-4921

Wilson, S. A., Wachira, S. J., Drew, R. E., Jones, D., and Pearl, L. H. 1993. Antitermination of amidase expression in *Pseudomonas aeruginosa* is controlled by a novel cytoplasmic amide binding protein. EMBO J. 12: 3637-3642

Wong, P. -K., Popham, D., Keener, J., and Kustu, S. 1987. *In vitro* transcription of the nitrogen regulatory operon *nifLA* of *Klebsiella pneumoniae*. J. Bacteriol. 169: 2876-2880

Worsey, M. J., and Williams, P. A. 1975. Metabolism of toluene and xylenes by *Pseudomonas putida (arvilla) mt-2*: evidence for a new function of the TOL plasmid. J. Bacteriol. 124: 7-13

Wozniak, D. J., and Ohman, D. E. 1991. *Pseudomonas aeruginosa* AlgB, a two-component response regulator of the NtrC family, is required for *algD* transcription. J. Bacteriol. 173: 1406-1413

Wu, C. W., Wu, F. Y. H., and Speckhard, D. C. 1977. Subunit location of the intrinsic divalent metal ions in RNA polymerase from *Escherichia coli.* Biochemistry 16: 5449-5454

Yager, T. D., and von Hippel, P. 1987. Transcription elongation and termination in *Escherichia coli. In Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology (Neidhardt, F. C. ed.), pp 1241-1275. American Society for Microbiology, Washington, D. C.

Yang, C., and Nash, H. 1989. Interaction of *E.coli* IHF protein with specific binding sites. Cell 57: 869-880

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119

Yanofsky, C. 1981. Attenuation in the control of expression of bacterial operons. Nature (London) 289: 751-758

Yanofsky, C. 1988. Transcription attenuation. J. Biol. Chem. 263: 609-612

Yen, K. -M., and Gunsalus, I. C. 1982. Plasmid organization: Naphthalene/Salicylate oxidation. Proc. Natl. Acad. Sci. USA. 79: 875-878

Zahn, K., and Blattner, F. R. 1987. Direct evidence for DNA bending at the Lambda replication origin. Science 236: 416-422

Zukowski, M. M., Miller, L., Cogswell, P., Chen, K., Aymerich, S., and Steinmetz, M. 1990. Nucleotide sequence of the *sacS* locus of *Bacillus subtilis* reveals the presence of two regulatory genes. Gene 90: 153-155

Appendix A

The Complete DNA Sequence of the 5.3kb *Hin*dIII/*Sal*I Amidase Operon of *Pseudomonas aeruginosa.*

Regulation of the *Pseudomonas aeruginosa* Amidase Operon by Transcription Antitermination

by

Munuhe Simon James Wachira

A thesis submitted for the Degree of Doctor of Philosophy in the University of London

Department of Biochemistry and Molecular Biology
University College London
January 1994

```
                        20                  40
     *    *     *    *     *    *     *    *     *    *
AAGCTTCCGT GCGAATGATG GCATGCATGC TATCTCAGGC TCGCACCATG

                                      >-35_promoter_site
                                              |
          60                    80            |     100
     *    *     *    *     *    *     *    *  |  *    *
TGCTTTCGCG ATCGCGCCGA TTACATAACG TTACACGAAC CTTGACAGCC

>-10_amiE_promoter_sequence   >amiE_transcription_initiation_site
                        |             |
                        | 120         |         140
     *    *     *|   *     *   |*     *    *     *    *
CCTTCCGACG GGGCTTATAA GTGGCGCCAT CAGGTCATGC GCATCAGCGT
                                      Met ArgIleSerVal>

                             >leader_transcription_terminator
                                              |
        160                   180             |        200
     *    *     *    *     *    *     *    *     *    *
CGATGTCGCG GGACCGAACC TAACGCATAC GCACAGAGCA AATGGGCTCT
 AspValAla GlyProAsn LeuThrHisThr HisArgAla AsnGlyLeu>

                                          >amiE_S/D_sequence
                                                    |
                    220                   240         |
     *    *     *    *     *    *     *    *     *   |*
CCCGGGGTTA CCCGGGAGGG CCTTTTTTTC GTCCCAAAAA ATAACAACAA
SerArgGlyTyr ProGlyGly ProPhePhe ArgProLysLys ***>

        260                   280                   300
     *    *     *    *     *    *     *    *     *    *
GAGGTGATAC CCATGCGTCA CGGCGATATT TCCAGCAGCA ACGACACCGT
             MetArgHis GlyAspIle SerSerSer AsnAspThrVal>

                    320                   340
     *    *     *    *     *    *     *    *     *    *
CGGAGTGGCC GTGGTCAACT ACAAGATGCC GCGCCTGCAC ACCGCGGCGG
 GlyValAla ValValAsn TyrLysMetPro ArgLeuHis ThrAlaAla>

        360                   380                   400
     *    *     *    *     *    *     *    *     *    *
AGGTCCTGGA CAACGCCCGG AAGATCGCCG ACATGATCGT CGGCATGAAG
GluValLeuAsp AsnAlaArg LysIleAla AspMetIleVal GlyMetLys>

                    420                   440
     *    *     *    *     *    *     *    *     *    *
CAGGGCCTGC CCGGCATGGA CCTGGTGGTG TTCCCGGAGT ACAGCCTGCA
GlnGlyLeu ProGlyMetAsp LeuValVal PheProGlu TyrSerLeuGln>

        460                   480                   500
     *    *     *    *     *    *     *    *     *    *
GGGCATCATG TACGATCCGG CGGAGATGAT GGAAACCGCG GTGGCGATCC
 GlyIleMet TyrAspPro AlaGluMetMet GluThrAla ValAlaIle>

                    520                   540
     *    *     *    *     *    *     *    *     *    *
CCGGCGAGGA AACCGAGAGA TTCTCCCGCG CCTGCCGCAA GGCCAACGTC
ProGlyGluGlu ThrGluArg PheSerArg AlaCysArgLys AlaAsnVal>

        560                   580                   600
     *    *     *    *     *    *     *    *     *    *
TGGGGCGTAG TCTCCCTCAC CGGCGAACGG CACGAGGAGC ATCCGCGCAA
TrpGlyVal ValSerLeuThr GlyGluArg HisGluGlu HisProArgAsn>

                    620                   640
     *    *     *    *     *    *     *    *     *    *
```

```
TGCGCCGTAC AACACCCTGG TGCTGATCGA CAACAACGGC GAGATCGTCC
 AlaProTyr AsnThrLeu ValLeuIleAsp AsnAsnGly GluIleVal>

       660                   680                   700
    *    *     *    *     *    *     *    *     *    *
AGAAGTACCG CAAGATCATT CCCTGGTGCC CCATCGAGGG CTGGTATCCC
GlnLysTyrArg LysIleIle ProTrpCys ProIleGluGly TrpTyrPro>

                  720                   740
    *    *     *    *     *    *     *    *     *    *
GGTGGCCAGA CCTACGTCAG CGAAGGGCCG AAGGGCATGA AGATCAGCCT
GlyGlyGln ThrTyrValSer GluGlyPro LysGlyMet LysIleSerLeu>

       760                   780                   800
    *    *     *    *     *    *     *    *     *    *
GATCATCTGC GACGACGGCA ATTACCCGGA GATCTGGCGC GACTGCGCGA
 IleIleCys AspAspGly AsnTyrProGlu IleTrpArg AspCysAla>

                  820                   840
    *    *     *    *     *    *     *    *     *    *
TGAAGGGCGC CGAGCTGATC GTGCGCTGCC AGGGCTACAT GTACCCGGCC
MetLysGlyAla GluLeuIle ValArgCys GlnGlyTyrMet TyrProAla>

       860                   880                   900
    *    *     *    *     *    *     *    *     *    *
AAGGACCAGC AGGTGATGAT GGCCAAGGCC ATGGCCTGGG CCAACAACTG
LysAspGln GlnValMetMet AlaLysAla MetAlaTrp AlaAsnAsnCys>

                  920                   940
    *    *     *    *     *    *     *    *     *    *
CTATGTGGCG GTGGCCAACG CGGCCGGCTT CGACGGTGTC TATTCCTACT
 TyrValAla ValAlaAsn AlaAlaGlyPhe AspGlyVal TyrSerTyr>

       960                   980                  1000
    *    *     *    *     *    *     *    *     *    *
TCGGCCACTC GGCGATCATC GGCTTCGACG GCCGTACCCT CGGTGAGTGC
PheGlyHisSer AlaIleIle GlyPheAsp GlyArgThrLeu GlyGluCys>

                 1020                  1040
    *    *     *    *     *    *     *    *     *    *
GGCGAGGAGG AAATGGGTAT CCAGTACGCC CAGCTGTCCC TTTCGCAGAT
GlyGluGlu GluMetGlyIle GlnTyrAla GlnLeuSer LeuSerGlnIle>

      1060                  1080                  1100
    *    *     *    *     *    *     *    *     *    *
CCGCGATGCG CGCGCCAACG ATCAGTCGCA GAACCACCTG TTCAAGCTCC
 ArgAspAla ArgAlaAsn AspGlnSerGln AsnHisLeu PheLysLeu>

                 1120                  1140
    *    *     *    *     *    *     *    *     *    *
TCCACCGCGG CTACAGCGGC TTGCAGGCGT CCGGCGACGG CGACCGGGGC
LeuHisArgGly TyrSerGly LeuGlnAla SerGlyAspGly AspArgGly>

      1160                  1180                  1200
    *    *     *    *     *    *     *    *     *    *
CTGGCGGAGT GTCCGTTCGA GTTCTACCGC ACCTGGGTCA CCGACGCCGA
LeuAlaGlu CysProPheGlu PheTyrArg ThrTrpVal ThrAspAlaGlu>

                 1220                  1240
    *    *     *    *     *    *     *    *     *    *
GAAGGCGCGC GACAATGTCG AGCGACTGAC CCGCTCGACC ACCGGCGTGG
 LysAlaArg AspAsnVal GluArgLeuThr ArgSerThr ThrGlyVal>

      1260                  1280                  1300
    *    *     *    *     *    *     *    *     *    *
CGCAATGCCC GGTCGGCCGG CTGCCCTATG AGGGACTGGA GAAGGAAGCC
AlaGlnCysPro ValGlyArg LeuProTyr GluGlyLeuGlu LysGluAla>
```

```
          >amiE_transcription_terminator
                          |
              1320        |              1340
       *       *       *       *   |  *       *       *       *       *       *
  TGACGGCAGA CGCCGCCAGC CCGGCGTGCC GTCGTGCGGC ACGCCGTCTC

                 >amiB_S/D_sequence
                          |
       1360               |    1380                      1400
       *       *       *       *  |  *       *       *       *       *       *
  CGCCATTTCC CCCCTGTGGC AGAAGGAGTT TCATCCATGC CTTTTCTGAG
                                       Met ProPheLeuSer>

              1420                  1440
       *       *       *       *       *       *       *       *       *       *
  CGACATGCTC GACCAGTCCC GCCGGCAGCA GGACGAGGAA CAGGCCCTGG
   AspMetLeu AspGlnSer ArgArgGlnGln AspGluGlu GlnAlaLeu>

       1460                1480                      1500
       *       *       *       *       *       *       *       *       *       *
  CGCGGGAGAA TCTTGCCGAG GCAAGTCTGC TCCAGGCCCA CCTGAGTCAC
 AlaArgGluAsn LeuAlaGlu AlaSerLeu LeuGlnAlaHis LeuSerHis>

              1520                  1540
       *       *       *       *       *       *       *       *       *       *
  CGCAGCGCCC TGCACAGCCG TTTCCGTTTC GACCCGGCGG CGGTGATGGA
  ArgSerAla LeuHisSerArg PheArgPhe AspProAla AlaValMetAsp>

       1560                1580                      1600
       *       *       *       *       *       *       *       *       *       *
  CTGCCTGCGC GCCGAGGTGC TCGGCCAGGA ACCGGCGCTA CAGGCCGTCG
   CysLeuArg AlaGluVal LeuGlyGlnGlu ProAlaLeu GlnAlaVal>

              1620                  1640
       *       *       *       *       *       *       *       *       *       *
  AGGACATGCT CAAGGTGGTT CGCGCGGACA TCGCCGACCC GCGCCGTCCG
 GluAspMetLeu LysValVal ArgAlaAsp IleAlaAspPro ArgArgPro>

       1660                1680                      1700
       *       *       *       *       *       *       *       *       *       *
  CTGTTCAGCG CGCTGTTCCT CGGCCCCACC GGAGTCGGCA AGACCGAGAT
  LeuPheSer AlaLeuPheLeu GlyProThr GlyValGly LysThrGluIle>

              1720                  1740
       *       *       *       *       *       *       *       *       *       *
  CGTGCGCGCC CTGGCCAGGG CCCTGCACGG CGACGCCGAG GGGTTCTGCC
   ValArgAla LeuAlaArg AlaLeuHisGly AspAlaGlu GlyPheCys>

       1760                1780                      1800
       *       *       *       *       *       *       *       *       *       *
  GGGTGGACAT GAACACCCTG TCCCAGGAGC ACTATGCCGC CGCCCTCACC
 ArgValAspMet AsnThrLeu SerGlnGlu HisTyrAlaAla AlaLeuThr>

              1820                  1840
       *       *       *       *       *       *       *       *       *       *
  GGTGCGCCGC CGGGCTACGT CGGGGCGAAG GAGGGCACCA CCCTGTTGGA
  GlyAlaPro ProGlyTyrVal GlyAlaLys GluGlyThr ThrLeuLeuGlu>

       1860                1880                      1900
       *       *       *       *       *       *       *       *       *       *
  GCAGGACAAG CTGGACGGCA GTCCCGGGCG CCCCGGCATC GTTCTCTTCG
   GlnAspLys LeuAspGly SerProGlyArg ProGlyIle ValLeuPhe>

              1920                  1940
       *       *       *       *       *       *       *       *       *       *
  ACGAACTGGA AAAGGCCAGC CCGGAAGTGG TCCATGCGTT GCTCAACGTA
 AspGluLeuGlu LysAlaSer ProGluVal ValHisAlaLeu LeuAsnVal>
```

```
          1960                  1980                  2000
      *    *     *    *     *    *     *    *     *    *
CTCGACAACG GCCTGCTACG GGTCGCTTCC GGCGAACGCA CCTACCATTT
LeuAspAsn GlyLeuLeuArg ValAlaSer GlyGluArg ThrTyrHisPhe>

                  2020                  2040
      *    *     *    *     *    *     *    *     *    *
CCGCAACACC CTGGTGTTCA TGACCAGCAA TCTCTGCGCC CATGAGATCC
 ArgAsnThr LeuValPhe MetThrSerAsn LeuCysAla HisGluIle>

          2060                  2080                  2100
      *    *     *    *     *    *     *    *     *    *
AGCGCTACGA CGAGCGTCGC CAGCGCCTGC CCTGGCGCCT GCTGCCGGTC
GlnArgTyrAsp GluArgArg GlnArgLeu ProTrpArgLeu LeuProVal>

                  2120                  2140
      *    *     *    *     *    *     *    *     *    *
GGCGGCGAGC GCCGGCGGCG GGACATCGAC GGGATGGTCC GGGCCCGGCT
GlyGlyGlu ArgArgArgArg AspIleAsp GlyMetVal ArgAlaArgLeu>

          2160                  2180                  2200
      *    *     *    *     *    *     *    *     *    *
GCTGAAGACC TTCTCGCCGG AGTTCGTCAA TCGTCTCGAT AGCGTGGTCA
 LeuLysThr PheSerPro GluPheValAsn ArgLeuAsp SerValVal>

                  2220                  2240
      *    *     *    *     *    *     *    *     *    *
CCTTCAACTG GATCGAACGC GACGTCGTCG CGCGCCTGGT CGAGCTGGAG
ThrPheAsnTrp IleGluArg AspValVal AlaArgLeuVal GluLeuGlu>

          2260                  2280                  2300
      *    *     *    *     *    *     *    *     *    *
GTGCAGCGGC TCAACCGGCG CCTGGAGAAG CATCGCTGCC GCCTGGAGGC
ValGlnArg LeuAsnArgArg LeuGluLys HisArgCys ArgLeuGluAla>

                  2320                  2340
      *    *     *    *     *    *     *    *     *    *
GACCCCGGAG GTGCTGGCGA AGATCGCCCG CGCCGGCTTC GACCGGCAGT
 ThrProGlu ValLeuAla LysIleAlaArg AlaGlyPhe AspArgGln>

          2360                  2380                  2400
      *    *     *    *     *    *     *    *     *    *
TCGGCGCCCG TGCGTTGCGC CGCAGCGTCC GTCATCATCT CGAGGTACCG
PheGlyAlaArg AlaLeuArg ArgSerVal ArgHisHisLeu GluValPro>

>rpoN_promoter
      |
      |           2420                  2440
      |   *    *     *    *     *    *     *    *     *    *
      CTGGCCGAGC ATCTGCTCGA TCACCACCAG CCGGGCGACG GGAACTGCAC
      LeuAlaGlu HisLeuLeuAsp HisHisGln ProGlyAsp GlyAsnCysThr>

       >rpoN_promoter
            |
          2460                  2480                  2500
      *   | *     *    *     *    *     *    *     *    *
GATCTACCTG GCGAGCCTGG AGCACGAGCG GGTTCGCTTC GTACGGCGCT
 IleTyrLeu AlaSerLeu GluHisGluArg ValArgPhe ValArgArg>

          >amiC_S/D_sequence
                  |
                  |  2520                  2540
      *    *      |*    *     *    *     *    *     *    *
GAGCGACAGT CACAGGAGAG GAAACGGATG GGATCGCACC AGGAGCGGCC
                                Met GlySerHis GlnGluArgPro>

          2560                  2580                  2600
      *    *     *    *     *    *     *    *     *    *
```

```
GCTGATCGGC CTGCTGTTCT CCGAAACCGG CGTCACCGCC GATATCGAGC
 LeuIleGly LeuLeuPhe SerGluThrGly ValThrAla AspIleGlu>

                2620                    2640
    *    *    *    *    *    *    *    *    *    *
GCTCGCAGCG CTATGGCGCA TTGCTCGCGG TCGAGCAACT GAACCGCGAG
ArgSerGlnArg TyrGlyAla LeuLeuAla ValGluGlnLeu AsnArgGlu>

      2660                    2680                    2700
    *    *    *    *    *    *    *    *    *    *
GGCGGCGTCG GCGGTCGCCC GATCGAAACG CTGTCCCAGG ACCCCGGCGG
GlyGlyVal GlyGlyArgPro IleGluThr LeuSerGln AspProGlyGly>

                2720                    2740
    *    *    *    *    *    *    *    *    *    *
CGACCCGGAC CGCTATCGGC TGTGCGCCGA GGACTTCATT CGCAACCGGG
 AspProAsp ArgTyrArg LeuCysAlaGlu AspPheIle ArgAsnArg>

      2760                    2780                    2800
    *    *    *    *    *    *    *    *    *    *
GGGTACGGTT CCTCGTGGGC TGCTACATGT CGCACACGCG CAAGGCGGTG
GlyValArgPhe LeuValGly CysTyrMet SerHisThrArg LysAlaVal>

                2820                    2840
    *    *    *    *    *    *    *    *    *    *
ATGCCGGTGG TCGAGCGCGC CGACGCGCTG CTCTGCTACC CGACCCCCTA
MetProVal ValGluArgAla AspAlaLeu LeuCysTyr ProThrProTyr>

      2860                    2880                    2900
    *    *    *    *    *    *    *    *    *    *
CGAGGGCTTC GAGTATTCGC CGAACATCGT CTACGGCGGT CCGGCGCCGA
 GluGlyPhe GluTyrSer ProAsnIleVal TyrGlyGly ProAlaPro>

                2920                    2940
    *    *    *    *    *    *    *    *    *    *
ACCAGAACAG TGCGCCGCTG GCGGCGTACC TGATTCGCCA CTACGGCGAG
AsnGlnAsnSer AlaProLeu AlaAlaTyr LeuIleArgHis TyrGlyGlu>

      2960                    2980                    3000
    *    *    *    *    *    *    *    *    *    *
CGGGTGGTGT TCATCGGCTC GGACTACATC TATCCGCGGG AAAGCAACCA
ArgValVal PheIleGlySer AspTyrIle TyrProArg GluSerAsnHis>

                3020                    3040
    *    *    *    *    *    *    *    *    *    *
TGTGATGCGC CACCTGTATC GCCAGCACGG CGGCACGGTG CTCGAGGAAA
 ValMetArg HisLeuTyr ArgGlnHisGly GlyThrVal LeuGluGlu>

      3060                    3080                    3100
    *    *    *    *    *    *    *    *    *    *
TCTACATTCC GCTGTATCCC TCCGACGACG ACTTGCAGCG CGCCGTCGAG
IleTyrIlePro LeuTyrPro SerAspAsp AspLeuGlnArg AlaValGlu>

                3120                    3140
    *    *    *    *    *    *    *    *    *    *
CGCATCTACC AGGCGCGCGC CGACGTGGTC TTCTCCACCG TGGTGGGCAC
 ArgIleTyr GlnAlaArgAla AspValVal PheSerThr ValValGlyThr>

      3160                    3180                    3200
    *    *    *    *    *    *    *    *    *    *
CGGCACCGCC GAGCTGTATC GCGCCATCGC CCGTCGCTAC GGCGACGGCA
 GlyThrAla GluLeuTyr ArgAlaIleAla ArgArgTyr GlyAspGly>

                3220                    3240
    *    *    *    *    *    *    *    *    *    *
GGCGGCCGCC GATCGCCAGC CTGACCACCA GCGAGGCGGA GGTGGCGAAG
ArgArgProPro IleAlaSer LeuThrThr SerGluAlaGlu ValAlaLys>
```

```
      3260                    3280                    3300
    *    *     *    *     *    *     *    *     *    *
ATGGAGAGTG ACGTGGCAGA GGGGCAGGTG GTGGTCGCGC CTTACTTCTC
MetGluSer AspValAlaGlu GlyGlnVal ValValAla ProTyrPheSer>

                3320                    3340
    *    *     *    *     *    *     *    *     *    *
CAGCATCGAT ACGCCCGCCA GCCGGGCCTT CGTCCAGGCC TGCCATGGTT
 SerIleAsp ThrProAla SerArgAlaPhe ValGlnAla CysHisGly>

      3360                    3380                    3400
    *    *     *    *     *    *     *    *     *    *
TCTTCCCGGA GAACGCGACC ATCACCGCCT GGGCCGAGGC GGCCTACTGG
PhePheProGlu AsnAlaThr IleThrAla TrpAlaGluAla AlaTyrTrp>

                3420                    3440
    *    *     *    *     *    *     *    *     *    *
CAGACCTTGT TGCTCGGCCG CGCCGCGCAG GCCGCAGGCA ACTGGCGGGT
GlnThrLeu LeuLeuGlyArg AlaAlaGln AlaAlaGly AsnTrpArgVal>

      3460                    3480                    3500
    *    *     *    *     *    *     *    *     *    *
GGAAGACGTG CAGCGGCACC TGTACGACAT CGACATCGAC GCGCCACAGG
 GluAspVal GlnArgHis LeuTyrAspIle AspIleAsp AlaProGln>

                3520                    3540
    *    *     *    *     *    *     *    *     *    *
GGCCGGTCCG GGTGGAGCGC CAGAACAACC ACAGCCGCCT GTCTTCGCGC
GlyProValArg ValGluArg GlnAsnAsn HisSerArgLeu SerSerArg>

      3560                    3580                    3600
    *    *     *    *     *    *     *    *     *    *
ATCGCGGAAA TCGATGCGCG CGGCGTGTTC CAGGTCCGCT GGCAGTCGCC
IleAlaGlu IleAspAlaArg GlyValPhe GlnValArg TrpGlnSerPro>

                3620                    3640
    *    *     *    *     *    *     *    *     *    *
CGAACCGATT CGCCCCGACC CTTATGTCGT CGTGCATAAC CTCGACGACT
 GluProIle ArgProAsp ProTyrValVal ValHisAsn LeuAspAsp>

           >amiR_S/D_sequence
                    |
      3660          |     3680                    3700
    *    *     *    |*    *     *    *     *    *     *    *
GGTCCGCCAG CATGGGCGGG GGACCGCTCC CATGAGCGCC AACTCGCTGC
                                  MetSerAla AsnSerLeu>
TrpSerAlaSer MetGlyGly GlyProLeu Pro>

                3720                    3740
    *    *     *    *     *    *     *    *     *    *
TCGGCAGCCT GCGCGAGTTG CAGGTGCTGG TCCTCAACCC GCCGGGGGAG
LeuGlySerLeu ArgGluLeu GlnValLeu ValLeuAsnPro ProGlyGlu>

      3760                    3780                    3800
    *    *     *    *     *    *     *    *     *    *
GTCAGCGACG CCCTGGTCTT GCAGCTGATC CGCATCGGTT GTTCGGTGCG
ValSerAsp AlaLeuValLeu GlnLeuIle ArgIleGly CysSerValArg>

                3820                    3840
    *    *     *    *     *    *     *    *     *    *
CCAGTGCTGG CCGCCGCCGG AAGCCTTCGA CGTGCCGGTG GACGTGGTCT
 GlnCysTrp ProProPro GluAlaPheAsp ValProVal AspValVal>

      3860                    3880                    3900
    *    *     *    *     *    *     *    *     *    *
TCACCAGCAT TTTCCAGAAT GGCCACCACG ACGAGATCGC TGCGCTGCTC
PheThrSerIle PheGlnAsn GlyHisHis AspGluIleAla AlaLeuLeu>
```

```
                    3920                3940
      *    *    *    *    *    *    *    *    *
GCCGCCGGGA CTCCGCGCAC TACCCTGGTG GCGCTGGTGG AGTACGAAAG
AlaAlaGly ThrProArgThr ThrLeuVal AlaLeuVal GluTyrGluSer>

      3960                3980                4000
   *    *    *    *    *    *    *    *    *    *
CCCCGCGGTG CTCTCGCAGA TCATCGAGCT GGAGTGCCAC GGCGTGATCA
 ProAlaVal LeuSerGln IleIleGluLeu GluCysHis GlyValIle>

                    4020                4040
      *    *    *    *    *    *    *    *    *    *
CCCAGCCGCT CGATGCCCAC CGGGTGCTGC CTGTGCTGGT ATCGGCGCGG
ThrGlnProLeu AspAlaHis ArgValLeu ProValLeuVal SerAlaArg>

      4060                4080                4100
   *    *    *    *    *    *    *    *    *    *
CGCATCAGCG AGGAAATGGC GAAGCTGAAG CAGAAGACCG AGCAGCTCCA
ArgIleSer GluGluMetAla LysLeuLys GlnLysThr GluGlnLeuGln>

                    4120                4140
      *    *    *    *    *    *    *    *    *    *
GGACCGCATC GCCGGCCAGG CCCGGATCAA CCAGGCCAAG GTGTTGCTGA
 AspArgIle AlaGlyGln AlaArgIleAsn GlnAlaLys ValLeuLeu>

      4160                4180                4200
   *    *    *    *    *    *    *    *    *    *
TGCAGCGCCA TGGCTGGGAC GAGCGCGAGG CGCACCAGCA CCTGTCGCGG
MetGlnArgHis GlyTrpAsp GluArgGlu AlaHisGlnHis LeuSerArg>

                    4220                4240
      *    *    *    *    *    *    *    *    *    *
GAAGCGATGA AGCGGCGCGA GCCGATCCTG AAGATCGCTC AGGAGTTGCT
GluAlaMet LysArgArgGlu ProIleLeu LysIleAla GlnGluLeuLeu>

      4260                4280                4300
   *    *    *    *    *    *    *    *    *    *
GGGAAACGAG CCGTCCGCCT GAGCGATCCG GGCCGACCAG AACAATAACA
 GlyAsnGlu ProSerAla>

>amiS_S/D_sequence
|
|                   4320                4340
|  *    *    *    *    *    *    *    *    *    *
AGAGGGGTAT CGTCATCATG CTGGGACTGG TTCTGCTGTA CGTTGGCGCG
                  Met LeuGlyLeu ValLeuLeuTyr ValGlyAla>

      4360                4380                4400
   *    *    *    *    *    *    *    *    *    *
GTGCTGTTTC TCAATGCCGT CTGGTTGCTG GGCAAGATCA GCGGTCGGGA
ValLeuPhe LeuAsnAlaVal TrpLeuLeu GlyLysIle SerGlyArgGlu>

                    4420                4440
      *    *    *    *    *    *    *    *    *    *
GGTGGCGGTG ATCAACTTCC TGGTCGGCGT GCTGAGCGCC TGCGTCGCGT
 ValAlaVal IleAsnPhe LeuValGlyVal LeuSerAla CysValAla>

      4460                4480                4500
   *    *    *    *    *    *    *    *    *    *
TCTACCTGAT CTTTTCCGCA GCAGCCGGGC AGGGCTCGCT GAAGGCCGGA
PheTyrLeuIle PheSerAla AlaAlaGly GlnGlySerLeu LysAlaGly>

                    4520                4540
      *    *    *    *    *    *    *    *    *    *
GCGCTGACCC TGCTATTCGC TTTTACCTAT CTGTGGGTGG CCGCCAACCA
AlaLeuThr LeuLeuPheAla PheThrTyr LeuTrpVal AlaAlaAsnGln>

      4560                4580                4600
```

```
    *    *    *    *    *    *    *    *    *    *
GTTCCTCGAG GTGGACGGCA AGGGCCTCGG CTGGTTCTGC CTGTTCGTCA
 PheLeuGlu ValAspGly LysGlyLeuGly TrpPheCys LeuPheVal>

                4620                4640
    *    *    *    *    *    *    *    *    *    *
GCCTCACCGC CTGCACCGTG GCGATCGAGT CGTTCGCCGG CGCCAGTGGT
SerLeuThrAla CysThrVal AlaIleGlu SerPheAlaGly AlaSerGly>

      4660                4680                4700
    *    *    *    *    *    *    *    *    *    *
CCGTTCGGCC TGTGGAACGC GGTCAACTGG ACAGTCTGGG CGTTGCTCTG
ProPheGly LeuTrpAsnAla ValAsnTrp ThrValTrp AlaLeuLeuTrp>

                4720                4740
    *    *    *    *    *    *    *    *    *    *
GTTCTGTTTC TTCCTGCTGC TGGGGCTGTC CCGCGGCATC CAGAAGCCGG
 PheCysPhe PheLeuLeu LeuGlyLeuSer ArgGlyIle GlnLysPro>

      4760                4780                4800
    *    *    *    *    *    *    *    *    *    *
TGGCCTACCT GACCCTGGCC AGCGCCATAT TCACCGCCTG GTTGCCCGGC
ValAlaTyrLeu ThrLeuAla SerAlaIle PheThrAlaTrp LeuProGly>

                4820                4840
    *    *    *    *    *    *    *    *    *    *
CTGCTGCTGC TCGGACAGGT GCTCAAGGCA TAGCAGGAAG TCGGAAAGGG
LeuLeuLeu LeuGlyGlnVal LeuLysAla>

      4860                4880                4900
    *    *    *    *    *    *    *    *    *    *
ATGACGGCTT GCCGCCATCC CGTCCCTTCC GAACGCCTAG CCGAGCGGCC

                4920                4940
    *    *    *    *    *    *    *    *    *    *
AGTTGATCAC CACGACGGCG TCGTTGTAGT CGTTGTCGGT GCCGTCTTCA

      4960                4980                5000
    *    *    *    *    *    *    *    *    *    *
GAGCCGACCA GGGCGAAGTT CAGCTCGTTG GTCAGGATTA CCTGTGCCGA

                5020                5040
    *    *    *    *    *    *    *    *    *    *
GACCAGATCC GAGGGGCGGC CGTTGACGCT GACCTGGACC TGTACCTTGC

      5060                5080                5100
    *    *    *    *    *    *    *    *    *    *
CACTGCTGCC GGAGTTGAGC ACCTGGGTGC CGATGACGGC GTTATTGGTG

                5120                5140
    *    *    *    *    *    *    *    *    *    *
CTTTGCCCGC TGAAGGTCGC GGCCGTGCTC GTTGTTGACC AGCACGTTCA

      5160                5180                5200
    *    *    *    *    *    *    *    *    *    *
CCGTCTGGGT TCCGGACGAG TTGGCGAAGG CGGTGACGCC GGAACCTGGT

                5220                5240
    *    *    *    *    *    *    *    *    *    *
TGTTGGCGGG AAGGGTGAAC ACTCCTTGTG GTTGCCATGG TGGTATCTCC

      5260                5280                5300
    *    *    *    *    *    *    *    *    *    *
ACTGAATACC TGGCCCCTTC CTTTTCAGGC AGCCGTCTGG CGCGCGGTAT

                5320                5340
    *    *    *    *    *    *    *    *    *    *
GGCGTGTCGG GAGAAATCCG CAGTCCTTGG CGGCAGGCGA TGCGCAGGCA
```

```
      5360                  5380
    *    *     *    *     *    *     *
GGAAGGACGC ATCGTTCAGC CAATCTACGC CGTCGAC
                                    ArgArg>
```

The EMBO Journal vol.12 no.9 pp.3637–3642, 1993

Antitermination of amidase expression in *Pseudomonas aeruginosa* is controlled by a novel cytoplasmic amide-binding protein

Stuart A.Wilson[1,2], Simon J.Wachira[3], Robert E.Drew[3], David Jones[1] and Laurence H.Pearl[1]

[1]Biomolecular Structure Group and [3]Microbial Genetics Group, Department of Biochemistry and Molecular Biology, University College London, London WC1E 6BT, UK

[2]Correspondence

Communicated by P.H.Clarke

Amide-inducible expression of the aliphatic amidase system of *Pseudomonas aeruginosa* can be reconstituted in *Escherichia coli* with only the amidase structural gene *amiE*, the negative regulator *amiC* and the positive regulator *amiR*, a transcription antitermination factor. Complementation experiments in *E.coli* suggest that negative control of amidase expression by AmiC is mediated by a protein–protein interaction with AmiR. Purified AmiC binds acetamide with a K_D of 3.7 μM in equilibrium dialysis studies, and therefore AmiC appears to be the sensory partner of the AmiC/AmiR pair of regulatory proteins, responding to the presence of amides. Sequence analysis techniques suggest that AmiC is a member of the structural family of periplasmic binding proteins, but has a distinct and novel cytoplasmic role.

Key words: binding proteins/gene regulation/signal transduction

Introduction

Pseudomonas aeruginosa is able to grow on short chain aliphatic amides by virtue of a chromosomally located amidase (EC 3.5.1.4) (Brammar and Clarke, 1964). Amidase activity is inducible by some low molecular weight amides, although the substrate and inducer specificities are quite distinct (Kelly and Clarke, 1962). The amidase genes have been cloned from a constitutive mutant PAC433 (Drew *et al.*, 1980) and from the wild-type strain PAC1 (Wilson and Drew, 1991). The *amiE* gene encoding the amidase enzyme was initially located (Clarke *et al.*, 1981) and sequenced (Brammar *et al.*, 1987) and two regulatory genes, *amiC* and *amiR*, which lie ~2 kb downstream from the *amiE* gene, have also been identified and sequenced (Cousens *et al.*, 1987; Lowe *et al.*, 1989; Wilson and Drew, 1991) (Figure 1).

The *amiR* gene was initially identified as a positive regulator of amidase expression (Farin and Clarke, 1978) and subsequently shown to encode a transcription antitermination factor, thought to function by allowing RNA polymerase to read through a rho-independent terminator identified between the *amiE* promoter and the *amiE* structural gene itself (Drew and Lowe, 1989). The second regulatory protein, AmiC, negatively regulates amidase expression and disruption of the *amiC* open reading frame leads to constitutive amidase expression. The *amiC* and *amiR* open reading frames overlap by 2 bp and are transcribed on the same mRNA (Wilson, 1991; Wilson and Drew, 1991). The AmiC protein has been overexpressed, purified and crystallized (Wilson *et al.*, 1991).

Previous complementation studies have shown that antitermination of *amiE* transcription by AmiR is independent of inducing amides (Cousens *et al.*, 1987; Wilson and Drew, 1991). This suggested that AmiC was the sensory protein and would be responsive to the presence of amides. The precise molecular mechanism of negative control by AmiC has not yet been fully elucidated.

Two well characterized bacterial regulatory systems, the *bgl* operon of *Escherichia coli* and the sac operon of *Bacillus subtilis*, also operate by transcription antitermination (Schnetz and Rak, 1988; Le Coq *et al.*, 1989). In the *bgl* operon the transcription antitermination factor BglG binds to a sequence in the nascent mRNA upstream of and overlapping with the transcription terminator and allows RNA polymerase to read through the terminator (Houman *et al.*, 1990). In both systems the antiterminator activity is negatively regulated by a membrane bound protein (BglF and SacX respectively) and these two proteins each form a part of a phosphoenolpyruvate dependent phosphotransferase system (Amster-Choder *et al.*, 1989; Le Coq *et al.*, 1989). The antitermination factors and negative regulators of both these systems show significant sequence homology. Under non-inducing conditions, the negative regulator BglF phosphorylates BglG, inhibiting its antitermination activity and under inducing conditions, BglF dephosphorylates BglG, activating it and also phosphorylating incoming β-glucosides.

The *P.aeruginosa* amidase operon is clearly distinct from the *bgl* and *sac* operons since the negative regulator AmiC is a soluble cytoplasmic protein, and neither AmiC nor AmiR show homology with their functional counterparts in the *bgl* and *sac* systems.

In this paper we show that a fully functional *P.aeruginosa* amidase induction system can be reconstructed using two and three-plasmid systems in *E.coli*, from the *amiE*, *amiC* and *amiR* genes alone. We also show that the AmiC protein is a specific amide binding protein and that its amide dependent inhibitory action on AmiR operates via a post-transcriptional interaction. We have identified AmiC as a member of the structural protein family comprising the bacterial periplasmic binding proteins, using sequence

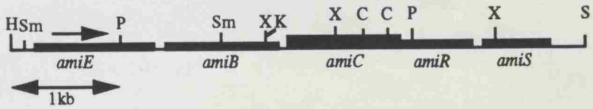


Fig. 1. Organization of the *P.aeruginosa* amidase operon. Restriction sites are indicated as follows: C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; P, *Pvu*II; S, *Sal*I; Sm, *Sma*I; X, *Xho*I. The *amiB* and *amiS* genes are uninvolved in regulation of amidase expression, but may form an active transport system for amides (Drew and Wilson, 1992).

analysis techniques. The cytoplasmic location of AmiC suggests a so far unique role for a member of this protein family.

Results

Reconstitution of an inducible amidase expression system in E.coli

To establish that all of the regulatory genes for inducible amidase expression had been identified, we have reconstructed the regulatory system in *E.coli* (which lacks an aliphatic amidase activity) using a two-vector system. Plasmid pMW21 which expresses the *amiC* and *amiR* genes at high levels, was constructed by insertion of a *Kpn*I–*Sal*I fragment of the amidase operon into the broad host range expression vector pMMB66HE (Figure 2). This vector allows expression of *amiC* and *amiR* from the *tac* promoter under the control of IPTG. A second reporter plasmid (pTM1) was constructed which carries the *amiE* gene in pACYC184 under control of its native promoter and transcription terminator (Figure 2). Amidase activity was measured in *E.coli* JA221 harbouring these plasmids, after growth under inducing, non-inducing and repressing conditions (Table I). pTM1 (containing only the *amiE* gene and regulatory elements) expresses a low level of amidase activity which shows no significant increase in the presence of inducing (lactamide) or repressing (butyramide) amides. *Escherichia coli* simultaneously harbouring pTM1 and pMW21 (containing the *amiC* and *amiR* genes) shows a low level of amide-inducible amidase activity in the absence of IPTG presumably due to leaky repression of the *tac* promoter, and high amide-inducible levels of amidase expression in the presence of IPTG. Thus we have reconstituted inducible amidase expression in *E.coli* using the *amiE*, *amiC* and *amiR* genes alone. Inducible amidase expression can also be reconstructed in *P.aeruginosa* using pMW21 in strain PAC327 (Brown, 1969). PAC327 is *amiRC*$^-$ and amidase negative. In the presence of IPTG, PAC327 harbouring pMW21 shows inducible amidase expression (Table I).

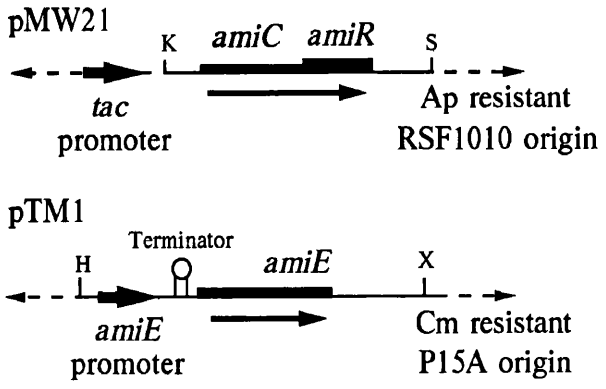


Fig. 2. Two-plasmid system. Structure of *amiC*+*amiR* expression plasmid pMW21, and *amiE* reporter plasmid pTM1. Restriction sites are as Figure 1. Selective antibiotics: Ap, ampicillin; Cm, chloramphenicol. *P.aeruginosa* derived sequences are shown as full lines, plasmid sequence as broken arrowed lines. The thin arrows indicate the direction of transcription.

Three-plasmid complementation in E.coli

Several possibilities exist for the mechanism of negative control of amidase expression by *amiC*. First, AmiC might be a conventional DNA binding repressor of *amiE* or *amiR* transcription; however, the AmiC sequence contains none of the motifs usually associated with sequence specific DNA binding proteins, nor does purified AmiC bind to DNA (Wilson, 1991). Alternatively, AmiC may interact directly with AmiR and modify its activity either by formation of a stoichiometric complex, or by an enzymatic modification of AmiR such as phosphorylation, as seen in the *bgl* and *sac* operons (s.v.).

To distinguish between these two modes of AmiC action, a three-plasmid complementation system was constructed in *E.coli* with *amiE*, *amiC* and *amiR* each carried on separate compatible plasmids. The *amiE* gene was carried on pACYC184 (pTM1) described above. The *amiC* gene (*Kpn*I–*Pvu*II) was carried on pMMB66HE (pSW41) and expressed from the vector *tac* promoter (Wilson and Drew, 1991) and *amiR* (*Xho*I–*Xho*I) was carried on pBGS19$^-$ (pTM2) with *amiR* expressed from the vector *lac* promoter (Figure 3).

Amidase activities expressed by *E.coli* with pTM1 (*amiE* alone), pTM1+pTM2 (*amiE*+*amiR*) and pTM1, pTM2 and pSW41 (*amiE*+*amiR*+*amiC*) are shown in Table II. As in the two-plasmid experiments, pTM1 alone expressed low levels of amidase activity. With the addition of plasmid pTM2, high levels of amidase activity were seen both in the presence and absence of amides. With the addition of the third plasmid pSW41, which produces very high constitutive levels of AmiC, there is almost complete shut down of amidase expression. Since AmiR and AmiC are transcribed from vector promoters, it is most unlikely that AmiC can be affecting the expression of the *amiR* gene by repression of its transcription, but is rather acting post-transcriptionally by interaction with the AmiR protein itself.

Table I. Amidase activity in *E.coli* and *P.aeruginosa* carrying recombinant plasmids

Strains/ plasmids	IPTG	Glucose	Glucose lactamide	Succinate	Succinate/ lactamide	Succinate/ butyramide
E.coli[a]						
pTM1	–	0.5	0.5	0.4	0.7	1.3
pTM1, pMW21	–	0.8	4.9	0.8	6.4	0.3
pTM1, pMW21	+	3.0	22.3	3.9	40.4	0.5
P.aeruginosa[b]						
–	–	–	–	0.4	0.3	–
pMW21	+	–	–	0.7	3.0	–

[a]*E.coli* strain JA221.
[b]*P.aeruginosa* strain PAC327 (*amiCR*$^-$).

In this three-plasmid complementation system, negative regulation by AmiC cannot be relieved by addition of inducing amides. Although expression of *amiR* from pTM2 produces a significant *in vivo* effect on amidase expression, it does not produce high levels of AmiR, and no new band is visible on SDS–PAGE gels (Wilson, 1991), whereas the AmiC expression vector pSW41 produces a major new band on SDS–PAGE gels. It is likely that the unrelievable inhibition of amidase expression in the three-plasmid system is due to the presence in the cell of saturating amounts of AmiC with respect to AmiR.

Amide binding by AmiC

Complementation systems containing *amiE* and *amiR* alone do not respond to amide inducers, whereas the addition of *amiC* confers amide inducibility, implicating the AmiC protein as the amide 'sensor', and suggesting that AmiC should bind amides. Equilibrium dialysis experiments were performed using purified AmiC (Wilson *et al.*, 1991) and [^{14}C]acetamide. A Scatchard plot (Figure 4) of the results from these experiments gives a value for K_D of 3.7 μM and a protein:ligand stoichiometry of 2:1. As the AmiC protein has been found to migrate as a dimer in gel filtration studies (Wilson *et al.*, 1991) this ratio suggests that one acetamide molecule binds to an AmiC dimer. Other amides, lactamide and propionamide which have been found to be strong inducers of amidase expression compete with acetamide for binding to AmiC, as does butyramide, an inhibitor of induction (Kelly and Clarke, 1962). In competition binding experiments, propionamide shows a K_D of 3.1 μM whereas lactamide and butyramide bind ~100-fold less tightly.

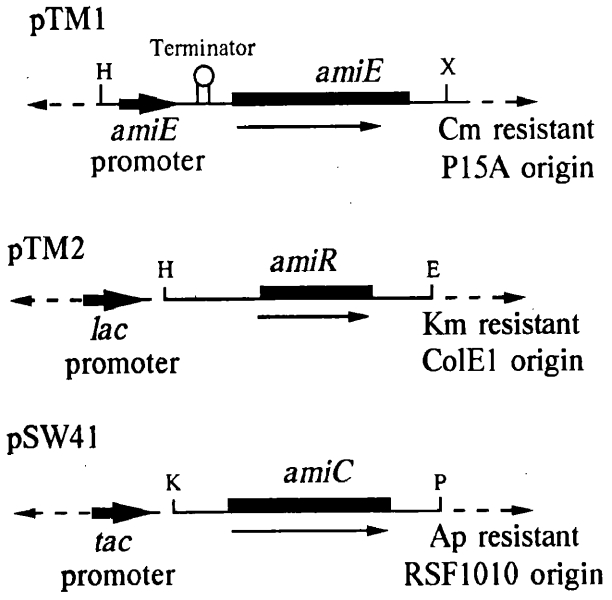


Fig. 3. Three-plasmid system. Structure of *amiE* reporter plasmid pTM1 (see Figure 2), *amiR* expression plasmid pTM2 and *amiC* expression plasmid pSW41. Restriction sites are indicated as in Figure 1; selective antibiotics as Figure 2 plus Km, kanamycin.

Sequence analysis of the AmiC protein

The derived amino acid sequence of the AmiC protein (Wilson and Drew, 1991) was compared with 26 706 sequences in the SwissProt sequence database, using the FASTA program (Pearson and Lipman, 1988). A relatively weak match (19.2% identity over 339 amino acids) was found with the sequence of the BraC branched-chain amino acid binding protein of *P.aeruginosa* (Hoshino and Kose, 1989). Figure 5 shows an alignment of the AmiC sequence with the BraC protein and several functionally related proteins. Although the individual identity between AmiC and the branched-chain amino acid binding proteins is low, AmiC shows a general overall similarity to the family of proteins over its entire length, with conformationally important residues such as glycine and proline, being frequently conserved. The similarity of AmiC to these proteins has been further tested by secondary structure prediction. A consensus of seven standard prediction algorithms (Eliopoulos *et al.*, 1982) was used to predict the secondary structure of AmiC and of the leucine-isoleucine-valine binding protein of *E.coli* (LivJ). These predictions were compared with the secondary structure actually observed in the crystal structure of LivJ (Sacks *et al.*, 1989) (Figure 4). While the predicted secondary structure of LivJ does not precisely match the observed structure, most secondary structural elements are correctly predicted in position and type, if not in length. Most interestingly, the predicted secondary structure of AmiC corresponds equally well to the observed LivJ secondary structure. As a final clue to the structure of AmiC, we have submitted the amino acid sequence to a 'threading' analysis (Jones *et al.*, 1992). This procedure measures the 'fit' of the AmiC sequence onto 102 three-dimensional folds from known X-ray structures, using a combination of solvent accessibility and pair-distance parameters. The distribution of threading scores for AmiC is shown in Figure 6. The match with most folds gives a roughly Gaussian distribution; however, the scores for two related folds, the *E.coli* LivJ protein (Sacks *et al.*, 1989) and the *Salmonella typhimurium* galactose binding protein (Mowbray and Petsko, 1983), are significantly better and outlie the overall distribution. Taken together these analyses suggest that the AmiC sequence is capable of adopting a three-dimensional structure which is very similar to LivJ and the *S.typhimurium* galactose binding protein, both members of a large and well characterized structural family of binding proteins (Quicho, 1991) which function as periplasmic receptors for small molecules in many bacteria.

Discussion

Reconstruction of the *P.aeruginosa* amidase regulatory system in *E.coli* confirms that only two regulatory proteins

Table II. Amidase activity in *E.coli* JA221 carrying *amiE*, *amiC* and *amiR* recombinant plasmids

Plasmids	Glucose	Glucose lactamide	Succinate	Succinate/ lactamide	Succinate/ butyramide
pTM1	0.5	0.5	0.4	0.7	1.3
pTM1, pTM2	30.8	31.8	31.0	33.6	39.1
pTM1, pTM2, pSW41	0.9	1.0	0.9	0.9	0.9

are required for amide-inducible expression, and identifies the AmiC protein as the component responsible for recognition of inducer molecules. This role for AmiC has been confirmed by demonstrating binding of acetamide and other inducers/induction inhibitors by purified AmiC *in vitro*. In the three-plasmid complementation system AmiC is able to block AmiR activity even though AmiR was expressed from a foreign promoter. We conclude therefore that AmiC does not function as a classical DNA binding repressor, but inhibits the action of AmiR post-transcriptionally (Wilson and Drew, 1991). Analysis of the AmiC sequence shows it to be a member of a structural family usually found as periplasmic receptors in a variety of bacterial transport and chemotaxis systems. AmiC lacks the periplasmic localization signal found at the N-terminus of these proteins, consistent with its cytoplasmic localization. The classic periplasmic binding proteins display a characteristic ligand dependent protein−protein interaction with a specific membrane bound protein complex (Furlong, 1987). As AmiC has a structural homology and displays a ligand binding function of similar affinity to these proteins, it too might act via a protein−protein interaction but with the cytoplasmic antitermination factor AmiR, and this interaction may be responsible for inhibition of the antitermination activity of AmiR.

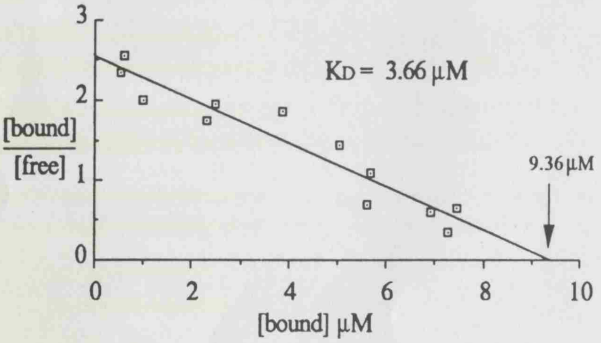


Fig. 4. Scatchard plot of [^{14}C]acetamide binding to AmiC. See text for experimental details. The binding constant was determined by linear regression of the experimental points. The arrow indicates the theoretical maximum acetamide binding.

We have attempted to determine whether AmiC covalently modifies AmiR by phosphorylation or by proteolysis, but no such activity has been detected (data not shown), suggesting that AmiC inhibits AmiR by direct binding,

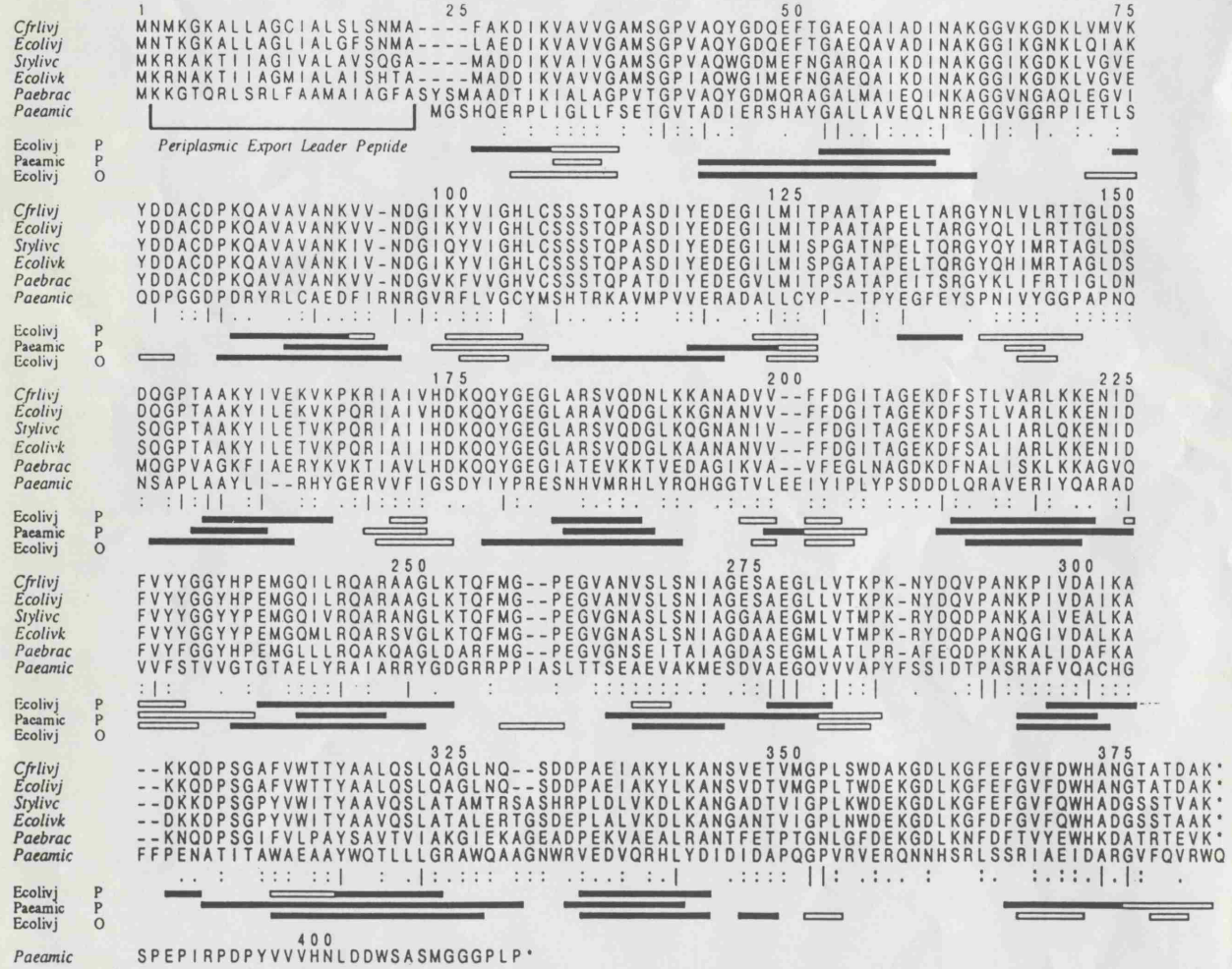


Fig. 5. Sequence alignment of AmiC with members of the periplasmic binding protein structural family. Sequences: Cfrlivj, *C.freundi* LivJ (Garvin and Hardies, 1991); Ecolivj, *E.coli* LivJ (Sacks *et al.*, 1989). Stylivc, *S.typhimurium* LivC (Ohnishi *et al.*, 1990); Ecolivk, *E.coli* LivK (Adams *et al.*, 1990); Paebrac, *P.aeruginosa* BraC (Hoshino and Kose, 1989); Paeamic, *P.aeruginosa* AmiC (Wilson and Drew, 1991). Sequence homologies are indicated below the sequences by |, totally conserved; :, strongly conserved; ., weakly conserved. Secondary structure predictions/observations are Ecolivj P, *E.coli* LivJ predicted secondary structure; Paeamic P, *P.aeruginosa* AmiC predicted secondary structure; Ecolivj O, *E.coli* LivJ observed secondary structure. Solid bars, α-helix; open bars, β-sheet. The N-terminal periplasmic export signal found in the periplasmic binding protein sequences is indicated.

blocking access of AmiR to the *amiE* RNA leader sequence. Previous genetic studies suggest that there is a stoichiometry between the levels of AmiC and AmiR. Overexpression of AmiR in a wild-type *P.aeruginosa* background leads to constitutive amidase expression (Wilson and Drew, 1991). In this situation there is only a single copy of *amiC* on the chromosome expressing relatively low levels of AmiC compared with the overexpressed plasmid-borne *amiR* gene. Similarly, if AmiC is overexpressed with respect to AmiR in *P.aeruginosa*, the system becomes uninducible and all amidase expression ceases. Coordinate overexpression of AmiC and AmiR in *E.coli* and in *P.aeruginosa* gives normal inducible amidase expression. Taken together these results suggest that the mechanism of AmiC inhibition of AmiR operates via a stoichiometric rather than a catalytic relationship between the two proteins, probably involving formation of an AmiC–AmiR complex.

On binding their specific ligand, some classical periplasmic binding proteins display a conformational change which enables them to bind to the membrane protein complex, and which can be detected spectroscopically (e.g. Zukin, 1979). We have attempted to observe a similar change in the conformation of AmiC on binding of acetamide, but no significant signal has been observed using fluorescence or near-UV circular dichroism (data not shown), however, both these techniques are only sensitive to relatively large changes in the environments of aromatic groups, and the lack of a significant signal does not indicate the absence of some more subtle structural change on binding of amides. Addition of amides does not alter the dimerization state of AmiC, which runs as a dimer of molecular weight 86 kDa, even in the presence of >100 mM acetamide. As AmiC is the receptor of amides in the amidase system, and its ability to inhibit the action of AmiR depends on the presence or absence of amides, it seems likely that binding of acetamide to AmiC will have some effect on its structure, if only in that part that interacts with AmiR.

A full understanding of the molecular mechanism of this unusual signal transduction/gene regulation system must await the successful purification of AmiR for studies of the AmiC–AmiR complex in the presence of inducing/induction inhibiting amides.

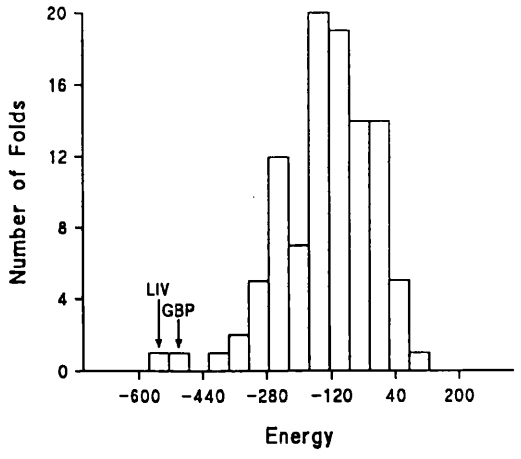


Fig. 6. Threading analysis of the AmiC sequence. Histogram of pseudo-energies for the AmiC sequence threaded on to 102 known protein folds. The most stable threadings on to the folds of *E.coli* LivJ and *S.typhimurium* galactose binding protein are indicated. The score for the LivJ threading is >100 kcal/mol more stable than that for the galactose binding protein and suggests a strong similarity between the AmiC and LivJ three-dimensional structures.

Materials and methods

Construction of plasmids

All plasmid purifications, transformations and cloning was carried out as described previously (Wilson and Drew, 1991). Plasmid pTM1 was constructed by insertion of a 2.3 kb *Hin*dIII–*Xho*I *amiE* fragment from pAS20 (Wilson and Drew, 1991) into the *Hin*dIII and *Sal*I sites of pACYC184 which has a p15A origin of replication and confers chloramphenicol resistance. Plasmid pTM2 carries a 1.5 kb *Hin*dIII–*Eco*RI fragment from pSW24 (Wilson, 1991) subcloned into pBGS19$^-$, which has a ColE1 origin of replication and confers resistance to kanamycin. The orientation of the *amiR* gene allows expression from the *lac* promoter in the plasmid.

Assay of amidase activity

Amidase activity in intact cells was measured by the transferase assay (Brammar and Clarke, 1964) with acetamide as the substrate. Activity levels presented in this article are the mean values of duplicate assays carried out on at least three separate occasions. One unit represents 1 μmol of acetohydroxamate formed per min per mg of bacteria.

Purification of AmiC

AmiC was purified as described previously (Wilson *et al.*, 1991) and protein concentration determined by the Bradford assay (Bradford, 1976).

Equilibrium dialysis with acetamide

[^{14}C]acetamide was synthesized by reaction of [^{14}C]ethyl acetate (Amersham International) with an excess of concentrated ammonia in ethanol and cold ethyl acetate. The product from an identical cold synthesis was characterized by assay with *P.aeruginosa* amidase enzyme and found to be >99.5% pure. In all equilibrium dialysis experiments a constant concentration of AmiC was used (18 μM) and the concentration of [^{14}C]acetamide was varied. In competition dialysis experiments, constant concentrations of AmiC (18 μM) and [^{14}C]acetamide (15 μM) were used, and the concentrations of competing amides were varied. To determine accurately the concentration of the radioactive acetamide, a titration experiment was performed in which non-radioactively labelled acetamide of known concentration was used as competitor of AmiC binding. The concentration of cold acetamide which gave a 50% reduction in the amount of bound acetamide was taken as an equimolar concentration to the labelled acetamide (data not shown). Dialysis experiments were performed with 50 μl volumes on either side of a 14 kDa cut-off dialysis membrane in a Teflon equilibrium dialysis module (Hoeffer Scientific) at 16°C overnight. Duplicate 20 μl samples were taken from both sides of the membrane and liquid scintillation counted.

Sequence alignment, secondary structure prediction and fold analysis

Sequence alignments were performed using the multiple alignment algorithm of Feng and Doolittle (1987) implemented in the program PileUp (University of Wisconsin) with small manual adjustment of the final alignments. Secondary structure prediction was performed using the consensus prediction method of Eliopoulos *et al.* (1982). Optimal fold threading of the AmiC sequence used the method of Jones *et al.* (1992). Briefly, the AmiC sequence was threaded on to 102 protein folds, and the pseudo-energy of each alignment calculated as the weighted sum of pairwise and solvation pseudo-energy terms.

Acknowledgements

We are grateful to Dr George Nicolaou for advice and help in chemical synthesis and to Tahir Malik for technical assistance. We thank Drs Andrew Hemmings, Bernard O'Hara and Christine Orengo, and Professor Patricia Clarke for useful discussion. This work was funded by a Wellcome Trust Grant to L.H.P. and R.E.D. S.J.W. was supported by a British Council Studentship, and D.J. was supported by an SERC Studentship.

References

Adams,M.D., Wagner,L.M., Graddis,T.J., Landick,R., Antonucci,T.K., Gibson,A.L. and Oxender,D.L. (1990) *J. Biol. Chem.*, **265**, 11436–11443.

Amster-Choder,O., Houman,F. and Wright,A. (1989) *Cell*, **58**, 847–855.
Bradford,M.M. (1976) *Anal. Biochem.*, **22**, 248.
Brammar,W.J. and Clarke,P.H. (1964) *J. Gen. Microbiol.*, **37**, 307–319.
Brammar,W.J., Charles,I.G., Matfield,M., Cheng-Pin.L., Drew,R.E. and Clarke,P.H. (1987) *FEBS Lett.*, **215**, 291–294.
Brown.P.R. (1969) Ph.D. thesis, University of London.
Clarke,P.H., Drew,R.E., Turbeville,C., Brammar,W.J., Ambler,R.P. and Auffret,A.D. (1981) *Biosci. Rep.*, **1**, 299–307.
Cousens,D.J., Clarke,P.H. and Drew,R.E. (1987) *J. Gen. Microbiol.*, **133**, 2041–2052.
Drew,R.E. and Lowe,N. (1989) *J. Gen. Microbiol.*, **135**, 817–823.
Drew,R.E. and Wilson,S.A. (1992) In Galli,E., Silver,S. and Witholt,E. (eds), *Pseudomonas: Molecular Biology and Biotechnology*. American Society of Microbiology, Washington DC, pp. 207–213.
Drew,R.E. Clarke,P.H. and Brammar,W.J. (1980) *Mol. Gen. Genet.*, **177**, 311–320.
Eliopoulos,E., Geddes,A.J., Brett,M., Pappin,D.J.C. and Findlay,J.B.C. (1982) *Int. J. Biol. Macromol.*, **4**, 263–268.
Farin,F. and Clarke,P.H. (1978) *J. Bacteriol.*, **135**, 379–392.
Feng,D.-F. and Doolittle,R.D. (1987) *J. Mol. Evol.*, **25**, 351–360.
Furlong,C.E. (1987) In Niedhart,C.F. *et al.* (eds), *Escherichia coli and Salmonella typhimurium—Cellular and Molecular Biology*. American Society for Microbiology, Washington DC, pp. 768–796.
Garvin,D.L. and Hardies,S.C. (1991) SwissProt Database entry P25399.
Hoshino,T. and Kose,K. (1989) *J. Bacteriol.*, **171**, 6300–6306.
Houman,F., Diaz-Torres,M.R. and Wright,A. (1990) *Cell*, **62**, 1153–1163.
Jones,D.T., Taylor,W.R. and Thornton,J.M. (1992) *Nature*, **358**, 86–89.
Kelly,M. and Clarke,P.H. (1962) *J. Gen. Microbiol.*, **27**, 305–316.
Le Coq,D., Crutz,A.M., Richter,R., Aymerich,S., Gonzy-Treboul,G., Zagorec,M., Rain-Guion,M.C. and Steinmetz,M. (1989) In Butler,L.O., Harwood,C. and Moseley,B.E.B. (eds), *Genetic Transformation and Expression*. Intercept Ltd, Andover, UK, pp. 447–456.
Lowe,N., Rice,P.M. and Drew,R.E. (1989) *FEBS Lett.*, **246**, 39–43.
Mowbray,S.L. and Petsko,G.A. (1983) *J. Biol. Chem.*, **258**, 7991–7997.
Ohnishi,K., Nakazima,A., Matsubara,K. and Kiritani,K. (1990) *J. Biochem.*, **107**, 202–208.
Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 2444–2448.
Quicho,F.A. (1991) *Curr. Opin. Struct. Biol.*, **1**, 922–933.
Sacks,J., Saper,M.A. and Quicho,F.A. (1989) *J. Mol. Biol.*, **206**, 171–191.
Schnetz,K. and Rak,B. (1988) *EMBO J.*, **7**, 3271–3277.
Wilson,S.A. (1991) Ph.D. thesis. University of London.
Wilson,S.A. and Drew,R.E. (1991) *J. Bacteriol.*, **173**, 4914–4921.
Wilson,S.A., Chayen,N.E., Hemmings,A.M., Drew,R.E. and Pearl,L.H. (1991) *J. Mol. Biol.*, **222**, 869–871.
Zukin,R.S. (1979) *Biochemistry*, **18**, 2139–2145.

Received on April 19, 1993