A STUDY OF GENE EXPRESSION IN *PSEUDOMONAS*

By

Rebecca Louise Allen

Submitted for the Degree of Doctor of Philosophy

University of Leicester

April 1989

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

For My Parents

STATEMENT

The accompanying thesis submitted for the degree of Doctor of Philosophy entitled "A study of gene expression in *Pseudomonas*" is based on work conducted by the author in the Department of Microbiology of the University of Leicester mainly during the period between October 1985 and September 1988.

All the work recorded in this thesis is original, unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other university.

Signed.....R.L.Allen..... Date..14.4.89..

ABSTRACT

An inherent problem in the study of the genetics of the interesting and potentially commercially useful properties of the pseudomonads is that of gene expression, since many of the genes encoding these properties are not well expressed in an *E.coli* background. The evidence available at the present time indicates that some *Pseudomonas* genes may possess different promoter sequences not recognised by *E.coli* RNA polymerase.

An *in vitro* coupled transcription/translation system based on *P.putida* has been developed. A comparison of *E.coli* and broad host range plasmid DNA in this and the equivalent *E.coli* system showed that although cloned *E.coli* and vector polypeptides were synthesised in both systems, there was a difference in the polypeptide products directed by broad host range plasmid DNA in the two systems. In particular RSF1010 directed the synthesis of a 73kD polypeptide uniquely in the *P.putida* system. This was shown to be a polypeptide involved in mobilisation of the plasmid.

A broad host range promoter-probe vector based on RSF1010 was constructed and used for the shotgun cloning of *P.putida* promoters. A small subset of fragments which were active as promoters in *P.putida* but exhibited much lower activity in *E.coli* were isolated, sequenced and analysed with respect to concensus *E.coli* and nitrogen-regulated promoter sequences. These isolated DNA fragments may represent promoters which have sequences specifically recognised by *Pseudomonas* RNA polymerase. An analysis of published *Pseudomonas* chromosomally-encoded promoters revealed putative *Pseudomonas*-specific concensus regions.

ABBREVIATIONS

Ap	Ampicillin
bp	base pair
BSA	Bovine serum albumin
Cb	Carbenicillin
Cm	Chloramphenicol
C23O	Catechol 2,3-Oxygenase
DEPC	Diethylpyrocarbonate
DTT	Dithiothreitol
EDTA	Diaminoethanetetra-acetic acid
FA	Folinic acid
IPTG	Isopropyl-β-D-thiogalactosylpyranoside
kb	kilobase
kD	kilodalton
Km	Kanamycin
LMM	Low Molecular Weight Mix
MOPS	3-[N-mophilino] propanesulfonic acid
ORF	open reading frame
PEG	Polyethyleneglycol
PEP	Phosphoenolpyruvate
SDS-PAGE	Sodium dodecylsulphate-polyacrylamide gel electrophoresis
Sm	Streptomycin
Sp	Spectinomycin
Su	Sulphonamide
TEMED	N,N,N,N'-Tetramethylethylenediamine
Tp	Trimethoprim
X-Gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

CONTENTS

1.1 THE GENUS OF *PSEUDOMONAS*

The genus of *Pseudomonas* represents a markedly heterogeneous group of species which have colonised a variety of habitats including the soil, fresh and salt water and some exist in association with plants and animals. The name *Pseudomonas* means "false unity" and the many members of the genus were so grouped because of their apparent lack of similarity to other genera. In an attempt to subject this diverse group of species to more stringent taxonomic analysis Stanier *et al* in 1966 investigated the nutritional properties of 267 aerobic pseudomonads. They put forward the following definition of the aerobic pseudomonads, thus excluding some bacteria which had previously belonged to the group:

> "Unicellular rods, with the long axis straight or curved, but not helical. Motile by means of one or more flagella. Gram-negative. Do not form spores, stalks or sheaths. The energy-yielding metabolism is respiratory, never fermentative or photosynthetic. All can use molecular oxygen as a terminal oxidant, and some can use denitrification as an alternative, anaerobic respiratory mechanism. All are chemo-organotrophs; some are facultative chemolithotrophs that use H_2 as an energy source."

They also noted that the GC ratio of the group varies from 58% to 69%. This great variation in base composition within a bacterial group, in addition to the variety of ecological niches that the species exploits, illustrates the great heterogeneity and DNA sequence divergence of the genus.

Stanier *et al* (1966) showed that, because of the number of compounds able to be used as carbon and energy sources, each species and groups of species of *Pseudomonas* could be classified according to their nutritional phenotype. Table 1 illustrates the extensive nutritional versatility of one *Pseudomonas putida* biotype. The similarites in nutritional phenotype between the species or groups of species tend to suggest that they have

TABLE 1

SUBSTRATES UTILISED BY *PSEUDOMONAS PUTIDA* BIOTYPE B

Carbohydrates and sugar derivatives
- D-Glucose
- D-Fructose
- Gluconate
- 2-Ketogluconate
- Saccharate
- Mucate

Fatty Acids
- Acetate
- Propionate
- Butyrate
- Valerate
- Isovalerate
- Caproate
- Heptanoate
- Caprylate
- Pelargonate
- Caprate

Dicarboxylic Acids
- Malonate
- Succinate
- Fumarate
- Glutarate

Hydroxy acids
- L-Malate
- DL-B-Hydroxybutyrate
- DL-Lactate
- DL-Glycerate

Miscellaneous Organic Acids
- Citrate
- α-Ketoglutarate
- Pyruvate
- Aconitate

Polyalcohols and Glycols
- Glycerol

Alcohols
- Ethanol
- n-Propanol
- n-Butanol
- Isobutanol

Non-nitrogenous aromatic and other cyclic compounds
- Benzoate
- ρ-Hydroxybenzoate
- Phenylacetate
- Quinate
- Testosterone

Aliphatic Amino Acids
- Glycine
- L-α-Alanine
- D-α-Alanine
- B-Alanine
- L-Serine
- L-Leucine
- L-Isoleucine
- L-Valine
- L-Aspartate
- L-Glutamate
- L-Lysine
- DL-Arginine
- DL-Ornithine
- γ-Aminobutyrate
- δ-Aminovalerate

Amino Acids and related Compounds containing a ring Structure
- L-Histidine
- L-Proline
- L-Tyrosine
- Phenylalanine
- L-Tryptophan
- L-Kynurenine
- Anthranilate

Amines
- Benzylamine
- Putrescine
- Spermine
- Histamine
- Tryptamine
- Butylamine
- α-Amylamine

Miscellaneous nitrogenous compounds
- Betaine
- Sarcosine
- Trigonelline

similar biochemical pathways, however this is not always the case; some species able to grow on a certain compound may possess different enzymes such that the metabolic pathways diverge. A knowledge of the biochemistry of a pathway is therefore required to determine the relatedness of two species able to metabolise one compound. In addition to these analyses Palleroni (1975) investigated classification of *Pseudomonas* at the genetic level by *in vitro* DNA-DNA and rRNA-DNA hybridisation. The results suggest the grouping of the species into five "RNA homology groups" (Table 2). This is the first complex group of bacteria whose internal subdivision has been successfully achieved by ribosomal RNA homology studies.

1.2 *PSEUDOMONAS* GENETICS

Compared with our knowledge of *E.coli* genetics, the study of genetics of *Pseudomonas* is in its infancy. Work on *Pseudomonas* has concentrated on *P.aeruginosa* and *P.putida*, both of which are members of RNA homology Group I and thus, is unlikely to be representative of the genus as a whole. The chromosomes of *P.aeruginosa* and *P.putida* have been mapped to a limited degree and significant differences have been found on comparison with the *E.coli* and *Salmonella typhimurium* chromosomes. In *E.coli* and *S.typhimurium* the genes for biosynthetic pathways tend to be contiguously arranged, but in *Pseudomonas* there tends to be little contiguous arrangement of genes encoding enzymes involved in biosynthetic pathways. For example, the tryptophan biosynthesis genes of *P.putida* and *P.aeruginosa* (Chakrabarty *et al* 1968, Gunsalus *et al* 1968, Shinomiya *et al* 1983) and the arginine biosynthesis genes *of P.aeruginosa* (Haas *et al* 1977) show a dispersed arrangement on the chromosome. The lack of clustering of genes for enzymes of biosynthetic pathways is not unique to *Pseudomonas*; it occurs in other bacteria and also in eukaryotes. In fact the clustering of these genes may indeed be a peculiarity of enteric bacteria. There is a tendency towards supra-operonic clustering in *Pseudomonas* (Holloway and Morgan 1986). Supra-

TABLE 2

RNA HOMOLOGY GROUPS OF SOME *PSEUDOMONAS* SPECIES

Group I	*P.aeruginosa*
	P.fluorescens
	P.putida
	P.syringae
	P.stutzeri
	P.mendocina
	P.alcaligenes
	P.pseudoalcaligenes
Group II	*P.cepacia*
	P.marginata
	P.caryophylli
	P.pseudomallei
	P.mallei
	P.solanacearum
	P.pickettii
Group III	*P.acidovorans*
	P.teststeroni
	P.facilis
	P.delafieldii
	P.saccharophila
Group IV	*P.diminuta*
	P.vesicularis
Group V	*P.maltophilia*

operonic clustering of catabolic functions occurs almost exclusively in *Pseudomonas* (Leidigh and Wheelis 1973). The plasmid-encoded catabolic genes for the degradative pathways tend to be clustered. For example, the genes for the metabolism of toluene/xylene of the TOL plasmid from *P.putida* are clustered into 2 operons. In the region from 0-14 minutes on the chromosomal map of *P.putida* lie all the genes for the dissimilation of catechol, mandelate, histidine, protocatechuate, phenylacetate, phenylalanine and quinate. *Pseudomonas* also displays contiguous arrangement of functionally related genes, for example, the structural gene for amidase, *ami*E, and its regulatory gene, *ami*R (Brammar *et al* 1967).

The *Pseudomonas* chromosome encodes a vast amount of genetic information for the utilisation of many different substrates each of which will only occasionally be available as nutrient sources in the soil environment. This genetic load could not be withstood if all the genes were constitutively expressed nor if all the pathways for catabolism were independent. It seems that such genes tend to be positively regulated and that there is much convergence of biochemical pathways, particularly on the product catechol which can then be metabolised by two different routes, the *ortho*- or *meta*-cleavage pathways.

1.2.1 Extrachromosomal Elements in *Pseudomonas*

There are many extrachromosomal elements in *Pseudomonas*. These include the degradative plasmids (see 1.3.3) particularly in *P.putida*. Also R plasmids (resistance factors), FP plasmids (sex factors) and plasmids conferring resistance to heavy metal ions are frequently found in *P.aeruginosa*. The plasmids FP2 and FP5 have been used to mobilise the *P.aeruginosa* chromosome (Haas *et al* 1977). R68.45 is an IncP-1 broad host range plasmid which was isolated from *P.aeruginosa* (Haas and Holloway 1978). It has the ability to mobilise the chromosome at high frequency for any marker, from many sites and in both orientations. This has made it

useful for the creation of circular linkage maps in various species, which was difficult when certain of the FP plasmids were used since they had the same site of origin for chromosome transfer. Very few chromosome-mobilising plasmids have been identified in *Pseudomonas* species other than *P. aeruginosa.* Among the R plasmids, those belonging in particular to incompatibility group IncP-1 exhibit an extensive host range and have been important for the construction of broad host range cloning vectors (see 1.4.1.1). Lysogeny is common in *P. aeruginosa* and phages have been extensively used for the mapping of the chromosome by generalised transduction (Holloway *et al* 1979); such phages are uncommon in other pseudomonads.

1.3 BIOLOGICAL PROPERTIES OF *PSEUDOMONAS*

The ability of the genus to metabolise such an array of chemical compounds is the major distinguishing feature between *Pseudomonas* and enterobacteria and it is for this reason that an enormous amount of interest and work has, in the last 20 years, been focused on this group of organisms, and indeed *P. aeruginosa* is one of the most well characterised organisms after *E. coli.* The major areas of work in the field of *Pseudomonas* genetics are a study of the pathogenicity of one of the few human pathogens in the group, *P. aeruginosa,* many of the plant-bacteria interactions and the xenobiotic degradation by some of the species.

1.3.1 Human Pathogenicity of *P. aeruginosa*

Most members of the genus *Pseudomonas* do not have the potential to be pathogenic for humans or animals, but one species in particular, *P. aeruginosa,* has adapted to exploit some unusual niches. It is highly ubiquitous both in nature and in almost any moist environments provided by humans. This is of particular importance in hospitals where *P. aeruginosa* can be isolated from respiratory equipment, floors, baths and even soap.

Thus the potential of spread to susceptible persons is great. The advent of the jacuzzi has created a further habitat and has given rise to cases of an irritating skin complaint called "jacuzzi rash"! *P.aeruginosa* does not infect healthy human beings, but it can be an important cause of mortality in hospital patients with impaired host defence. It acts opportunistically causing local infections in wounds and is particularly important in burns patients, where it can lead to septicaemia and death. Another class of people susceptible to serious and debilitating infection with *P.aeruginosa* are those who suffer from cystic fibrosis (CF). Such people produce viscous mucous in the lungs blocking the bronchioles and providing a habitat for the opportunistic pathogen (Deretic *et al* 1987a). Initially *Staphylococcus aureus* and *Haemophilus influenzae* will colonise the area, but these can be effectively eliminated by the use of antibiotics. *P.aeruginosa* replaces these organisms and takes on a mucoid phenotype (this mucoidy is only encountered when the organism invades the lungs of cystic fibrosis sufferers) due to the production of the exopolysaccharide alginate. Mucoid strains of *P.aeruginosa* form microscopically visible microcolonies embedded in a gelatinous material which adhere to the pulmonary mucosa. It is thought that this structure plays an important role in explaining the higher virulence of mucoid *P.aeruginosa* and the high resistance of them to the strongest anti-pseudomonal antibiotics and also their inability to be taken up by phagocytes (Govan and Harris 1986). The organisms and probably the frustrated immune system cause extensive tissue damage and the alginate combined with CF lung mucous leads to increased bronchopneumonia often followed by mortality. The genetic control of mucoidy or alginate production in *P.aeruginosa* is being studied and recent work (Deretic *et al* 1987b) has shown that one of the genes involved, *algD*, coding for GDPmannose dehydrogenase is subject to strong transcriptional activation in mucoid *P.aeruginosa* strains. The *algD* gene is under positive control by the *algR* gene (Deretic *et al* 1987c).

P.aeruginosa, unlike the non-human pathogenic *P.putida*, possesses fimbriae which promote colonisation of host surfaces. Potential virulence factors of *P.aeruginosa* which are being actively studied are exotoxin A, elastase, alkaline protease, exoenzyme S and phospholipase C (a haemolysin).

1.3.2 *Pseudomonas*-Plant Interactions

A large number of *Pseudomonas* species exhibit a symbiotic or parasitic relationship with plants. The phytopathogenic species belong to the pseudomallei-cepacia RNA homology group (Group II) and the fluorescent species of Group I. Many species of *Pseudomonas* confer beneficial properties on plants. For example, *P.putida* produces yellow-green fluorescent siderophores under iron-limiting conditions. These compounds are high-affinity ironIII-chelating substances and in the soil environment bacteria excreting such compounds deprive other microorganisms of iron. One such siderophore-secreting strain *P.putida*, WCS358, can increase crop yields presumably by competing successfully for available iron with microorganisms deleterious to the crop. The genetics of the siderophore biosynthesis of this strain is being investigated (Marugg *et al* 1988). Many species are phytopathogenic for some important crop plants. For example, *P.syringae pv syringae* causes brown spot disease of *Phaseolus vulgaris* and *P.marginalis* is a pathogen of witloof chicory. The potential for the genetic manipulation of these plant-associated pseudomonads for use in growth promotion or alteration of virulence determinants of the organisms is enormous.

With the recent growth in the field of biotechnology much interest has been directed towards these organisms, particularly with a view to creating and releasing into the wild, genetically modified organisms which would confer beneficial properties on crops. For instance, *P.syringae* and *P.fluorescens* have been genetically altered such that they lack the ice

nucleation protein which normally causes extensive frost damage to plant surfaces by acting as a focus for ice nucleation when temperatures are within 5°C below freezing (Warren *et al* 1987). Limited field trials with strawberry plants carried out in the USA have been successful in terms of the effective competition of these ice nucleation deficient species with the resident microbial flora, but have not taken place without protest from people opposed to the deliberate release of genetically modified organisms into the wild.

1.3.3 Xenobiotic Degradation by *Pseudomonas*

A xenobiotic is a chemical compound which is foreign to the biosphere (introduced by humans) of which many are environmental pollutants due to their inability to be degraded. Many bacterial species play a part in the biodegradation of chemical compounds in the soil, but with the industrialisation of the world, waste chemical products which have never before been encountered in nature cannot be metabolised by the native microbial flora. *Pseudomonas* species are able to metabolise and derive energy from a wide range of aromatic compounds and for this reason there are commercial biotechnological interests in *Pseudomonas*. By studying the biochemistry of these pathways it is hoped that ultimately they can be modified either to produce a commercially valuable novel compound or to metabolise (and thus degrade) other xenobiotic (and pollutant) materials.

An example of a way in which speciality chemicals can be produced is the combination of metabolic pathways from different species. The dye indigo was produced from cheap substrates by combining the naphthalene pathway from *P.putida* and the tryptophan pathway of *E.coli* into a single gene cluster. *E.coli* harbouring this recombinant pathway exhibit the regulated expression of the indigo pathway genes (Ensley *et al* 1983). An

example of the selective evolution of *Pseudomonas* enzymes is shown by the derivation of mutant strains which exhibited altered substrate specificities for the amidase enzyme of *P.aeruginosa*. This enzyme is encoded by the *ami*E gene which is positively activated by the product of the closely linked *ami*R gene (Farin and Clarke 1978). The wild type enzyme has a narrow substrate specificity, supporting growth on a small number of amides as sole carbon and nitrogen source. These amides, in addition to acting as substrates also act as inducers of the enzyme. Acetamide and propionamide are the best subtrates and inducers of the wild type enzyme. Brown *et al* (1969) derived a mutant able to utilise butyramide and Brown and Clarke (1972) isolated a mutant which could utilise acetanilide. Betz and Clarke (1972) derived a set of mutants able to utilise phenylacetamide. This biochemical versatility of the pseudomonads illustrates the great potential for the generation of strains with altered substrate specificities which may have importance for the generation of cheap food materials by utilisation of industrial waste compounds for bacterial growth.

P.putida has in particular been intensively investigated with regard to its possible biotechnological exploitation since it exhibits one of the widest nutritional versatilities of the pseudomonads. Many of the catabolic pathways that this species possesses are encoded by the chromosome, for example the benzene oxidation genes (Irie *et al* 1987a), and those for the assimilation of benzoate to β-ketoadipate (Aldrich and Chakrabarty 1988). In addition, large plasmids of greater than 150kD encoding degradative pathways have been isolated from several strains of *P.putida*. These are more easy to manipulate than chromosomally-encoded genes, since they can be transferred between strains and the genes of the whole pathway can be cloned more easily.

1.3.3.1 The TOL Plasmid

The TOL plasmid, pWWO, is a 117kb plasmid isolated from *P.putida* mt-2 (Williams and Murray 1974) and it encodes genes for the degradation of toluene and xylenes to intermediates of the TCA cycle. The TOL plasmid is self-transmissible and possesses replication functions for replication in several pseudomonads. Other plasmids for the degradation of toluene have been isolated from other strains of *P.putida* (Williams and Worsey 1976). These carry the same genes, but the organisation of the operons is different. This pathway has been intensively studied and much is known about its biochemistry and genetic regulation. In pWWO the pathway is divided into two regulatory blocks of genes. The *xylCAB* operon is responsible for the degradation of toluene and xylenes to aromatic carboxylic acids (the "upper" pathway) and the *xylDLEGF* operon converts these carboxylic acids to TCA cycle intermediates (the "lower" or "meta" pathway) (Franklin *et al* 1981). The involvement of two regulatory proteins encoded by genes *xylR* and *xylS* is necessary for induction of the pathways (Worsey *et al* 1978).

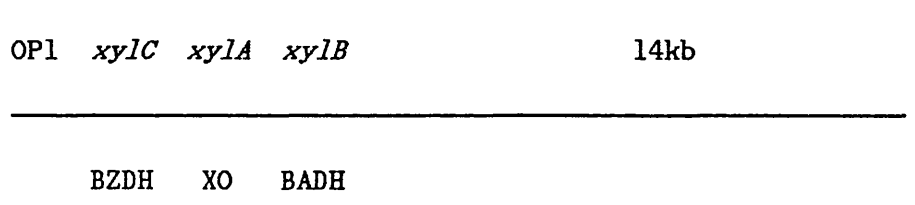
Many of the TOL genes were elucidated by the generation of mutations in the structural genes (Worsey *et al* 1978), whilst their molecular characterisation involved the use of more sophisticated techniques, such as those of transposon mutagenesis and gene cloning. In this way Franklin *et al* (1981) mapped the genes on the TOL plasmid discovering that the two operons were physically separated by 14kb of DNA. The regulatory genes *xylR* and *xylS* have been mapped and cloned (Inouye *et al* 1981, Franklin *et al* 1983, Inouye *et al* 1983). The organisation of the genes is illustrated in Figure 1.

Worsey *et al* (1978) proposed a regulatory model of the pathway whereby the product of *xylR* can combine with either *m*-xylene or *m*-methylbenzyl alcohol to induce both regulatory blocks. The product of the *xylS* gene can combine with *m*-toluate to induce the pathway for the degradation of the

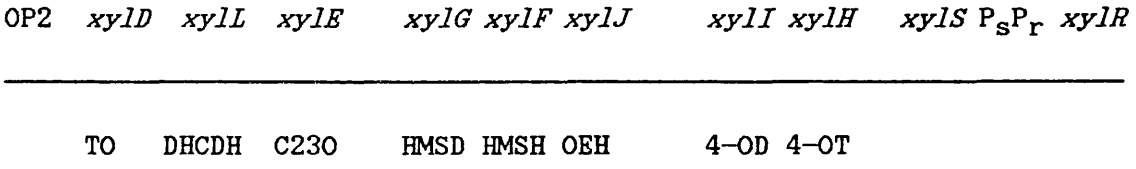
FIGURE 1

TOL GENES OF pWWO

Upper operon



Lower operon



Abbreviations

BZDH	benzaldehyde dehydrogenase
XO	xylene oxygenase
BADH	benzyl alcohol dehydrogenase
TO	toluate/benzoate oxygenase
DHCDH	dihydroxycyclohexadiene carboxylate dehydrogenase
C23O	catechol 2,3-oxygenase
HMSD	hydroxymuconic semialdehyde dehydrogenase
HMSH	hydroxymuconic semialdehyde hydrolase
OEH	2-oxopent-4-enoate hydratase
4-OD	4-oxalocrotonate decarboxylase
4-OT	4-oxalocrotonate tautomerase
OP	operator/promoter
P_s	promoter of *xylS*
P_r	promoter of *xylR*

carboxylic acids. Although Worsey's model has remained basically correct, the regulatory mechanisms have been extended in the light of more recent experiments. It seems that both *xylS* and *xylR* are required for activation of the lower pathway by *m*-xylene or *m*-methylbenzyl alcohol (Inouye *et al* 1981, Inouye *et al* 1983). From more recent experimental results it appears that the expression of *xylS* and *xylR* is under tight regulatory control and a new regulatory model has been proposed (Inouye *et al* 1987a). XylR together with m-xylene or m-methylbenzylalcohol (m-MBA) induces the *xylCAB* operon and *xylS*. The XylS product then activates the *xylDLEGF* operon. Hence, *xylR* indirectly causes activation of *xylDLEGF* via the activation of *xylS*; m-xylene is degraded by the enzymes of the *xylCAB* operon and the products (m-MBA, m-tolualdehyde and m-toluate), together with XylS induce the *xylDLEGF* operon. m-MBA interacts with XylR to induce *xylS* giving rise to activation of *xylDLEGF*; m-MBA can also interact with XylS alone, when the XylS level is high, to activate *xylDLEGF*. The operon can also be activated by the interaction of m-toluate with XylS even when the XylS protein is at a non-induced level. The expression of *xylR* is autoregulated, such that XylR represses the activity of *xylR*. Further to this, it has been found that overproduction of XylS can activate the *xylDLEGF* operon in the absence of XylR and inducer (Spooner *et al* 1987, Inouye *et al* 1987b)

The product of the *xylR* gene has been identified in the *E.coli* maxicell system as a 68kD protein (Spooner *et al* 1986) and as a 67kD protein by Inouye *et al* (1985). Sequence data (Spooner *et al* 1986) revealed that the expected size of the *xylS* product was 36.5kD. This was confirmed when the *xylS* gene was cloned under control of the *tac* promoter and the 36kD product detected in the *E.coli* maxicell system (Spooner *et al* 1987). The TOL-specified genes were generally found to be poorly expressed in *E.coli* (Franklin *et al* 1981). The transcription initiation points are the same in both species but the amount of mRNA synthesised from the *xylCAB* and *xylDLEGF* operons in *E.coli* under induced conditions was found to be only 5-

50% of that synthesised by *P.putida* (Inouye *et al* 1984a, Inouye *et al* 1984b). This difference is either due to a reduction in the amount of XylR and XylS activator proteins or a real difference in transcription from the two operons such that *E.coli* RNA polymerase inefficiently transcribes the operons due to its inability to recognise the promoter signals. It does not appear to be due to a limited supply of activator proteins since the amount of *xylR* mRNA is approximately the same in both species (Inouye *et al* 1983). The promoters of the two operons and of the activator genes have been determined and they do not exhibit typical *E.coli*-like sequences (see 1.5.1.3.1).

1.3.3.2 Other Degradative Plasmids

In addition to the TOL plasmid, many other plasmids encoding degradative pathways have been isolated. For example, these include those which encode degradation of camphor (CAM), n-octane (OCT), salicylate (SAL), naphthalene (NAH), 2-hydroxypyridine (2-HP), nicotine/nicotinate (NIC). Most of these are very poorly characterised genetically in comparison to the TOL plasmid, however there exists a growing amount of information on the genetics of the NAH7 plasmid of *P.putida* (Dunn and Gunsalus 1973). This plasmid is self-transmissible between *Pseudomonas* species. The degradative pathway, as for the TOL plasmid, is organised into two operons, encoding enzymes for the catabolism of naphthalene to salicylate (*nah* operon) and the other encodes enzymes for the catabolism of salicylate to tricarboxylic acid intermediates (*sal* operon). Induction of the pathways requires the inducer, salicylate, and the product of the *nahR* gene (Yen and Gunsalus 1985). NahR appears to bind to the operons in the presence or absence of salicylate, but only when salicylate is present is transcription activated. Homologous sequences upstream of the *nah* and *sal* operons are possibly the region where NahR binds. The promoter sequences of *nah, sal* and *nahR* have a structure closely homologous to that of most

E.coli promoters (Schell 1986), an observation which may explain why expression of these genes is high in both *E.coli* and *P.putida*.

1.4 GENETICAL ANALYSIS OF *PSEUDOMONAS*

The biochemistry of many of the degradative pathways has been elucidated (Clarke and Ornston 1975) but information about the regulation of the pathways requires a knowledge of the genetics of the pathways. For this sort of analysis a variety of genetical techniques are needed. Elaborate methods for the genetic manipulation of *E.coli* have been devised, but the use of these for the manipulation of *Pseudomonas* have been limited. The major reason for this being that many pseudomonad degradative genes tend to be poorly expressed in *E.coli* due presumably to the differing metabolic and physiological backgrounds. There is, therefore, a need for techniques based on *Pseudomonas* such that pathways can be manipulated in the native host rather than having to work with the heterologous host in which problems with expression arise.

1.4.1 Host-Vector Systems

1.4.1.1 Cloning Vectors for use with *Pseudomonas*

In the last few years host-vector systems have been developed for *Pseudomonas*. These are highly valuable in the rapidly growing field of analysis of the degradative pathways exhibited by the group. For the construction of vectors for the cloning of pseudomonal activities, it was necessary to make use of the group of plasmids referred to as broad host range (see Chapter 4), since those commonly employed as *E.coli* cloning vectors will not replicate and be maintained in *Pseudomonas*. Most of the broad host range plasmids in current use are based on the IncQ plasmid RSF1010 (or R300B), the IncP plasmid RK2 (or RP4/RP1) or the IncW plasmid pSa.

RSF1010 (or R300B or R1162) is a useful plasmid for the development of

broad host range cloning vectors due to its small size (8.7kb) compared with other broad host range plasmids and its ability to replicate and be maintained in a large number of Gram-negative species. It is of little use itself as a cloning vector since, in common with other broad host range plasmids, there are very few suitable restriction sites for cloning. Consequently vectors have been constructed which possess additional antibiotic resistance genes possessing restriction sites allowing insertional inactivation. Bagdasarian *et al* (1981) produced a series of vectors based on RSF1010, some of which, pKT248, pKT210 and pKT215 possess the chloramphenicol resistance gene from either R621ala or pSa. Others, pKT231 (Km^r from R6-5), pKT230 (Km^r from pACYC177) have a kanamycin resistance determinant. These vectors are all about 12-13kb in size. Sharpe (1984) constructed a series of vectors based on R300B containing regions of DNA from pBR322 and pBR328 providing antibiotic resistance markers with potentially useful cloning sites. One of these vectors, pGSS33, is particularly useful being 13.4kb (reasonably small) and expressing resistance to ampicillin, tetracycline, chloramphenicol and streptomycin, all of which have restriction sites for insertional inactivation.

Vectors based on RK2 (IncP) are more difficult to construct since the plasmid is 56kb in size and the replication and maintenance functions of the molecule are spread round the plasmid such that functional derivatives of a reasonable size ($<$ 20kb) are difficult to derive. The vectors pRK290 and pRK2501 are low copy number vectors useful for the cloning of genes whose products at high concentrations are host lethal. pRK290 was constructed by deletion of DNA non-essential for replication (Ditta *et al* 1980). It is 20kb in size, non-self transmissible, but can be mobilised by appropriate conjugal transfer functions provided *in trans*. It has a copy number of 5-8 in *E.coli* and has unique *Bgl*II and *Eco*RI sites for DNA insertion. pRK2501 is an 11.1kb plasmid expressing resistance to Km and Tc which provide restriction enzyme sites for insertional inactivation (Kahn

et al 1979).

Tait *et al* (1983a) constructed a series of 5 vectors based on the IncW plasmid pSa (29.6kb) by deletion of the conjugal transfer functions. They all contain the origin of replication from pSa, the replication gene (*rep*A) and express resistance to spectinomycin (Sp) and Km. Some possess Cm^r of Tn9. They are all 15kb or less in size. Leemans *et al* (1982) also derived broad host range vectors based on pSa, the simplest construction involving the recircularisation of a *Bgl*II fragment to give pGV1106, an 8.4kb plasmid lacking the conjugal transfer functions and expressing resistance to Km and Sm/Sp. Other vectors have been derived from this by the integration of other antibiotic resistance genes. For example, the addition of Tc^r from pBR325 yielded pGV1120 and the addition of Cm^r from pACYC184 giving pGV1124. Some of the vectors possess the ColE1 origin of replication in addition to that of pSa and these plasmids are amplifiable by Cm.

Broad host range vectors should either be self-transmissible or better non-self transmissible but mobilisable. This circumvents problems arising from the low efficiency of transformation of *Pseudomonas* species providing an effective method for introduction of DNA into these species particularly when it is necessary to maximise the number of recombinants in a cloning experiment.

Most vectors available for use with *Pseudomonas* are based on broad host range plasmids, but cloning vectors have been derived from *Pseudomonas* plasmids. These include the series of vectors developed from the *Pseudomonas* plasmid pVS1 (Itoh and Haas 1985). These vectors have a narrow host range, not being maintained in *E.coli*, but they are maintained in a range of *Pseudomonas* species.

1.4.1.2 *Pseudomonas* Host Strains

Strains defective in the host-specific restriction system have been isolated for *P.aeruginosa* and *P.putida* (Bagdasarian *et al* 1981) so that DNA

prepared in *E.coli* is not restricted on entry into these strains. These are *P.aeruginosa* PAO1162 and *P.putida* mt-2 KT2440. It is necessary to have high efficiency methods for the uptake of DNA by these strains. Although conjugation provides this, it is often desirable that the strains can be transformed with DNA since this is more direct and simpler. Although *Pseudomonas* species are quite refractory to transformation, methods developed for *E.coli* have been adapted and an efficiency of about one order of magnitude lower than that for the commonly used *E.coli* strains can be obtained (Bagdasarian *et al* 1981). It is also necessary that the host strains do not carry endogenous plasmids which might interfere with restriction analysis. Strain 2440 has no endogenous plasmids whilst PAO1162 possesses the FP2 sex factor but this is so large that it is not normally recovered in preparations of plasmid DNA and therefore does not interfere with the *in vitro* analysis of plasmids (Bagdasarian *et al* 1981).

1.4.2 Broad Host Range Cosmids

Many of the broad host range vectors developed for use in *Pseudomonas* have been modified further for the creation of broad host range cosmids. These are plasmids containing the *cos* site of λ such that large (35-40kb) fragments can be cloned and packaged *in vitro*. For the cloning of a specific gene the packaged plasmids in λ phage can be used to infect *E.coli* containing a conjugal plasmid such that the cosmid clones can be transferred to a host which is mutant in the gene in question and the recombinant cosmids possessing the gene can be identified by complementation. Such a method has been used for the cloning of the vanillate demethoxylase genes from a *Pseudomonas* spp. using the RSF1010-based cosmid pJRD215, a 10.2 kb plasmid carrying Km and Sm resistance genes and at least 23 unique restriction sites (Davison *et al* 1986). The above method was also adopted by Frey *et al* (1983) for the cloning of the genes specifying *myo*-inositol transport from *Pseudomonas*

strain JD34 using cosmids pMMB33 and pMMB34. These are RSF1010-based cosmids constructed from pKT231 which are 13.75kb in size and express Km^r.

1.4.3 Broad Host Range Expression Vectors

In addition to the need for broad host range cloning vectors, there is often a requirement for overproduction of cloned gene products and to this end, several broad host range expression vectors have been derived. These allow tightly regulated expression, since constitutive high expression of some genes may affect cellular physiology causing a selective pressure for loss of the cloned gene. Mermod *et al* (1986) have constructed a 14kb broad host range expression vector, pNM185, based on the promoter of the *xylDLEGF* operon and carrying the *xylS* gene. This vector was derived from the RSF1010-based vector pKT231 and regulated expression was found to occur in 16 different genera of Gram-negative bacteria. It confers regulated expression by virtue of specific induction of the meta promoter by the *xylS* product and coinducer (benzoate or toluate). Levels of induction of the promoter are more than 20 fold in *P.putida* and 600 fold in *E.coli*. This vector has been further improved (Ramos *et al* 1988) by the isolation of a plasmid from pNM185 which has a mutation in the *xylS* gene such that the product can be activated by 3-methylsalicylate and this activation caused a 3.5 fold increase on the promoter induction level observed when *m*-toluate was used to activate XylS. In addition a further mutation was isolated such that temperature range of expression of the gene was extended and therefore the vector could be used with species whose temperature optima for growth is 40°C. A multiple cloning site was cloned into the vector adjacent to the meta promoter to increase its versatility. The resultant plasmid, pERD21, is 13.8kb and specifies Km^r.

As well as the use of *Pseudomonas* gene promoters for the derivation of broad host range expression vectors, several have been produced which are based on the *tac* (hybrid *trp-lac*UV5) promoter (see 6.2.1) which has been

shown to function efficiently in pseudomonads as well as *E.coli* (Bagdasarian *et al* 1983). Vectors pMMB22 and pMMB24 were derived from the RSF1010-based vector pKT240 by the insertion of the *tac* promoter and the *lac*I^q gene (Bagdasarian *et al* 1983). The latter allows good repression of the *tac* promoter and activation by IPTG giving controlled expression. These vectors have been modified (Deretic *et al* 1987d) by insertion of a polylinker from pUC12 containing the *lacZ'* gene adjacent to the *tac* promoter. The plasmids so created were pVDtac39 (13.2kb, Ap^r) and pVDtac24 (12.7kb, Ap^r).

1.4.4 Broad Host Range Promoter-probe Vectors

A number of promoter-probe vectors are available for use in *E.coli*, but similar vectors for use in other species are limited. The promoter-probe vector pKT240 was derived by Bagdasarian *et al* (1983). This is a vector based on RSF1010 with Km^r and Ap^r genes incorporated from pHSG415 with a promoterless Sm^r gene. It has a copy number of 10-20 in *E.coli* and 20-40 in *P.aeruginosa* (Jeenes *et al* 1986). By cloning DNA fragments upstream of this gene their relative promoter activities can be determined by plating cells harbouring the recombinants on agar plates containing varying concentrations of streptomycin. A promoter-probe vector has been produced for the particular use in *Xanthomonas campestris*, since pKT240 was found to be unsuccessful (Osbourn *et al* 1987). This vector, pIJ3100 is 13.2kb, is based on RSF1010 and expresses resistance to Sm and has a promoterless chloramphenicol acetyltransferase gene. During the course of this work a series of promoter-probe vectors have been described which are based on RSF1010 and carry the structural genes for chloramphenicol resistance, penicillin resistance or firefly luciferase without their native promoters (Greener and Helinski 1987). Also recently a broad host range vector based on RSF1010 has been constructed carrying the promoterless *xylE* gene (Konyecsni and Deretic 1988). This plasmid is essentially the same as the promoter-

probe vector whose construction is described in this work.

1.4.5 Transposon-Donor Plasmids

Transposon-mediated mutagenesis is an extremely powerful tool in molecular biology for mapping genes by virtue of disruption of the gene into which the transposon inserts. It creates very stable mutations and by creating a battery of transposon insertions (usually random) into a piece of DNA the start and end of a gene can be mapped very easily using restriction analysis. Transposons can also facilitate cloning of a specific gene by providing a source of mobile restriction sites. There are numerous and sophisticated methods available in *E.coli,* but these cannot usually be adapted for use in other hosts. Transposon-donor plasmids have been developed based on the IncI-α plasmid ColIb for use in *P.putida* and other Gram-negative bacteria (Boulnois *et al* 1985). pLG221 carries the transposon Tn5 encoding Km^r and pLG223 Tn10 encoding Tc^r and in both cases the transposon lies in the colicin Ib gene. These transposon-donor plasmids have the ability to promote conjugation between *E.coli* and other Gram-negative bacteria, but they behave like suicide vectors in *P.putida* since they are not maintained in this host. Thus selection of recipients with kanamycin isolates those cells which have inherited a copy of Tn5 by virtue of its integration into the chromosome. In this way the chromosome of several Gram negative bacteria can be mutated.

1.4.6 Gene Expression Systems

Several systems for the study of polypeptides expressed by cloned genes have been developed in *E.coli.* These include the minicell (Meagher *et al* 1977) and maxicell (Sancar *et al* 1979) semi-*in vivo* systems and the *in vitro* coupled transcription/translation system (Zubay 1973, Pratt *et al* 1981). These provide a rapid and reliable method for analysing plasmid-encoded polypeptides. The coupled transcription/translation system

consists of an S30 supernatant (a supernatant derived from the centrifugation of disrupted cells at 30 000xg) plus the addition of a multitude of components thought to be required for protein synthesis. This system allows the expression of genes encoded by exogenous DNA. At the present time maxicell strains and strains for the production of minicells are only available for *Bacillus subtilis* (Shimotsu *et al* 1983, Reeve *et al* 1973) and *E.coli*. The *in vitro* system has been successfully modified for *B.subtilis* (Leventhal and Chambliss 1979, McLaughlin *et al* 1981), *Streptomyces lividans* (Thompson *et al* 1984), *Rhizobium meliloti* (Dusha *et al* 1986) and *Agrobacterium tumefaciens* (Kartasova *et al* 1981).

In addition to the *in vivo* systems available for *E.coli* (minicells and maxicells), the *in vitro* gene expression system of *E.coli* (Zubay 1973, Pratt *et al* 1981) has come to play an important part in the analysis of gene products. Its main use in the laboratory has been the identification of cloned gene products. It can also be used for the location of genes on restriction fragments, something which is not possible with the *in vivo* systems. Commercially ideas have been mooted about the possibility of using an *in vitro* system for the large-scale production of particular gene products rather than using whole cells (Riordan 1987). A major problem with the use of the system on a large scale is that the *in vitro* system is subject to a total shutdown of protein synthesis after a certain length of time.

An advantage of the *in vitro* gene expression system over the *in vivo* system is that although the initial preparation of an extract is labour-intensive and time-consuming, once it is made it is stable at -70°C for several years and the assay is simpler and more reproducible (as long as the DNA is of good quality) than that of minicells or maxicells. In addition, the *E.coli* coupled transcription/translation system is now available commercially, eliminating the need for the initial production of an extract. There were no gene expression systems based on any species of

Pseudomonas and, given the problems encountered with the expression of some *Pseudomonas* genes in *E.coli*, the development of such a system is considered to be a valuable addition to the techniques now available for the genetic analysis of *Pseudomonas*.

1.5 TRANSCRIPTIONAL AND TRANSLATIONAL BARRIERS TO THE EXPRESSION OF *PSEUDOMONAS* GENES IN THE HETEROLOGOUS HOST

In the last ten years enormous strides have been made in the derivation of techniques for the genetic manipulation of the pseudomonads, although we are still very limited compared with the sophistication of techniques available for use with *E.coli*. Over this time, with the growing interest in the pseudomonads, various *Pseudomonas* genes have been cloned. Due to the paucity of methods for the examination of the products in the native host, *E.coli*-based systems have been by necessity relied upon for the analysis of these genes in terms of their regulation and protein products. A major problem in obtaining this information has been that of expression of cloned *Pseudomonas* determinants. For this reason the orthodoxy soon evolved that while *E.coli* genes are expressed well in *Pseudomonas*, many *Pseudomonas* genes are poorly expressed in *E.coli*. However in the light of further investigations of several of the genes whose recalcitrance to expression in the heterologous host had given rise to this opinion, it has been revealed that many of them require the presence of an activator protein for expression. Since in many cases the properties being examined were unique to the *Pseudomonas* species from which the genes were derived, a suitable activator molecule was not present in *E.coli*. For example, the cloning in *E.coli* of the *nah* genes in the absence of *nahR* (Schell 1983) and exotoxin A from *P.aeruginosa* (Gray *et al* 1984) resulted in a failure of these genes to be expressed. This is not however the reason for the low level of expression exhibited by all *Pseudomonas* genes in *E.coli*, for example the TOL genes. This barrier to expression could potentially be at the trans-

criptional or translational level or both.

There are numerous reports of *Pseudomonas* genes being poorly expressed in *E.coli*, but fewer examples of the relative expression of *E.coli* genes in *Pseudomonas*. This poor expression of *Pseudomonas* genes in *E.coli* has been reported for both chromosomal and plasmid-encoded genes. Buckel and Zehelein (1981) found that expression of the *P.fluorescens* D-galactose dehydrogenase gene in *E.coli* was greatly reduced compared with its native host. However, under control of the *E.coli lac* promoter the level of expression of this gene was substantially increased. The histidine utilisation (*hut*) genes from *P.putida* were expressed in *E.coli* at about 50% of the level of expression in *P.putida* (Consevage *et al* 1985). The *lysA* gene cloned from *P.aeruginosa* was expressed in *E.coli* at only 2% of the level of expression in *P.aeruginosa* (Martin *et al* 1986). Minton *et al* (1983) reported that a cloned *Pseudomonas* carboxypeptidase G2 gene was inefficiently expressed in *E.coli*, compared with expression in *P.putida* which was 30 fold higher. Conversely the *alk* genes of the OCT plasmid of *P.putida* have been reported to be quite efficiently expressed in *E.coli* and the regulatory control in both species is the same (Eggink *et al* 1987). The genes from the TOL plasmid of *P.putida* show varying degrees of expression in *E.coli* and this appears to reflect their promoter sequence (see 1.5.1.3.2). As mentioned above *E.coli* genes were reportedly well expressed in *Pseudomonas*, for example the expression of the cloned *E.coli argE* and *argF* genes in *P.fluorescens* was tested and they were found to be expressed (Mergeay *et al* 1978).

To test the hypothesis that *Pseudomonas* genes are poorly expressed in *E.coli*, whereas *E.coli* genes are well-expressed in *Pseudomonas* more directly, Jeenes *et al* (1986) performed a two way experiment examining the expression of biosynthetic genes from both *E.coli* and *P.aeruginosa* in the respective heterologous host. They found that the *P.aeruginosa* genes *argA*, *argF* and *proC*, none of which require a positive activator were all poorly

expressed in *E.coli*, exhibiting only 0.3-5% of the level of expression in *P.aeruginosa*. Conversely the *argA* and *argC* biosynthetic genes of *E.coli* were expressed constitutively at moderate levels and the *prcB* and *prcC* genes at high levels in *P.aeruginosa*. These results for biosynthetic genes lend support to the orthodoxy that *Pseudomonas* genes are poorly expressed in *E.coli*, but *E.coli* genes are well expressed in *Pseudomonas*.

The expression of genes from other Gram negative species in *Pseudomonas* has also been of interest, especially those of species of close similarity to *Pseudomonas*. The expression of *Acinetobacter calcoaceticus cat*BCDE genes in *E.coli* and *P.putida* has also been investigated (Shanley *et al* 1986). In *A.calcoaceticus* these genes are regulated by induction, but in *P.putida* they are constitutively expressed. In *E.coli* when the genes were cloned into pBR322 such that they were not under the control of a vector promoter there was expression of *cat*CDE, but not of *cat*B. This is thought to be due to an internal promoter within the operon. Under control of the *lac* promoter, the genes were expressed in *E.coli* suggesting that the barrier to expression was not at the translational level. Chatterjee and Chatterjee (1987) found that the TOL genes cloned on an RP4-TOL hybrid could be efficiently expressed in *Caulobacter crescentus*, a bacterium which shares several important characteristics with the pseudomonads (Poindexter 1981).

1.5.1 Factors affecting Transcription

1.5.1.1 RNA Polymerase Molecule

The enzyme which catalyses the synthesis of mRNA complementary to the DNA template is RNA polymerase. The components of the holoenzyme are $\alpha_2\beta\beta'\sigma$. The genes which code for these have been identified and sequenced and the molecular weights of the protein products have been predicted. This is more accurate than previous values based on electrophoretic mobility by SDS-PAGE. In addition to these subunits a whole array of accessory proteins

have been identified with sizes in the range of 10kD to 105kD (Ishihama *et al* 1983). The DNA-dependent RNA polymerases from *P.aeruginosa, P.putida* and *Pseudomonas* BAL-31 (subsequently reclassified as *Alteromonas espejiana*) have been purified and the subunits analysed (Allan and Kropinski 1987, Johnson *et al* 1971, Zimmer and Millette 1975) (Table 3). The subunit structures are similar to those of other eubacterial RNA polymerases, namely $\alpha_2\beta\beta'\sigma$, but the sizes of the specific subunits do differ. Although the genes encoding these proteins have not been cloned and sequenced so that no accurate value for molecular weights has been obtained, there seem to be some differences in the apparent molecular weights of some of the subunits (particularly α and σ) compared with those of *E.coli.* The preference of various RNA polymerase molecules for specific DNA templates was tested. Zimmer and Millette (1975) found that the template preference of the *P.putida* RNA polymerase was the same as *E.coli*, in that there was no preference for the native template and T7 phage templates were more efficiently transcribed than any other. Allan and Kropinski (1987) report a similar situation with the *P.aeruginosa* molecule with T7 DNA being preferentially utilised compared with ΦPLS27 (a podovirus which infects *P.aeruginosa*). This is most likely due to the strength of T7 promoters (Deuschle *et al* 1986, see 1.5.1.3).

Although little investigation has been carried out on RNA polymerase from *Pseudomonas,* it seems likely, in view of the similarities in subunit structure and the ability of the *Pseudomonas* transcription/translation machinery to express *E.coli* genes, that there are no significant differences in the core enzyme. Indeed, Gragerov *et al* (1984) showed that although the subunit sizes of RNA polymerase from *E.coli* and *P.putida* differ, the structures of DNA-binding centres are highly similar since they show identical DNA contacts with several *E.coli* promoters.

TABLE 3

SUBUNIT COMPOSITIONS OF RNA POLYMERASE FROM *E.COLI* AND *PSEUDOMONAS*

Subunit composition of RNA polymerase of *E.coli*

Subunit	Molecular weight (in Daltons)
α	36 512
β	150 618
β'	155 163
σ	70 263
ω	10 000

Subunit composition of RNA polymerase of *P.putida*

Subunit	Molecular weight (in Daltons)
α	44 000
β	155 000
β'	165 000
σ	98 000

Subunit composition of RNA polymerase of *P.aeruginosa*

Subunit	Molecular weight (in Daltons)
α	45 000
β	148 000
β'	157 000
σ	87 000

Subunit composition of RNA polymerase of *Pseudomonas* BAL-31

Subunit	Molecular weight (in Daltons)
α	38 000
β	155 000
β'	165 000
σ	89 000

1.5.1.2 Sigma Factors

It is the σ subunit of RNA polymerase holoenzyme which confers on it the ability to recognise and therefore bind to specific promoter sequences. In *E.coli* there is a major σ subunit (σ^{70}) which allows RNA polymerase to recognise -35 and -10 sequences typical of *E.coli* promoters. The first evidence for the existence of multiple σ factors came from the studies of *B.subtilis* phages which encode σ factors responsible for the expression of middle and late phage genes, the early genes being transcribed by the host polymerase holoenzyme (Losick and Pero 1981). In addition to these phage genes, *B.subtilis* expresses minor σ factors which have specific roles in terms of the genes whose promoters they recognise (Gillman *et al* 1981, Johnson *et al* 1983). *B.subtilis* has a complex developmental process involving sporulation in which there is successive expression of specific genes. The evidence so far suggests that alternate σ factors are involved in the expression of genes whose products are required at a certain stage of development and genes which control the response to environmental shock, including nutrient deprivation, heat and chemical shock (Doi and Wang 1986).

1.5.1.2.1 Evidence for the existence of multiple σ factors in *E.coli*.

The first evidence for the existence of alternative σ factors in *E.coli* came from studies of the heat shock response. It was observed that when *E.coli* cells were subjected to elevation of growth temperature from 30°C to 40-45°C there was an increased synthesis of 17 proteins (heat shock proteins) (Neidhardt *et al* 1984). This is a universal effect occurring across a spectrum of prokaryotes and eukaryotes and at least two of the heat shock proteins are conserved across this large phylogenetic range (Gross *et al* 1987). The expression of heat shock genes in *E.coli* is governed by the product of the *rpoD* gene which was found to be a sigma subunit (σ^{32}) of RNA polymerase. Heat shock genes do not exhibit the

canonical -35/-10 promoter sequences of *E.coli*, but they have promoter sequences which are different to those recognised by the *E.coli* RNA polymerase holoenzyme ($E\sigma^{70}$). It has been shown (Gross *et al* 1987) that the σ^{32} confers on RNA polymerase the ability to recognise these novel promoter signals. A concensus sequence for these promoters has been reported by Cowing *et al* (1985), it has a -35 region of TNtCNCcCTTGAA and a -10 region of CCCCATtTA (the upper case indicating those bases which are strictly conserved and the lower case indicates bases that are less conserved). This concensus sequence was constructed from the sequence of 3 heat shock promoters. Straus *et al* (1987) and Skelly *et al* (1987) showed that σ^{32} is directly responsible for controlling gene expression, since following an increase in temperature, the amount of σ^{32} increases transiently. Ethanol has the same effect as heat shock.

E.coli encodes a protein which is analogous to the NtrA protein of *Klebsiella pneumoniae*, which is known to act as a σ factor (σ^{60}) for the transcription of nitrogen-regulated (ntr) genes (Hirschman *et al* 1985, see 1.5.1.3.1) in particular the nitrogen-fixation (nif) genes. The product of *glnF* in *E.coli* has been purified (Hunt and Magasanik 1985) following overproduction due to cloning the gene under control of the *tac* promoter and has been shown to bind to core RNA polymerase. Genes which are activated by the NtrA protein have a promoter sequence substantially different to that recognised by $E\sigma^{70}$. With the discovery of an ntr-type promoter for the gene for formate dehydrogenase (involved in anaerobic metabolism) in *E.coli*, it has been further established that this gene is dependent on NtrA for expression (Birkmann *et al* 1987).

Comparison of nucleotide sequnces upstream of flagellar and chemotaxis genes of *E.coli* and *S.typhimurium* has revealed considerable homology with promoter regions of genes of *B.subtilis* which are known to be activated by the minor σ subunit σ^{28} (Helmann and Chamberlin 1987). These enteric genes are not transcribed by $E\sigma^{70}$, hence these results suggest that there exists

another σ factor in *E.coli* analogous to the σ^{28} of *B.subtilis.*

Homologies exist between the peptide sequences of σ^{32} and σ^{70}. Fujita *et al* (1987) have used antiserum to a synthetic peptide which has homology to both of these σ factors (and is thus the putative core-binding site) and probed whole cell lysates. They found that in addition to σ^{32} and σ^{70}, several other polypeptides crossreact with the antibody and these have molecular weights of 75, 27, 23 kD. It is possible that these may be new σ factors but equally they may be degradation products of the known σ factors. None have been shown to associate with core RNA polymerase.

1.5.1.2.2 Evidence for the existence of multiple σ factors in *Pseudomonas.*

Since *E.coli* genes have no apparent barrier to expression in *Pseudomonas* this suggests that *Pseudomonas* must possess a σ factor analogous to σ^{70}. That many pseudomonad genes are not expressed in *E.coli* implies that there may be alternative σ factors responsible for the expression of these genes. There is little direct evidence for the existence of alternate σ factors in *Pseudomonas.* The observation that the heat shock response, studied in detail in *E.coli*, also occurs in *P.aeruginosa* (Allan *et al* 1988) and that it has a very similar kinetic and biochemical profile suggests that a σ subunit analogous to σ^{32} of *E.coli* must exist in *Pseudomonas.* Several of the induced proteins are immunologically similar in *P.aeruginosa* and *E.coli.* Four of the heat shock proteins synthesised in *E.coli* are also synthesised during *P.aeruginosa* heat shock. An attempt to identify the σ factor responsible was carried out by raising antibodies to RNA polymerase of *P.aeruginosa.* This approach was based on the supposition that, since σ^{32} and σ^{70} of *E.coli* show some amino acid sequence homology, then analogous σ factors of *Pseudomonas* should too. An additional product should be evident following heat shock. A polypeptide of about 40kD was observed which may be the factor analogous to σ^{32} of *E.coli.*

There is some indirect evidence for the presence in *Pseudomonas* of a protein analogous to the *ntr*A gene product and this comes from the finding that the *Klebsiella pnuemoniae nif* gene cluster can be expressed in *P.putida* (Postgate and Kent 1987) and that the *xyl*CAB operon from the *P.putida* TOL pathway is activated by NtrA in *E.coli* (Dixon 1986). Evidence for the existence of *ntr*A gene products in species closely related to *Pseudomonas* comes from the finding that some flagellar genes in *Caulobacter crescentus* have nif-like promoters (Mullin *et al* 1987). There have been several reports of the existence of ntr-like promoter sequences upstream of *Pseudomonas* genes (see 1.5.1.3.1).

1.5.1.3 Promoter Sequence

The majority of promoters in *E.coli* are recognised by the RNA polymerase holoenzyme Eσ^{70} and these exhibit the typical concensus sequence of TTGACA (-35) and TATAAT (-10) (Takanami *et al* 1976, Gilbert 1976). No naturally occurring promoter has yet been isolated with this concensus sequence. The analysis of relatively few promoters led to the recognition of these highly conserved -35 and -10 sequences. There is now a whole collection of known *E.coli* promoters and from these a series of possible sequence requirements for a promoter has been put forward. McClure (1985) investigated 112 promoter sequences and stated that all promoters which use σ^{70} for activation must possess at least 2 of the 3 highly conserved bases of the -10 region (TA---T), and at least one of the most highly conserved bases of the -35 region (TTG---). Promoters which show little homology to the concensus sequence tend to be positively activated. Harley and Reynolds (1987) calculated from 263 promoter sequences the frequency of each base in the -35 and -10 regions. This confirms McClure's findings. Russell and Bennett (1982) found that for both the *tet* and *lac* promoter the optimal spacing of the -35 and -10 regions was 17bp. Harley and Reynolds (1987) determined the spacing between the -35 and -10 regions of the 263 sequences

finding that most (53%) had a spacing of 17bp. Although spacings of 16bp (21%) and 18bp (19%) apparently can be tolerated, very few sequences had spacings of 15bp (1.5%), 20bp (0.4%) or 21bp (2.7%). The reason for this is probably that a base pair spacing increase or decrease of more than 1 from the concensus spacing creates too much of a conformational strain, since each base pair difference causes a 36° rotation in the relative positions of the two hexanucleotide sequences such that the three dimensional structure of the promoter is altered and cannot be recognised so efficiently by the RNA polymerase.

With the advent of DNA base sequence manipulation it has been possible to construct synthetic promoter sequences and to test their activities *in vivo* and *in vitro*. De Boer *et al* (1983) constructed 2 hybrid promoters from the *lac*UV5 -10 region (concensus -10) and the *trp* -35 region (concensus -35) forming the *tac*I promoter which has *E.coli* concensus -10 and -35 regions with a spacing between them of 16bp. The *tac*II promoter has the -35 sequence from the *trp* promoter and a -10 sequence of TTTAAT. This has one base difference to that of the *E.coli* concensus and that difference is in one of the highly conserved bases. The spacing between the -35 and -10 sequences is 17bp. The activity of these promoters in the activation of transcription was measured by cloning them on a plasmid adjacent to a promoterless galactokinase gene and measuring the activity of this enzyme in *E.coli*. They found that the *tac*I and *tac*II promoters were 11 and 7 times stronger respectively than the *lac*UV5 promoter. The *tac* promoter works well in *P.putida* (Bagdasarian *et al* 1983), *E.coli* and *P.aeruginosa* (Soldati *et al* 1987).

The events occurring in the promoter region can be divided into the recognition of the promoter by the RNA polymerase molecule; isomerisation of the promoter-RNA polymerase "closed" complex into a transcriptionally functional "open" complex; initiation of RNA synthesis and transition into

the elongation mode and clearance of the promoter. Recent experiments indicate that the -35 and -10 hexanucleotide sequences alone are not sufficient for determining promoter strength and that the optimisation of events other than the formation of a "closed" complex can modulate gene expression. Brunner and Bujard (1987) measured the degree of promoter recognition by host RNA polymerase and compared this with the observed promoter strength *in vivo* for several phage and *E.coli* promoters. They found that there was little correlation between the two indicating that promoter strength *in vivo* is determined by more than just one functional parameter (the recognition of primary promoter sequence).

Deuschle *et al* (1986) made a detailed analysis of 14 promoter sequences, by cloning them into a promoter-probe vector and assessing their activity *in vitro* and *in vivo* relative to the promoter of β-lactamase which was present with the structural gene as an internal standard on each plasmid. The promoters analysed were P_{lac}, P_{lacuv5}, P_{tacI} (trp/lacuv5), P_{con} (concensus), various phage T7 and T5 promoters and P_L from λ. Figure 2 illustrates their results. It is evident that the concensus promoter has a very low activity compared to that of the phage promoters which do not have the concensus sequence (Table 4). Indeed the one with the highest activity, the P_{A1} promoter of phage T5, has a sequence of TTGACT (-35) and GATACT (-10) and is 19 times more active than the concensus sequence (P_{con}). A comparison of nucleotide sequence data from the entire 70bp region covered by RNA polymerase of the strong T5 early promoters indicates, in addition to homologies in the -35 and -10 regions, there are significant homologies in the -43 and in the +1 and +7 regions.

Kammerer *et al* (1986) have redefined the promoter as a 70bp sequence (the total region covered by the RNA polymerase molecule of which 20bp are transcribed) including the -35 and -10 sequences as a core, but involving both upstream and downstream regions. They found that the *in vivo* and *in vitro* promoter strength of a phage promoter and an *E.coli* concensus

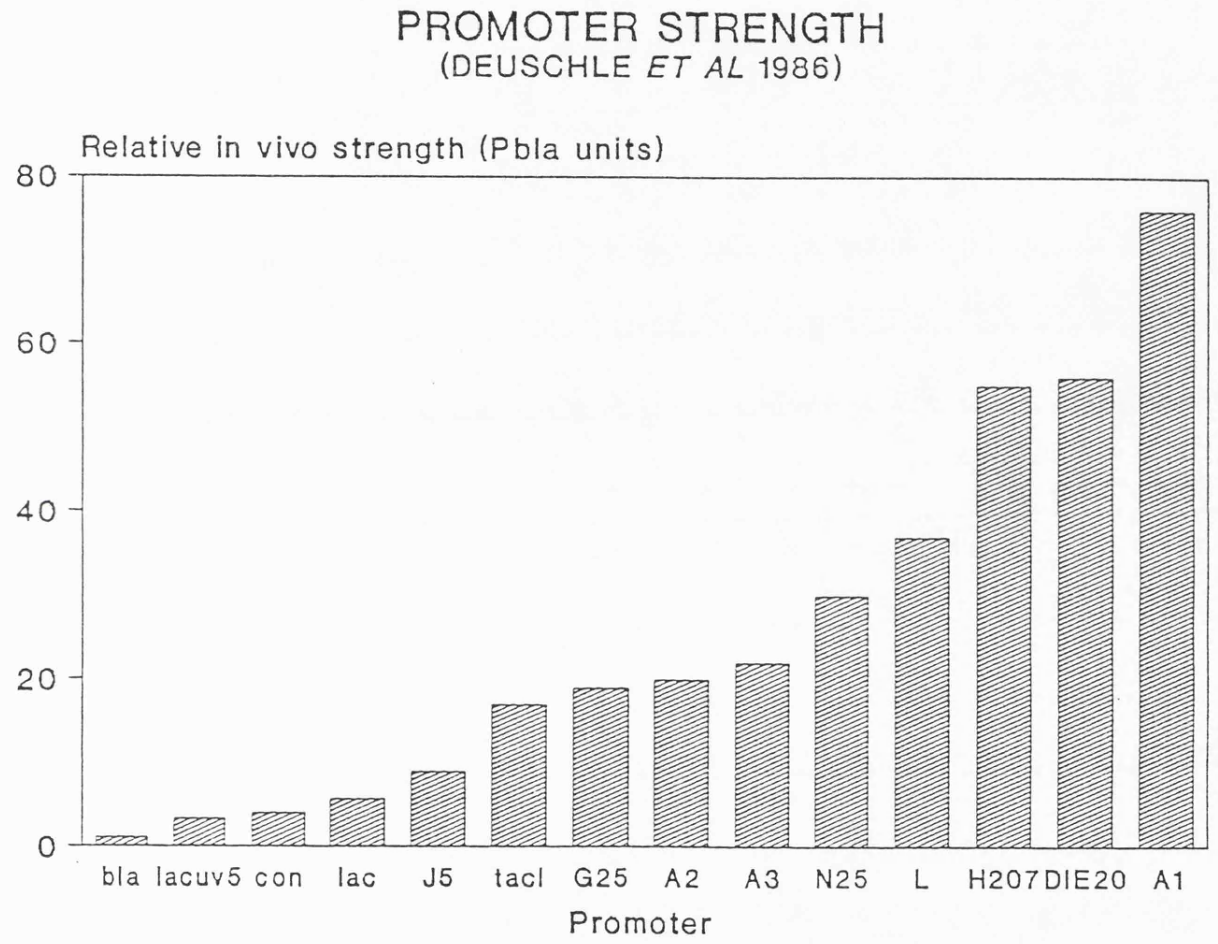

PROMOTER STRENGTH
(DEUSCHLE ET AL 1986)
Relative in vivo strength (Pbla units)
80
60
40
20
0
bla
lacuv5
con
lac
J5
tacI
G25
A2
A3
N25
L
H207
DIE20
A1
Promoter

FIGURE 2

TABLE 4

NUCLEOTIDE SEQUENCE OF PROMOTERS USED IN STUDY BY DEUSCHLE *ET AL* (1986)

PROMOTER	SOURCE	-35	-10	SPACING
P_{bla}	*E.coli*	TTCAAA	GACAAT	17
P_{lacuv5}	*E.coli* mutant	TTTACA	TATAAT	18
P_{con}	synthetic	TTGACA	TATAAT	17
P_{lac}	*E.coli*	TTTACA	TATGTT	18
P_{J5}	T5	TTTGACA	TATACT	17
P_{tac}	synthetic	TTGACA	TATAAT	16
P_{G25}	T5	TTGATA	TATAAT	17
P_{A2}	T7	TTGACA	TAAGAT	18
P_{A3}	T7	TTGACA	TACGAT	18
P_{N25}	T5	TTGCTT	TATAAT	17
P_{L}	λ	TTGACA	GATACT	17
P_{H207}	T5	TTGCTA	TATAAT	17
P_{DIE20}	T5	TTGACA	TAAGAT	17
P_{A1}	T7	TTGACT	GATACT	17

promoter could be altered by replacing the sequence downstream of the transcription start with downstream regions from other promoters. They also found that the sequences involved in promoter recognition and formation of a stable promoter-RNA polymerase complex were located upstream of -4 from the transcription start including the -35 and -10 sequences. Hence, the downstream regions must be involved in the later stages of transcription initiation. It appears then that promoters which exhibit the same strength can have very different structures, and promoters which have a high activity *in vivo* may have a sequence close to the concensus but not necessarily. Hence one should be careful in predictions of *in vivo* activity of promoters based solely on similarity to predicted concensus sequences.

1.5.1.3.1 NTR (or NIF-like) Promoters

A global control by nitrogen regulation (ntr) exists in several bacterial genera. Genes subject to this control include the nitrogen fixation (nif) genes of *K.pneumoniae* and the nitrogen assimilation genes of *E.coli*, *Salmonella* and *K.pneumoniae*. They are governed by a complex regulatory circuit. The promoters of these genes are not homologous to the -35 and -10 sequences typical of *E.coli* promoter sequences, but possess a sequence, $CTGGCACN_5TTGCA$, centered around -24 and -12 (Beynon *et al* 1983, Buck 1986). This sequence is not recognised by $E\sigma^{70}$, but by an alternative σ factor, σ^{60} or NtrA. This alternative σ factor is not by itself sufficient to promote transcription, a positive activator is required, this is either NtrC (GlnG), a general nitrogen regulatory protein, or NifA specific to the nif genes. These two proteins can substitute for one another to varying degrees. The binding site of NifA lies considerably upstream of the promoter sequence (Buck *et al* 1986) and more than 100bp from the transcription start site. Unlike *E.coli* promoter sequences where the spacing between the -35 and -10 regions can vary by ±1bp without abolishing promoter activity there appears to be an absolute requirement

for a spacing of 10bp between the base at -24 and that at -12 for a nif promoter to be active (Buck 1986).

1.5.1.3.2 Sequences of Known *Pseudomonas* Promoters

Very few promoter sequences of *Pseudomonas* genes have so far been determined, but from these alone there is evidence of heterogeneity in the structure of *Pseudomonas* promoters. Since some *Pseudomonas* genes were found to be expressed well in *E.coli* it seemed likely that at least these *Pseudomonas* promoters possessed *E.coli*-like promoter sequences. Some of the first *Pseudomonas* promoters to be investigated were those which direct the expression of genes carried on the degradative plasmids of *P.putida*. The promoter sequences for the *xyl* genes of the TOL plasmid have been determined, as have those sequences of the *nah, sal* and *nahR* promoters. The organisation of the TOL and NAH/SAL genes are remarkably similar; they are plasmid-encoded, both possess two physically separate operons and all the operons are regulated in a positive manner by the product of another gene encoded close to one of the operons on the plasmid. However in spite of this, at the level of promoter nucleotide sequence, there is little homology. The *nah* and *sal* operons and the *nahR* gene all have *E.coli*-like promoter sequences (Schell 1986) and are well-expressed in *E.coli*, whereas the *xyl* promoters are more heterogeneous. It was found that the genes for degradation of toluene and xylenes were not well expressed in *E.coli* suggesting that these promoters had different features compared with those of *E.coli* promoters. Inouye *et al* (1984a) identified the promoter of the *xylCAB* operon and determined its nucleotide sequence. This promoter showed no homology with the canonical -35 and -10 sequences typical of *E.coli* promoter sequences, but has high homology to the nif concensus sequence. Mermod *et al* (1984) and Inouye *et al* (1984b) mapped the transcriptional start points of the *xylDLEGF* operon. Mermod *et al* (1984) reported that there were two transcription initiation start sites and two overlapping

promoter sequences. Inouye *et al* (1984b) reported only one transcription initiation site. Neither of these promoter regions bore homology to either the *E.coli*-like promoter sequence or to the nif-like sequence. The promoter sequences of the regulatory genes *xylS* and *xylR* have also been determined (Inouye *et al* 1985, Spooner *et al* 1986, Inouye *et al* 1987a). The *xylS* gene promoter, like that of the *xylCAB* promoter, showed considerable homology to the nif concensus sequence. This *xylS* gene was inefficiently expressed in *E.coli* (Spooner *et al* 1986) and this may be due in part to its promoter sequence. In addition the ribosome binding site may be poorly effective in *E.coli* due to its close proximity to the translation initiator codon. The *xylR* gene promoter has homology to the *E.coli* concensus and the levels of mRNA are approximately the same in both *E.coli* and *P.putida* and the mRNA start point is the same in both species (Inouye *et al* 1985). The genes for the degradation of 3-chlorocatechol are also encoded as a positively regulated operon on a conjugative plasmid (Frantz and Chakrabarty 1987). The promoter of this *clcABD* cluster has no apparent homology to either the *E.coli* or the nif concensus. Hence, even within the plasmid-borne genes there appears to be significant heterogeneity of *Pseudomonas* promoter sequences.

There has been a suggestion that the difference in the promoter structures of the TOL and NAH genes may be related to host range in so far as perhaps the NAH plasmid has a broader host range in terms of its replication properties due to the possession of *E.coli*-like promoter sequences. Indeed it appears that the broad host range plasmid RSF1010 possesses *E.coli*-like promoter signals presumably for the active expression of genes in a wide range of hosts. It is conceivable that, although the NAH plasmid cannot transfer itself (or be transferred) to a number of other species, perhaps in past evolutionary time it had the ability to do that, whereas possibly the TOL plasmid represents *Pseudomonas* genes that were once chromosomally-located, but which are now carried on a plasmid.

These plasmid-borne genes are highly regulated and evidence from *E.coli* indicates that positively regulated genes tend to have promoters which differ significantly from the concensus (McClure 1985). Therefore, although these provide important information on promoters which *P.putida* can recognise, they may not be analogous to constitutive chromosomally-located *P.putida* promoters. In an attempt to determine what elements are important in a promoter sequence for constitutive expression of *Pseudomonas* genes some mutant promoters of the TOL genes have been isolated and analysed. Mermod *et al* (1984) generated several mutants which gave rise to constitutive expression of the *xyl*DLEGF operon and the nucleotide sequences and mRNA initiation points of these were compared with the wild type promoter. The mutations were of three types: duplication of a specific sequence; a C to T transition at particular position and a C to T transition at an alternate position. The effect of these mutations was the alteration of the transcription initiation point compared with the wild type; presumably a different promoter was utilised. Since these mutant promoters gave rise to constitutive expression in both *E.coli* and *P.putida*, it appears that optimisation of the promoter regions for activity in both hosts had occurred not specifically for activity in *Pseudomonas*. The C to T transitions occurred at either +5 or -18 relative to the transcription initiating position (+1) in *P.putida*. Inouye *et al* (1986) isolated a promoter which exhibited high constitutive expression of *xyl*EGF in *P.putida* and when this region was cloned on a promoter-probe vector it gave rise to increased activity in *P.putida* compared with *E.coli*. This was apparently due to an increased synthesis of mRNA in *P.putida* compared with *E.coli*. These results suggested that a new transcription initiation site had been generated in the region upstream of the *xyl*E gene. Three transcription initiation points were identified in both *E.coli* and *P.putida*, one of these had homology to -35 and -10 regions of *E.coli* promoters, but these regions were displaced 5bp from their usual positions.

There were no regions homologous to *E.coli* promoter sequences corresponding to the other two transcription starts.

Although a large number of *Pseudomonas* genes have been cloned and sequenced, including the regions upstream of the structural gene for most, the promoter region has not been determined by the mapping of the transcriptional start point and therefore this limits any assessment of the likely promoters of these genes. *Pseudomonas* genes for which the promoter sequences are known are illustrated in Table 5. The importance of the determination of transcription starts in both *E.coli* and *Pseudomonas* for the localisation of the possible promoter is illustrated well by the finding of Itoh *et al* (1988). These workers found that there are two transcription starts for the *argF* gene in both *E.*coli and *P.aeruginosa*. One of these has similarity to the -35 and -10 concensus, but this is only poorly active in *E.coli*. However the other promoter, which has no homology to the *E.coli* or nif concensus is used in *P.aeruginosa* but not in *E.coli*.

Some *Pseudomonas* genes exhibit homology in their promoter regions to the *E.coli* -35 and -10 concensus elements. These include the *nah/sal/nahR* genes of the NAH plasmid (Schell 1986) and the *ompF* gene of *P.aeruginosa* (Duchene *et al* 1988). Generally it seems that those promoters which are highly homologous to *E.coli*-like promoters are expressed well in *E.coli*. Also *Pseudomonas* genes which are placed under control of an *E.coli* promoter are expressed to a high level in *Pseudomonas*. Two examples are one of the genes involved in 3-chlorocatechol degradation, *clcD*, from a *Pseudomonas* species (Frantz *et al* 1987) and the D-galactose dehydrogenase gene of *P.fluorescens* (Buckel and Zehelein 1981). Further information on the structure of some of the promoter sequences of *Pseudomonas* genes has come from the recognition of homology between them and the ntr-regulated promoters of the enteric bacteria. The promoter of the *xylCAB* operon has such a promoter and it has been shown to have a functional significance for the regulation of the genes. Dixon (1986) found that the operon is

TABLE 5

PSEUDOMONAS GENES FOR WHICH THE PROMOTER REGION HAS BEEN DETERMINED BY MAPPING OF THE TRANSCRIPTION START

exoA	*P.aeruginosa*	Grant and Vasil 1986
toxR	*P.aeruginosa*	Wozniak *et al* 1987
algD	*P.aeruginosa*	Deretic *et al* 1987
argF	*P.aeruginosa*	Itoh *et al* 1988
ompF	*P.aeruginosa*	Duchene *et al* 1988
PAK and PA103 pilins	*P.aeruginosa*	Johnson *et al* 1986
plc	*P.aeruginosa*	Grant and Vasil 1986
catBC	*P.putida*	Aldrich and Chakrabarty 1988
CPG2	*Pseudomonas sp.*	Minton and Clarke 1985
nah/sal/nahR	NAH plasmid	Schell 1986
xylCAB	TOL plasmid	Inouye *et al* 1984a
xylDLEGF	TOL plasmid	Mermod *et al* 1984, Inouye *et al* 1984b
xylR	TOL plasmid	Spooner *et al* 1986, Inouye *et al* 1987
xylS	TOL plasmid	Inouye *et al* 1985
clcABD	*Pseudomonas* plasmid	Frantz and Chakrabarty 1987

activated by the *ntr*A gene (which encodes σ^{60}) and *xyl*R in *E.coli* and the general nitrogen activator products of *nif*A and *ntr*C can substitute for XylR. Also *xyl*S has an ntr-like promoter and its activation in the presence of m-xylene depends on both NtrA and XylR in *E.coli* (Inouye *et al* 1987a).

Several other pseudomonad genes exhibit homology to the ntr-regulated promoters including the carboxypeptidase G2 gene of a *Pseudomonas* species (Minton and Clarke 1985) and the pilin genes of *P.aeruginosa* (Johnson *et al* 1986, Pasloske *et al* 1988). Activity of the CPG2 promoter is very much greater in *P.putida* than *E.coli* (Minton *et al* 1983). For these promoters it has not been determined whether they are dependent on NtrA for activation. Evidence from those promoters which are homologous to the nif promoters suggests that genes with such promoters are poorly expressed in *E.coli* compared with their level of recognition in *Pseudomonas*.

There is a third class of *Pseudomonas* promoters which exhibit neither homology to *E.coli* promoter sequences nor to the ntr-like promoters. These include *xyl*DLEGF of the *P.putida* TOL pathway, *arg*F of *P.aeruginosa* (Itoh *et al* 1988), *alg*D of *P.aeruginosa* (Deretic *et al* 1987c), exotoxinA of *P.aeruginosa* (Grant and Vasil 1986) and *tox*R of *P.aeruginosa* (Wozniak *et al* 1987). Because of this lack of homology to known promoter signals, and because both chromosomally and plasmid-located genes belong to this group, it can be postulated that they may represent a novel promoter signal unique to *Pseudomonas*. However although this group can be considered to consist of *Pseudomonas*-specific promoters, it also includes some promoters, for example *tox*A, which are positively activated and, when the positive activator is present will be functional in *E.coli*. These may, therefore, be more like typical positively regulated *E.coli* promoters whose promoter sequence is known to deviate considerably from the *E.coli* concensus -35 and -10 regions. It has been noted (Deretic *et al* 1987c) that *tox*A and *alg*D exhibit a high degree of homology in their promoter regions and this may be important in their mechanism of activation.

Some *Pseudomonas* genes which have been cloned and sequenced, but the transcription initiation points not determined, also show no homology to *E.coli* promoter sequences upstream of the open reading frame. These include the trans-zeatin producing gene (*ptz*) of a *P.syringae* plasmid (Powell and Morris 1986) and the indole acetic acid synthesis genes of *P.savastanoi* (Yamada *et al* 1985).

Although few promoter sequences of *Pseudomonas* genes have been defined, there have been suggestions of possible *Pseudomonas* concensus promoter sequences. Mermod *et al* (1984) identified two overlapping promoters that initiated transcription of the *xylD*LEGF operon of the TOL pathway. They compared these regions with the *xylC*AB promoter region and with two constitutive *xylD*LEGF promoter mutants and proposed a concensus sequence of A - A G G C - T N_{7-12} G C T/A A T A N_{3-7} A(transcriptional start). This concensus has little significance for several reasons;

(a) only 5 sequences were used for its derivation

(b) the comparison was between both positively regulated and constitutive mutant promoters for which one would not expect a high degree of homology

(c) in the light of the finding that the *xylC*AB operon exhibits a nif concensus sequence and that this is the functional promoter, and since the *xylD*LEGF promoter does not have homology to a nif sequence, the two sequences are not strictly comparable. Hence to use this concensus as being representative of an archetypal *Pseudomonas* promoter, as has been done for some cloned *Pseudomonas* genes (Irie *et al* 1987b), is meaningless.

Of perhaps more validity is the concensus promoter sequence proposed by Inouye *et al* (1986). This is based on a comparison of 8 promoters from *P.putida* plasmids which are subject to constitutive action. These include 3 promoters from a mutant promoter region upstream of the *xylE* gene (Inouye *et al* 1986), the 2 promoters recognised in *xylR* (Inouye *et al* 1985), the 2 constitutive mutants of *xylD*LEGF (Mermod *et al* 1984) and the *nahR* gene promoter from the NAH7 plasmid (Schell 1986). This concensus has the

sequence AA--AAATGGTAAATAT with the T at the extreme 3' end occurring at 4 to 8 base pairs upstream of the transcriptional start. This sequence was not found in the promoter regions of inducible operons from *P.putida* plasmids, for example *xylCAB*, *xylDLEGF*, *nah/sal*.

Aldrich and Chakrabarty (1988) have made a comparison of the promoters of positively regulated *Pseudomonas* genes including *cat*BC, *nah*A, *nah*G, *xylCAB*, *xylDLEGF*, *clc*ABD, *algD* and *tox*A. Although no concensus was proposed homologies were found to exist in the region from -7 to -12 and there was a sequence ATCC which was present in the -10 regions of the *algD*, *cat*BC and *tox*A chromosomal promoter regions.

At the present time there is insufficient data to determine the promoter sequences that can be recognised by the RNA polymerase of *Pseudomonas* in addition to *E.coli*-like and nif-like sequences.

1.5.1.4 Termination Signals

A further way in which transcription could differ between prokaryotic species at the level of DNA primary structure in addition to promoter recognition is differences in the sequences which specify the termination of transcription and cause the RNA polymerase to release the nascent mRNA and dissociate it from the DNA strand. Such terminators occur in *E.coli* at the end of an operon or gene, before the start of the structural gene giving rise to attenuation important in the controlled regulation of biosynthetic genes, and also within an operon giving rise to the phenomenon of polarity. The presence of these structures therefore plays an important role in gene expression in *E.coli*. The molecular mechanism of transcription termination has not yet been fully elucidated, although some evidence has come from the discovery of certain sequences which cause transcription termination. Two classes of transcription termination signals have been discovered in *E.coli*; rho-independent and rho-dependent terminators. *E.coli* rho-independent terminators have a region of GC-rich dyad symmetry followed

by a run of A's in the DNA. There is some circumstantial evidence from nucleotide sequence data that rho-independent terminators exist in *Pseudomonas* and related species, since similar structures have been found 3' to the termination codon of genes such as *clcD* of *Pseudomonas* strain B13 (Frantz *et al* 1987), *argF* of *P.aeruginosa* (Itoh *et al* 1988), *ompF* of *P.aeruginosa* (Duchene *et al* 1988), cephalosporin acylase of a *Pseudomonas* strain (Matsuda *et al* 1987), *aerC* and *aerA* of *Aeromonas sobria* (Husslein *et al* 1988), *Caulobacter* 28.5kD flagellin gene (Gill and Agabian 1983), and *gap* of *Zymomonas mobilis* (Conway *et al* 1987b). It is possible that *Pseudomonas* may possess a unique signal for transcription termination which may modify gene expression in the heterologous host.

1.5.2 Factors affecting Translation

1.5.2.1 Ribosome Binding Site

The ribosome binding site in *E.coli* is a particular sequence which is complementary to the 3' end of 16s rRNA (Shine and Dalgarno 1974). Shine and Dalgarno (1975) sequenced the 3' terminal regions of 16s rRNA from several species, both Gram-negative and Gram-positive and found a remarkable homology between them. The species examined were *E.coli, P.aeruginosa, C.crescentus, B.subtilis* and *B.stearothermophilus.* The possibility that a barrier to expression is due to differences in the ribosome binding site and its recognition from different species seems unlikely in view of this strong sequence conservation.

Evidence that has come from the sequencing of *Pseudomonas* genes indicates that their putative ribosome binding sites tend to have as much homology to the predicted ribosome binding site of *P.aeruginosa* as that of *E.coli* and hence would be expected to be efficient in both species (Table 6).

TABLE 6

RIBOSOME BINDING SITES OF *PSEUDOMONAS* GENES

Ribosome binding site of *P.aeruginosa* (Shine and Dalgarno 1975) AAGGAGAG

Ribosome binding site of *E.coli* (Shine and Dalgarno 1975) GAAGGAGG

PSEUDOMONAS GENE	RBS	REFERENCE
catB P.aeruginosa	GGA	Aldrich and Chakrabarty 1988
catC P.aeruginosa	AAGGAGA	" " " "
ompF P.aeruginosa	GGGA	Duchene *et al* 1988
plc P.aeruginosa	GGGA	Pritchard and Vasil 1986
plcR P.aeruginosa	GAGGAG	Shen *et al* 1987
argF P.aeruginosa	AAGG	Itoh *et al* 1988
PAK pilin *P.aeruginosa*	GGAGA	Johnson *et al* 1986
PA103 pilin *P.aeruginosa*	GGAGA	" " " "
PAO1 pilin *P.aeruginosa*	GGAGA	" " " "
toxA P.aeruginosa	AGGAG	Grant and Vasil 1986
toxR P.aeruginosa	AGAG	Wozniak *et al* 1987
amiE P.aeruginosa	AGGAGG	Brammar *et al* 1987
algD P.aeruginosa	AGGTGA	Deretic *et al* 1987c
ORF1 *P.putida*	GTGAGA	Irie *et al* 1987
ORF2 *P.putida*	AAGGAA	" " " "
ORF3 *P.putida*	GGTG	" " " "
ORF5 *P.putida*	GAGGA	" " " "
inaW P.fluorescens	AAGGG	Warren *et al* 1986
bphC P.pseudoalcaligenes	AAGGAGA	Furukawa *et al* 1987
bphD P.pseudoalcaligenes	AAGAAGA	" " " "
clcD Pseudomonas sp.	GGAGAGA	Frantz *et al* 1987

Table 6 ctd.

acyI	*Pseudomonas sp.*	GAGG	Matsuda *et al* 1987
acyII	*Pseudomonas sp.*	GAGG	" " " "
CPG2	*Pseudomonas sp.*	AGGAGA	Minton and Clarke 1985
ORFA	*P.syringae*	TGGA	Mellano and Cooksey 1988
ORFB	*P.syringae*	GAGG	" " " "
ORFC	*P.syringae*	AAGGAG	" " " "
ptz	*P.savastanoi*	AAGGG	Warren *et al* 1986

1.5.2.2 mRNA Decay

Much effort has been directed towards the study of gene expression as a function of mRNA synthesis and it is probably true that it is at the transcriptional level that most genes are regulated. However for this reason little is known of how the degradation of message and its post-transcriptional modification can affect its translation. Prokaryotic mRNA molecules are rapidly degraded, they have an average half-life of 1 to 2 minutes. This, together with the fact that in prokaryotes transcription and translation are tightly coupled, creates problems in the investigation of this area. However some evidence now exists for mRNA affecting its own translation. Higgins *et al* (1982) discovered the presence of REP (repetitive extragenomic palindromic) sequences in several intergenic regions of operons of *E.coli* and *Salmonella typhimurium*. These consist of inverted repeats which have the capacity to form stem loops of up to 35bp. Subsequently it has been shown (Stern *et al* 1984) that these sequences are highly conserved and very numerous with up to 500 copies on the *E.coli* and *S.typhimurium* chromosomes. REP sequences have been found in intercistronic regions of operons or at the 3' untranslated end of transcripts upstream of the terminator. Newbury *et al* (1987a) have shown that these REP sequences have the effect of stabilising upstream mRNA *in vivo* by preventing degradation of RNA by 3'-5'exonuclease activity by RNaseII and polynucleotide phosphorylase (these are the enzymes which cause bulk degradation of mRNA). It is thought that the secondary structure creates a block for the processive exonucleases. Newbury *et al* (1987b) have further shown that the differential gene expression of the malEFG operon of *E.coli* is partly due to the REP-dependent stabilisation of upstream mRNA. malE mRNA is significantly more stable than full-length malEFG mRNA. The presence of ribosomes on a mRNA molecule may stabilise it by protection from endonucleases, but this is probably just a passive effect.

There is no evidence for the existence of similar features in *Pseudo-*

monas genes. Differential stability of mRNA molecules or of proteins synthesised in *E.coli* and *Pseudomonas* could be expected to influence the relative levels of gene expression in these two hosts.

1.5.2.3 Codon Usage

There is a high correlation between codon usage, tRNA abundance and the level of gene expression in *E.coli*. This has led to the idea that codon usage could be a modulator of gene expression (Grosjean and Fiers 1982). This hypothesis has been tested by Holm (1986) by comparing the distribution of codons and tRNA's to the kinetics of translational elongation. She concludes that for *E.coli* codon usage is modulated for expression, since highly transcribed genes tend to possess codons which are most well represented in the tRNA pool, but gene expression is not modulated by codon usage. Differences in gene expression of *E.coli* genes is not a result of variable codon usage. It is possible that the choice of codons utilised is to reduce mistranslation.

It has been reported that some *Pseudomonas* genes exhibit an unusual codon usage with a preference for G/C in the third position of codons. This has been noted for various genes, for instance the *ami*E gene of *P.aeruginosa* (Brammar *et al* 1987), *xyl*E gene of the TOL pathway (Nakai *et al* 1983), the mercuric reductase gene (*mer*A) from the *Pseudomonas* transposon Tn501 (Brown *et al* 1983) isolated from *P.aeruginosa* (since the GC composition of the gene is 65% and similar to that of the *P.aeruginosa* chromosome, this gene is considered to be a true *Pseudomonas* gene) and the *alg*D gene of *P.aeruginosa* (Deretic *et al* 1987c). In view of the high GC content of *P.aeruginosa* in particular it perhaps is not surprising that there is a predominance of G/C in the third position of codons. However the evidence of Holm (1986) suggests that this should not account for a dramatic difference in the translation of genes between *E.coli* and *P.putida*.

1.5.2.4 Translational Specificity of Bacterial Ribosomes

There is some evidence (Stallcup and Rabinowitz 1973, Lodish 1970) that certain species of bacteria exhibit specificity of translation of mRNA from other species. This has been observed in comparisons of *E.coli*, *Bacillus* and *Clostridium*. An examination of how *P.fluorescens* behaves in this regard showed that like *E.coli*, the ribosomes of *P.fluorescens* could translate mRNA from 9 different sources: f2 RNA, T4 early, *E.coli*, *P.fluorescens*, *Azotobacter vinelandii*, *C.pasteurianum*, *Streptococcus faecalis*, *B.subtilis* and *Peptococcus aerogenes*. These results indicate that Gram negative species will translate mRNA from both Gram negative and Gram positive organisms, but Gram positive bacteria will only translate mRNA from Gram positive organisms. More recent experiments on the structure of Gram positive and negative ribosomes have revealed that antibodies directed against S1 protein from *E.coli* cross-react with a high molecular weight protein of *Proteus vulgaris*, *Aeromonas hydrophila* and *P.aeruginosa* (Hahn 1986). This protein is of approximately the same size as the S1 from *E.coli*. There was no crossreaction with ribosomal proteins from *B.stearothermophilus*, *B.subtilis*, *Lactobacillus casei* or *S.faecalis*. This suggests that other Gram negative species are structurally similar to *E.coli* and this would account for the observed non-specificity of translation between them. However other workers (Muralikrishna and Suryanarayana 1987) have found that *P.fluorescens* does not have an S1-like protein nor does *Acetobacter pasteurianum*, whereas *Micrococcus luteus* does and this protein is approximately the same molecular weight as S1.

1.5.3 Summary of Transcriptional and Translational Problems associated with Cloned *Pseudomonas* Genes in the Heterologous Host

Many factors could potentially influence the level of gene expression in the heterologous host since it is a highly complex process involving the interplay of many specific components in a strictly regulated manner. Of

the two major processes involved in the expression of a gene, regulation at the transcriptional level plays a far greater role in the determination of gene expression than do translational events. There is little evidence of a differential effect on RNA translation in *E.coli* and *Pseudomonas*. Some evidence points to the contrary, in that mRNA levels from specific *Pseudomonas* genes in the two hosts tend to exhibit a large variation suggesting that the problem lies at the transcriptional point (Inouye *et al* 1984a, Inouye *et al* 1984b). Transcription has a great effect on whether a gene will give rise to its protein product or not. Evidence from cloning *Pseudomonas* genes, which are poorly expressed in *E.coli*, under control of *E.coli* promoter signals giving rise to increased expression in *E.coli* supports this view (Buckel and Zehelein 1981, Frantz *et al* 1988). The major evidence for differential transcriptional effects in the two hosts comes from the absence of recognisable *E.coli* or *nif* promoter sequences for both positively activated and repressible genes. Also the recent evidence that *E.coli* possesses alternate sigma factors creates the possibility that *Pseudomonas* may possess another sigma factor not present in *E.coli* and that this alternate sigma factor allows recognition of sequences found 5' to some *Pseudomonas* genes. The structure of such promoter sequences has not yet been elucidated, possibly because of the lack of sufficient experimental data, particularly the lack of transcription initiation data from many *Pseudomonas* genes so far cloned. The core RNA polymerase molecules of both *E.coli* and *Pseudomonas* are highly similar. There may be some difference in the α subunit as judged by molecular weight comparisons, but, since *E.coli* genes can be expressed in *Pseudomonas* and the DNA-binding centres of the RNA polymerases from *P.putida* and *E.coli* form identical contacts with *E.coli* promoters, this suggests that the core RNA polymerase enzyme of *E.coli* is not limited by its ability to transcribe *Pseudomonas* genes. Another strong possibility is that there may be termination signals specific to *Pseudomonas*, although there is, as yet, no evidence for this,

indeed, the mechanisms of transcription termination in the well-characterised species *E.coli* are still poorly understood.

1.6 PROJECT OUTLINE AND AIMS

There were two aspects to the project. Firstly to generate an *in vitro* coupled transcription/translation system based on a *Pseudomonas* species, and secondly a study of promoter signals in *Pseudomonas*. The first involved the adaptation of the published method for a coupled transcription/translation system of *E.coli* for the production of an active system for *Pseudomonas*. The second involved the isolation and characterisation of nucleotide sequences of fragments of DNA from *P.putida* which exhibited high promoter activity in *P.putida* but low activity in *E.coli* with a view to determining whether there is some thing peculiar to the sequence that inhibits recognition by *E.coli* RNA polymerase, but allows efficient recognition by *Pseudomonas* RNA polymerase. Although regulatory control by positive activation seems to be a usual method for *Pseudomonas* gene systems (with the ever-increasing numbers of *Pseudomonas* genes being studied), it is likely that not all biosynthetic genes in the group will be regulated. Many of the so-called "housekeeping" genes are likely to be constitutively expressed, unlike the specific and unusual activities of *Pseudomonas* species that are under current investigation. Hence, due to the dearth of information concerning *Pseudomonas* chromosomal gene promoters for constitutively-expressed genes and the possibility that these may contain a novel class of promoter signal peculiar to the pseudomonads, the aim was to isolate chromosomal promoter sequences which have high activity in *Pseudomonas*, but low activity in *E.coli* and to analyse these in terms of a possible concensus sequence for such *Pseudomonas* promoters.

P.putida was the chosen species for both the generation of an *in vitro* coupled transcription/translation system and the promoter study since it is the *Pseudomonas* species which exhibits the widest nutritional versatility -

and hence excites the most interest for the commercial harnessing of its unique properties.

CHAPTER 2

MATERIALS AND METHODS

2.1 BACTERIAL STRAINS AND PLASMIDS

Bacterial strains and plasmids used in this work are shown in Table 7. Strains were generally grown in Luria Broth and maintained on Luria agar plates, stored at 4°C and routinely subcultured onto fresh plates. For maintenance for periods of months strains were stored in 15% glycerol in Luria broth at -20°C. JM101 was maintained on M9 minimal medium for retention of the F' plasmid. Plasmids were stored as DNA preparations in sterile water at -20°C.

2.2 GROWTH MEDIA

Media used for growth of bacteria is shown in Table 8. For L-agar plates Bacto agar was added to 1.5%. For M-9 minimal agar Bacto agar was added to 1.5% and for M-9 minimal top agar Bacto agar was added to 0.75%. Brain Heart Infusion Broth and Isosensitest agar (Oxoid) were made up according to manufacturer's instructions. All media were made up with distilled water, autoclaved and stored at room temperature.

Antibiotics were used at the following concentrations; for *E.coli:* Ap 100μgml^{-1}, Km 50μgml^{-1}, Sm 25μgml^{-1}, for *P.putida*: Cb 2mgml^{-1}. Tp was incorporated into the Isosensitest agar at a concentration of 100μgml^{-1}.

2.3 TRANSFORMATION OF *E.COLI*

2.3.1 Preparation of Competent Cells

A 1:200 dilution of a stationary phase overnight culture of *E.coli* was made into 10ml of L-broth and grown up at 37°C to mid-log phase (OD_{600nm} of 0.55). The cells were pelleted at 3krpm in a bench centrifuge for 5 minutes at 4°C, washed in 4ml 10mM NaCl at 4°C, then resuspended in 4ml ice-cold 100mM CaCl and placed on ice for 20 minutes. The cells were then spun down at 2.5krpm for 5 minutes at 4°C and resuspended in 1ml of 100mM CaCl and the competent cells stored on ice until use.

TABLE 7

BACTERIAL STRAINS

E.COLI		
MRE600	RNase 1$^-$	Pratt *et al* 1981
LE392	*F*$^-$, *hsdR514*(r_k^-, m_k^-), *supF*$_{58}$ *lacY*$_1$ or Δ*lacIZY*, *galK*$_2$, *galT*$_{22}$, *metB*$_1$, *trpR*$_{55}$, λ$^-$	Maniatis *et al* 1982
MC1061	*araD139*, Δ(*ara-leu*)*7697*, Δ*lacX74*, *galU*, *galK*, *hsdR*, *hsdM*$^+$, *rpsL*	Casadaban and Cohen 1980
DS410	*minA*, *minB*, *rpsL*, *sup*$^+$	Dougan and Sherratt 1977
JM83	*ara*, Δ(*lac-proAB*), *rpsL*, Φ80, *lacZ*Δ*M15*	Vieira and Messing 1982
JM101	*supE*, *thi*, Δ(*lac-proAB*), *F*$^-$, *traD*$_{36}$, *proAB*, *lacIqZ*Δ*M15*	Yanisch-Perron *et al* 1985
P.PUTIDA		
mt-2 KT2440	*hsdR*$^-$, *hsdM*$^+$	Murray *et al* 1972
P.AERUGINOSA		
PAO1	wild type	Holloway 1969
PAO1161	*hsdR*$^-$, *leu*	GS Sharpe, Personal Comm.

Table 7 ctd.

PLASMIDS

pAT153	Derivative of pBR322 Ap^r,Tc^r	Twigg and Sherratt 1980
pCF32	*xylE* cloned in RSF1010-based vector Km^r,Sm^r,Cb^r	Spooner *et al* 1987
pGB123	Capsule biosynthesis genes from *E.coli* K1 cloned in pUC8	IS Roberts and GJ Boulnois, Personal Comm.
pGSS15	R300B-based broad host range cloning vector Tc^r,Ap^r	Sharpe 1984
pGSS33	R300B-based broad host range cloning vector Tc^r,Ap^r,Cm^r,Sm^r	Sharpe 1984
pJW252	Pnuemolysin gene from *S.pneumoniae* cloned in pUC8	Walker *et al* 1987
pKK223.3	*tac* promoter expression vector Ap^r	Brosius and Holy 1984
pKT231	RSF1010-based broad host range cloning vector Km^r,Sm^r	Bagdasarian *et al* 1981
pLG221	ColIb-based transposon donor plasmid Km^r	Boulnois *et al* 1985
pRLA5	Broad host range promoter-probe vector Cb^r	This work
R300B	IncQ broad host range plasmid Sm/Su^r	Barth and Grinter 1974
RSF1010	IncQ broad host range plasmid Sm/Su^r	Guerry *et al* 1974
RSF1010Δ18	Deletion derivative Sm^r	Bagdsasarian *et al* 1982
RSF1010Δ20	Deletion derivative Sm^r	" " "
pKT228	RSF1010::Tn3 Sm^r,Ap^r	" " "
pKT229	RSF1010::Tn3 Sm^r,Ap^r	" " "

TABLE 8

GROWTH MEDIA

MRE600 Medium

	per litre
KH_2PO_4	5.6g
K_2HPO_4	28.9g
Magnesium Acetate	1mM
Yeast Extract	1g
Vit. B1	1mg
Arginine	0.05g
Methionine	0.05g
Leucine	0.05g
Histidine	0.05g

10x M-9 Salts

	per litre
Na_2HPO_4	60g
KH_2PO_4	30g
NaCl	5g
NH_4Cl	10g

M-9 Minimal Medium

	per litre
10x M-9 salts	100ml
$MgSO_4.7H_2O$	1mM
20% Glucose	10ml
1% Vit. B1	1ml
10mM $CaCl_2$	10ml

2.3.2 Transformation with plasmid DNA

Plasmid DNA was added to 100μl of competent cells, held on ice for one hour and then heat shocked at 42°C for 2 minutes. 800μl L-broth added and the tubes incubated at 37°C for 1-1.5 hrs, then plated onto appropriate selective plates and incubated overnight at 37°C. A plasmid negative control was routinely employed to check for contamination.

2.3.3 Transformation with Bacteriophage M13

Competent JM101 cells were transformed with M13 DNA by incubating 150μl cells with the DNA on ice for one hour, then heat shocking at 42°C for 2 minutes. 150μl of a mid-log phase culture of JM101 cells were added to this and the whole added to 3ml M-9 minimal top agar (held at 45°C) containing 30μl 100mM IPTG and 30μl 2% X-Gal in Dimethylformamide and immediately poured onto M-9 minimal agar plates. When the top agar was set the plates were incubated overnight at 37°C.

2.4 TRANSFORMATION OF *P.PUTIDA* (Bagdasarian and Timmis 1982)

A 1:200 dilution of a stationary phase culture was made into 10ml L-broth (with appropriate antibiotics) and grown up to mid-log phase (OD_{600nm} 0.55). The cells were pelleted at 3krpm in a bench centrifuge for 5 minutes at 4°C and washed in 10ml of Buffer I at 4°C, then resuspended in 10ml Buffer II and held on ice for 30 minutes. The cells were then spun down at 2.5krpm for 5 minutes at 4°C and resuspended in 1ml of Buffer II. DNA was added to 200μl cells and placed on ice for 1hr, then heatshocked at 42°C for 2 minutes and 1ml L-broth was added and incubated at 30°C for 1.5 hrs then plated onto selective media.

Buffer I: 10mM MOPS pH 7.0

10mM RbCl

100mM $MgCl_2$

Buffer II: 100mM MOPS pH 6.5

10mM RbCl

100mM $CaCl_2$

2.5 MOBILISATION OF BROAD HOST RANGE PLASMIDS INTO *P.PUTIDA* (Franklin 1985)

10ml cultures of *P.putida* 2440 and of the *E.coli* strain JM83 containing the helper plasmid pLG221 and the plasmid to be mobilised were grown up to mid log phase in L-Broth (+ appropriate antibiotics for maintenance of the two plasmids) from a 1:200 dilution of an overnight culture. 1.5ml of each were spun down together in a microfuge, resuspended in 50µl of L-Broth and spread onto sterile Whatman nitrocellulose filters (0.45µm) which had been placed on the surface of non-selective L-agar plates. Mating was allowed to proceed for ≈5 hours at 37°C, then the cells were washed off the filters with L-broth, spun down and resuspended in 1ml of L-broth and plated at an appropriate concentration onto selective agar plates and incubated for 36 hours at 30°C.

2.6 CHROMOSOMAL DNA PREPARATION FROM *P.PUTIDA*

10ml overnight culture was pelleted at 3krpm for 5 minutes at 4°C in a bench centrifuge. The cells were resupended in 3ml of 50mM Tris HCl pH 8/ 20% sucrose and 200µl of 200mM EDTA pH 8 added, followed by 1ml of $10mgml^{-1}$ of freshly prepared lysozyme. This was allowed to stand for 5 minutes at room temperature. 1ml of 10% SDS was added and the mixture allowed to stand until the lysate was totally clear, this usually occurred instantaneously. It was then extracted with phenol/chloroform mixture about 10 times until the upper layer was totally clear, spinning at 5krpm for 10 minutes at room temperature after each extraction. The resultant DNA was then ethanol-precipitated by adding 10% sodium acetate and 2 vols ethanol down the side of the glass universal and looping out the DNA which

formed at the interphase with a glass hook. This was then dissolved in 1ml of water and subsequently purified by caesium chloride/ethidium bromide density gradient centrifugation.

Phenol/Chloroform

100g phenol were dissolved in 100ml chloroform + 0.1% 8-hydroxyquinolene, saturated with 0.2M Tris HCl pH 8.0 and stored at 4°C.

2.7 LARGE-SCALE PLASMID PREPARATION (Birnboim and Doly 1979)

For preparation of broad host range plasmids 2x500ml overnight cultures were employed and the DNA prepared separately for both, then pooled onto one caesium chloride gradient. For other plasmids one 500ml culture was sufficient.

The bacterial pellet from each 500ml culture was resuspended in 10ml of Solution I containing 5mgml^{-1} lysozyme (freshly added to autoclaved and stored Solution I). This mixture was transferred to a Sorvall 35ml tube and stood at room temperature for 5 minutes. 20ml of freshly prepared Solution II were added and the contents mixed by gentle inversion of the capped tube. This was stood on ice for 10 minutes. 15ml of an ice-cold solution of 5M potassium acetate (pH 4.8) were added and mixed by inverting the capped tube sharply several times. This was stood on ice for 10 minutes. The mixture was then centrifuged at 18krpm in a Sorvall SS-34 rotor for 30 minutes at 4°C. Equal quantities of the supernatant were transferred to two siliconised Corex tubes and 0.6vols of isopropanol added, mixed well and allowed to stand at room temperature for 15 minutes. The DNA was recovered by centrifugation at 6krpm for 20 minutes at 20°C. The supernatant was discarded and the DNA pellet dried in a vacuum desiccator.

The dry pellets were dissolved in 12.5% sucrose/25mM Tris HCl pH 8 and made up to 16ml per gradient. 16g caesium chloride were dissolved in this and the solution transferred to a 35ml polyallomer ultracentrifuge tube and

1ml 10mgml^{-1} ethidium bromide added and mixed. The solution was overlaid with paraffin oil and the tube capped and crimped. Gradients were spun at 40krpm at 20°C for 20 hours, then the plasmid bands removed using a guage 18 needle and syringe inserted into the side of the tube. Ethidium bromide was removed from the DNA solution by several extractions with isopropanol equilibrated with a saturated solution of caesium chloride. The resultant solution was dialysed for several hours at room temperature against several changes of water. The DNA was precipitated at -20°C with ethanol, then recovered by centrifugation at 6krpm 20 minutes at room temperature and resuspended in sterile nanopure water and stored at -20°C.

Solution I: 50mM glucose

25mM Tris HCl pH 8.0

10mM EDTA pH 8.0

This was made up in distilled water, autoclaved and stored at room temperature

Solution II: 0.2M NaOH

1% SDS

This was prepared freshly from sterile stock solutions of 10M NaOH and 10% SDS.

2.8 MINIPREPARATION OF PLASMID DNA

1.5ml of media containing the appropriate antibiotics was inoculated with cells harbouring the plasmid and grown overnight. This overnight culture was pelleted for 1 minute in a microfuge. The supernatant was removed by aspiration, then the pellet was resuspended in 100μl of an ice-cold solution of lysis buffer (Solution I) containing 4 mgml^{-1} lysozyme freshly added. The mixture was stood at room temperature for 5 minutes, then 200μl of a freshly-prepared ice-cold solution of 0.2M NaOH/1% SDS added and mixed by gently inverting the tube two or three times. This was stood on ice for 5 minutes. 150μl of ice-cold 5M potassium acetate pH

4.8 were added and vortexed gently in an inverted position for 10 seconds then stored on ice for 5 minutes. The mixture was then centrifuged for 5 minutes and the supernatant carefully removed with a drawn-out pasteur pipette into a fresh tube and an equal volume of phenol/chloroform added. The mixture was vortexed, then centrifuged for 2 minutes and the supernatant transferred to a fresh tube and 2 vols ethanol added, mixed and placed at -20°C for 20 minutes. The DNA was centrifuged for 5 minutes then washed in 500µl 70% ethanol, the supernatant removed and the pellet dried briefly in a vacuum desiccator. The DNA was resuspended in 30µl sterile nanopure water and stored at -20°C.

2.9 ETHANOL PRECIPITATION OF DNA

DNA was precipitated by the addition of 3M sodium acetate to a final concentration of 0.3M and 2 vols of ethanol, followed by incubation at -20°C for 30 minutes. The DNA was centrifuged in a microfuge, the supernatant removed by aspiration using a drawn-out Pasteur pipette and dried by dessication using a vacuum desiccator. The resultant pellet was resuspended in sterile water and stored at -20°C.

2.10 DIGESTION OF DNA

For restriction analysis 0.5µg plasmid DNA was digested with 5 units of restriction endonuclease at 37°C (or optimum temperature for digestion as indicated by manufacturer) for 1-2 hours in 20µl of the appropriate buffer. This was analysed by running the samples with 10% agarose gel loading buffer on a 1% agarose gel in TBE buffer (containing $0.5mgl^{-1}$ ethidium bromide) against TBE buffer (containing $0.5mgl^{-1}$ ethidium bromide). Visualisation of DNA bands was carried out on a longwave uv transilluminator. Restriction endonucleases were purchased from Bethesda Research Laboratories (BRL).

Partial digests were carried out by digesting samples of DNA for

varying time periods then terminating the reactions by the addition of EDTA pH 8 to a final concentration of 10mM and placing the reactions on ice. The buffers used for the restriction enzyme digestions are shown in Table 9. They were made up at 10x concentration and stored at -20°C. Digestion with *Bgl*II, *Eco*RI, *Hin*dIII and *Xho*I was carried out in Core buffer. Isolation of DNA fragments from agarose gels was carried out by electroelution followed by ethanol-precipitation and resuspension in sterile water.

Agarose Gel Loading Buffer: 50% Glycerol
5x TBE
$1 mg ml^{-1}$ bromophenol blue

2.11 DEPHOSPHORYLATION OF LINEAR DNA MOLECULES

Linearised DNA was ethanol precipitated and resuspended in 50µl 1x CIP buffer. 20 units of calf intestinal phosphatase (Pharmacia) were added and incubated at 37°C for one hour, then extracted once with phenol/chloroform and several times with water-saturated diethyl ether. The ether was removed by incubation at 70°C for about 10 minutes; the DNA was ethanol-precipitated, resuspended in water and stored at -20°C.

1x CIP Buffer: 50mM Tris HCl pH 9.0
1mM $MgCl_2$
0.1mM $ZnCl_2$
10mM spermidine

2.12 LIGATION OF DNA MOLECULES

For the ligation of DNA fragments to vector, 1 unit of T4 ligase was added to the ligation reaction and incubated overnight at 14°C in a total volume of 20µl.

TABLE 9

RESTRICTION ENZYME ASSAY BUFFERS

10x Core buffer

500mM Tris HCl pH 8

100mM $MgCl_2$

500mM NaCl

10x SalI buffer

80mM Tris HCl pH 7.6

1.5M NaCl

60mM $MgCl_2$

2mM Na_2EDTA

10x AccI buffer

500mM Tris HCl pH 8

100mM $MgCl_2$

10x SmaI buffer

50mM Tris HCl pH 8

150mM KCl

60mM $MgCl_2$

Assayed at 30°C

10x ClaI buffer

100mM Tris HCl pH 8

100mM $MgCl_2$

$1mgml^{-1}$ BSA

10x SphI buffer

500mM Tris HCl pH 7.5

500mM NaCl

60mM $MgCl_2$

60mM 2-mercaptoethanol

10x KpnI buffer

60mM Tris HCl pH 7.5

60mM NaCl

60mM $MgCl_2$

1x Ligation Buffer: 0.5M Tris HCl pH 7.5

0.1M $MgCl_2$

0.1M DTT

1mM ATP pH 7.0

2.13 GENE EXPRESSION SYSTEMS

2.13.1 *E.coli* Minicell Procedure

2.13.1.1 Preparation of Minicells from *E.coli* DS410 (Meagher *et al* 1977)

500ml of an overnight culture of DS410 containing the appropriate plasmid grown in Brain Heart Infusion (BHI) with the required antibiotics were pelleted at 2krpm at 4°C for 5 minutes in a Sorvall GSA rotor. The minicell-containing supernatant was then spun at 8krpm for 15 minutes at 4°C in a Sorvall GSA rotor to pellet the minicells. The pellet was resuspended in M9 medium, 3ml of M9 to each 250ml of starting culture. Using a pasteur pipette 3ml of cell suspension were carefully layered over a 20ml sucrose gradient, which had been formed by placing a solution of 20% sucrose in M9 in a glass universal at -20°C until frozen then thawed overnight at 4°C prior to use the next day. The gradient was spun at 5krpm for 20 minutes at 4°C in a bench centrifuge with a swing out rotor. The upper fraction of minicells was removed with a pasteur pipette and the minicells pelleted by spinning at 10krpm for 10 minutes at 4°C. The minicells were resuspended as before and layered onto fresh gradients and respun. The upper minicell fraction was removed, the OD at 600nm determined and the presence of any contaminating whole cells was checked under the microscope. The minicells were pelleted as before then resuspended in M9 to give a final OD_{600nm} of 2. Minicells if not required immediately can be stored for several months in M9+30% glycerol at -20°C in 100µl aliquots then before use the minicells should be harvested and resuspended in M9 medium.

2.13.1.2 Detection of polypeptides synthesised by minicells harboring plasmids

100μl aliquots of prepared minicells were centrifuged for 2 minutes to pellet the cells which were then resuspended in 200μl M9 and recentrifuged. They were then resuspended in 200μl M9/0.4% glucose/Difco methionine assay media (0.525g MAA in 5ml distilled water; for every ml of M9/0.4% glucose 15μl MAA was added). D-cycloserine was added to a final concentration of 200$\mu g ml^{-1}$ and incubated at 37°C for 90 minutes. The cells were then pelleted in a minifuge for 2 minutes, resuspended in 200μl of prewarmed M9/MAA/Glucose broth and 50μCi ^{35}S-methionine added. They were incubated at 37°C for 45 minutes then pelleted, resuspended in 200μl M9/glucose + 200$\mu g ml^{-1}$ "cold" methionine and incubated at 37°C for 15 minutes. They were then pelleted, resuspended in 50μl SDS-PAGE loading buffer, boiled for 10 minutes and analysed by SDS-PAGE.

2.13.2 Coupled Transcription/Translation (Zubay) Procedure

(Based on methods by Collins (1979) and Pratt *et al* (1981))

All glassware used was thoroughly sterilised and all solutions used for the preparation of the S30 extract and for the assay of coupled transcription/translation were prepared with sterile Diethyl pyrocarbonate (DEPC)-treated water (except the growth media) for the removal of contaminating RNase.

2.13.2.1 Growth of Cells (for both MRE600 and 2440)

5 litres of medium (either MRE600 medium or L-both) were inoculated to an OD_{450nm} of 0.07 from a 100ml overnight culture of the bacterium, then grown to an OD_{450nm} of 2-3. The cells were harvested at 4°C by pelleting at 5krpm in a Sorvall GSA rotor for 10 minutes. The cell pellets were washed 3 times in 250ml S30 buffer + 7.2mM mercaptoethanol per 10g cells at 10k rpm for 20 minutes at 4°C and then frozen overnight at -20°C.

Preparation of DEPC-treated water

To 1l of distilled water 1ml of DEPC is added with stirring, allowed to stir for 1hr then autoclaved and stored at room temperature.

S30 Buffer

A 10x S30 buffer was made from stock solutions of 1M Tris acetate pH 8.2, 1.4M Mg acetate, 6M K acetate, diluted just before use and DTT added from a stock 100mM solution to give a 1x S30 buffer.

1x S30 Buffer: 10mM Tris Acetate pH 8.2
14mM Magnesium acetate
60mM Potassium acetate
1mM DTT

2.13.2.2 Preparation of S30 Extract

2.13.2.2.1 By French Press

Cell pellets were thawed at 4°C for ≈30 minutes then slowly homogenised in S30 buffer + 7.2mM mercaptoethanol (100ml per 10g cells) then centrifuged for 30 minutes in a Sorvall GSA rotor at 10krpm and 4°C. The cell pellets were weighed and resuspended in S30 buffer, 60ml per 50g cells. This homogenate was passed through a chilled French Press at 8 400psi and 100µl of 0.1M DTT added per 10ml as soon as it was collected. This was centrifuged in a Sorvall SS-34 rotor at 15.5krpm for 30 minutes at 4°C and the supernatant removed, recentrifuged as before then retained on ice for the preincubation step.

2.13.2.2.2 By Alumina Grinding (Matthaei and Nirenberg 1961)

The frozen cell pellets were cracked into small pieces and ground at 4°C with 1.5 times the cell mass of alumina using a chilled pestle and mortar (of unglazed porcelain) until the mixture was perfectly homogeneous. The paste was removed by the addition of 2vols of S30 buffer per vol of paste. The alumina, whole cells and cell debris were removed by

centrifugation in a Sorvall SS-34 rotor at 13krpm for 15 minutes at 4°C. The supernatant was centrifuged twice in a Sorvall SS-34 rotor at 15.5krpm for 35 minutes at 4°C and retained on ice for the preincubation step.

2.13.2.3 Preincubation of S30 Extract

7.5ml of preincubation mix per 25ml S30 extract were added in a conical flask and incubated at 37°C for 80 minutes in a shaking waterbath. This mixture was transferred to 1" wide dialysis tubing and dialysed vs 50vols S30 buffer at 4°C 4x45 minutes. The mixture was then rapidly aliquoted in 200µl amounts, snap-frozen in liquid nitrogen then stored at -80°C.

Preincubation Mix

Pyruvate kinase	132 units
2.2M Tris acetate pH 8.2	1ml
3M magnesium acetate	23µl
38mM ATP pH 7.0	2.63ml
0.42M phosphoenol pyruvate (PEP) pH 7.0	1.5ml
Mixture of 20 amino acids, each at 50mM	6µl
0.55M Dithiothreitol (DTT)	60µl
DEPC-water to a final volume of 7.5ml	

Preparation of dialysis membrane

Dialysis membrane was boiled for 15 minutes in 0.1M Na_2CO_3/0.01M EDTA pH 8.0, then rinsed thoroughly in DEPC-treated water and stored at 4°C in DEPC-treated water.

2.13.2.4 Assay of coupled transcription/translation

An aliquot of extract was removed from -80°C and allowed to thaw on ice for 1 hour. The DNA (0.5µg in 5µl water), ^{35}S-methionine + LMM, water and magnesium acetate were mixed and incubated for 4 minutes at 37°C while the

S30 extract was spun down in the microfuge. 5μl of the S30 supernatant were added to the mixture and incubation continued for 20-30 minutes. 5μl of 44mgml^{-1} methionine (prewarmed at 37°C) were added as a chase and incubation continued for 5 minutes. Samples were then placed on ice, aliqouts removed for TCA precipitation and an equal volume of SDS-loading buffer added to the remainder. Samples were stored at -20°C until ready to be analysed. Samples were boiled for 10 minutes before running 20μl on an SDS-PAGE gel. Gels of 14% were routinely used.

Assay Reaction;

0.5-1μg Caesium chloride/ethidium bromide density gradient-purified DNA

0-5μl 0.1M magnesium acetate

LMM+^{35}S-methionine: 9.5μl LMM containing 1μl (15μCi) ^{35}S-methionine

5μl S30 extract

DEPC-treated water to a total volume of 30μl

Low Molecular Weight Mix (LMM)

Constituents of the LMM are shown in Table 10. Stock solutions were stored in aliquots at -20°C (excluding Tris, PEG, inorganic mix which were stored at room temperature). The LMM could be made up and stored at -20°C for several weeks. All stock solutions were sterilised by autoclaving, except DTT, ATP, CTP/GTP/UTP, PEP, FA, cAMP, tRNA and amino acid mixtures which were filter-sterilised.

2.13.2.5 TCA-Precipitation of ^{35}S-labelled polypeptides

2μl aliquots of the incubation mix were spotted onto filter paper wicks (3 replicates per incubation), allowed to dry then placed in 10% TCA + 100μgml^{-1} methionine for 30 minutes at 4°C. The TCA solution was replaced by 5% TCA + 100μgml^{-1} methionine and heated to 90°C for 10 minutes. The wicks were then washed 3 times in 10% TCA + methionine at 4°C, then rinsed in acetone and allowed to dry. The counts were measured by placing the dry

TABLE 10

LOW MOLECULAR WEIGHT MIX (LMM)

Stock solution	Volume (μl)	Final concentration in incubation reaction
2.2M Tris acetate pH 8.2	40	64mM
0.55M DTT	5	2mM
38mM ATP pH 7.0	50	1.38mM
CTP/GTP/UTP mix each at 88mM pH 7.0	15	0.96mM
0.4M PEP pH 7.0	100	29mM
19 amino acid mix each at 55mM	10	400μM
40% Polyethylene Glycol 6000	75	2.2%
2.7mgml^{-1} Folinic acid	20	39μgml^{-1}
50mM cAMP pH 7.0	20	0.726mM
17.4 mgml^{-1} tRNA	15	
Inorganic mix:	40	
1.4M ammonium acetate		41mM
2.8M potassium acetate		81mM
0.38M calcium acetate		11mM

wicks in scintillation vials containing "Optiphase" scintillation fluid and measuring β-emission with a scintillation counter.

2.14 DISCONTINUOUS SDS-PAGE AND FLUOROGRAPHY. (Laemmli 1970)

Gel plates of 18x20cm were cleaned throughly with ethanol and assembled with side spacers of 1.5mm and the edges sealed with 2.5cm wide yellow tape. The sides were clamped with bulldog clips. For a 14% separating gel 24.7ml Buffer A, 15.9ml of 44% acrylamide mix, 9.2ml water and 1.74ml of a freshly prepared $10mgml^{-1}$ solution of ammonium persulphate were mixed together in a 50ml conical flask, taking care not to generate bubbles. 2.5ml of this mixture were removed into a plastic universal and 20μl TEMED added and the mixture immediately poured between the vertical gel plates to form a plug along the bottom. This was allowed to set then 152μl TEMED were added to the remaining acrylamide solution and this was poured between the plates until it reached a depth of about 3cm below the edge of the smaller plate. A layer of Butan-2-ol was gently layered on the surface to exclude air and to aid the polymerisation process and the gel allowed to polymerise for about 30 minutes. The Butan-2-ol was removed by washing the surface of the gel with distilled water, then the area above the gel was dried with filter paper. For a stacking gel of 5.3% 15ml Buffer B, 3.7ml 44% acrylamide mix, 12ml water and 0.75ml ammonium persulphate were mixed in a 25ml conical flask. 60μl TEMED were added and the solution immediately poured on top of the separating gel. A comb with well width of 6mm inserted between the plates avoiding the trapping of bubbles and the gel was allowed to polymerise for a further 30 minutes.

Having removed the clips and yellow tape the comb was gently removed and the plates assembled in the gel tank. The top and bottom reservoirs were filled with running buffer and bubbles were removed from the bottom of the gel by use of a needle and syringe. The gel wells were cleaned of any superfluous polyacrylamide using a plastic pasteur pipette. Samples were

boiled for 10 minutes then loaded with a Hamilton syringe and run at a constant current of 25mA through the stacking gel and 50mA through the separating gel. 25nCi (in 5µl) of [^{14}C] Methylated markers were diluted 1:1 with SDS-loading buffer and loaded onto the gel as molecular weight standards. When the bromophenol blue marker reached the bottom, the gel was disassembled from the tank and the upper plate removed and the stacking gel taken off and discarded. The gel was then placed in a bath of 10% glacial acetic acid and allowed to fix for 15 minutes. The acid was poured off, the gel rinsed in distilled water and then soaked in Amersham "Amplify" (a fluorographic enhancer) for 20 minutes. The "Amplify" was drained off and the gel transferred to Whatman 3MM paper, covered with "Saranwrap" and dried at 80°C under vacuum on a gel drier for about 1.5 hours. Dried gels were exposed to Amersham "Hyperfilm" for 2 days at room temperature, then the film developed.

Acrylamide Mix: 44% Acrylamide (Biorad 99.9% pure)
0.8% Bis-acrylamide

Buffer A: 0.75M Tris Base
0.2% SDS
pH adjusted to 8.8 with conc. HCl

Buffer B: 0.25M Tris base
0.2% SDS
pH adjusted to 6.8 with conc. HCl

Running Buffer: 0.025M Tris base
0.19% SDS
0.192M Glycine
pH should be 8.3-8.6 without adjustment

SDS-Loading Buffer: 0.0625M Tris HCl pH 6.8

2% SDS

20% v:v Glycerol

0.72M 2-mercaptoethanol

1mgml^{-1} bromophenol blue

2.15 KINASE-LABELLING OLIGONUCLEOTIDES WITH [γ-^{32}P]ATP

2.15.1 Forward Labelling reaction

5μg DNA (at 500μgml^{-1} in TE pH7.5) were mixed with 5μl 10x Forward buffer, 100 μCi [γ-^{32}P]ATP, 2μl T4 polynucleotide kinase (of 10Uμl^{-1}) and water added to give a final volume of 50μl. This was incubated at 37°C for 1 hr then at 70°C for 10 minutes to inactivate the enzyme.

10x Forward buffer: 0.7M Tris HCl pH 7.5

0.1M $MgCl_2$

1M KCl

50mM DTT

Prior to the ligation of the oligonucleotide to the plasmid vector, it was necessary to remove the 5'OH group and replace it with a 5'P. The oligonucleotides were therefore incubated with T4 polynucleotide kinase as above but with the replacement of [γ-^{32}P]ATP with "cold" ATP at a concentration of 10mM in the 10x Forward buffer.

2.15.2 Exchange Labelling Reaction (Maniatis *et al* 1982)

1-50pmoles of 5' ends of DNA were mixed with 5μl 10x Exchange buffer, 3μl 5mM ADP, 100μCi [γ-^{32}P]ATP, 2μl T4 polynucleotide kinase (of 10Uμl^{-1}) and water added to give a final volume of 50μl. This was incubated at 37°C for 1 hr then at 70°C for 10 minutes to inactivate the enzyme.

10x Exchange buffer: 0.5M imidazole chloride pH 6.6

0.1M $MgCl_2$

50mM DTT

1mM spermidine

1mM EDTA pH 8.0

Labelled probes for use in hybridisation experiments were purified by passing through an anion exchange column of diethylaminoethyl cellulose, DE52, in TE pH 7.5 made in a 1ml disposable plastic tip with a small plug of polyallomer wool. The column was washed with 2ml TE, the probe added and excess label eluted by washing with 3ml TE, then with 0.2M NaCl in TE. The labelled probe was eluted by washing with 1ml 0.6M NaCl in TE.

2.16 ACRYLAMIDE GELS FOR ANALYSING ^{32}P-LABELLED DNA FRAGMENTS

2.16.1 Denaturing Gel

A 20% linear sequencing size gel was used for the separation of oligonucleotides of less than 30bp. The urea was dissolved in the acrylamide solution, TBE and water; 250μl freshly-made ammonium persulphate were added and the solution degassed under vacuum. 100μl TEMED were added and the gel poured as in 2.22.3. The comb was inserted, the gel allowed to polymerise, then run against 1xTBE at 2000V (approximately 30W) constant voltage.

Gel Components: 25ml 40% acrylamide solution (see 2.22.3)

25g urea

5ml 10xTBE

Made up to 50ml with water.

The samples to be analysed by denaturing polyacrylamide gel electrophoresis were not purified on a DE52 column, but were ethanol precipitated twice with tRNA, removing the supernatant with a drawn-out pasteur. The DNA

pellets were resuspended in 3μl water and 2μl Formamide dye mix (see 2.22.3). They were boiled for 2 minutes then loaded on the gel using a Gilson pipette. The upper siliconised plate was removed and the gel surface covered with "Saranwrap", then exposed to X-ray film for about one hour.

2.16.2 Non-Denaturing Gel

A 20% linear 18x20cm gel was used for the separation of fragments between 10 and 100bp. The acrylamide, TBE and water were mixed and the ammonium persulphate added. 200μl of TEMED were then added and the gel poured as in 2.14, but without the inclusion of a stacking gel. The comb was inserted and the gel allowed to polymerise for about one hour. The gel was run against 1xTBE at 200V. The upper glass plate was then removed and the gel surface was covered with "Saranwrap", then exposed to X-ray film for about one hour.

Gel Components: 27.2ml 44% acrylamide
25.5ml water
1.28ml 3% ammonium persulphate (freshly prepared)
6ml 10xTBE

The samples to be analysed by non-denaturing polyacrylamide gel electrophoresis were not purified on a DE52 column, nor ethanol precipitated. 5μl of Ficoll loading buffer were added directly to the incubation reaction and the sample loaded on the gel with a Gilson pipette.

Ficoll Loading Buffer: 25% w:v Ficoll
0.25% w:v bromophenol blue
0.25% w:v xylene cyanol

Oligonucleotides used in this study were prepared by J.Keyte using an automated DNA synthesiser.

2.17 DNA DOT BLOTTING PROCEDURE

1µl DNA was dotted onto a filter: either nitrocellulose or "Hybond N" and allowed to dry. 1µl Dot Blot solution was spotted over the DNA and allowed to dry. The filter was then either baked at 80°C for 2 hours, if nitrocellulose was used, or wrapped in "Saranwrap" and exposed to uv light for 5 minutes on the transilluminator, if "Hybond N" was used.

DNA Dot Blot solution: 2M NaCl
0.2M NaOH
20mM EDTA pH 8.0

2.18 SOUTHERN BLOTTING PROCEDURE (Southern 1975)

A 1% agarose gel was run with samples to be blotted and 1kb ladder as size marker run on either side of DNA samples. This was photographed on the transilluminator with linear rules along the side and bottom and the gel was cut to size. The gel was placed in a pan of 0.5M NaOH/1M NaCl to denature the DNA for 15 minutes with occasional shaking, this was drained off and repeated. The gel was then washed in distilled water and neutralised with 0.5M TrisHCl pH 7.5/3M NaCl for 15 minutes, then drained and repeated. The gel was rinsed in 6xSSC, then placed on top of a piece of Whatman paper that had been wetted in 6xSSC and placed onto a glass plate. A piece of wetted (in 6xSSC) nitrocellulose or "Hybond N" exactly the same size as the gel was placed onto its surface, taking care not to trap any air bubbles between. 3 wetted pieces of Whatman paper exactly the same size as the gel were placed on the top of the filter, then 3 dry ones and twenty pieces of paper towel cut to size on the top. A glass plate was placed over this and weighted down. Blotting was allowed to proceed at room temperature for about 2 hours, then the gel was removed and examined under uv to check that transfer of DNA was complete. The filter was air dried then baked at 80°C, if nitrocellulose was used, or covered in "Saranwrap" and exposed to

uv light on the transilluminator for 5 minutes, if "Hybond N" was used, to bind the DNA to the filter. The filter was rinsed in 6xSSC, then placed in an hybridisation chamber with preincubation mix and placed at the temperature of prehybridisation for one hour. This solution was replaced with hybridisation solution, the probe added and placed at the temperature of hybridisation overnight.

The filter was washed in the appropriate solutions for required stringency several times at the hybridisation temperature, then air dried and exposed to X-ray film overnight at -70°C with an intensifying screen.

20x SSC: 3M NaCl
0.3M Trisodium Citrate
pH adjusted to 7.0

50x Denhardt's: 1% Ficoll (w:v)
1% PVP (w:v)
1% BSA FrV (w:v)

2.19 HYBRIDISATION USING END-LABELLED PROBES

The temperature of hybridisation used for experiments involving the end-labelled 25bp probe was 37°C. Filters were prehybridised in the hybrisisation solution for one hour at 37°C, then this was replaced by fresh hybridisation solution and the labelled probe and hybridisation allowed to proceed overnight at 37°C. The filters were washed twice in the low stringency washing solution and twice more in the higher stringency washing solution at 37°C.

Hybridisation Solution: 5x Denhardt's
6x SSC
0.1% SDS

Low stringency washing solution: 3x SSC

0.1% SDS

High stringency washing solution: 0.5x SSC

0.1% SDS

2.20 CATECHOL 2,3-OXYGENASE ASSAY (Murray *et al* 1972)

10ml overnight cultures grown in L-broth + appropriate antibiotics were pelleted in a bench centrifuge at 3krpm for 10 minutes at 4°C and resuspended in 2ml 75mM potassium phosphate buffer/10% acetone, then sonicated on ice using a titanium probe of intermediate size (127mm length x 9.5mm diameter) at 50W 2x10s with a 30s interval between bursts. The sonicates were centrifuged in a minifuge to remove whole cells and cell debris and the supernatant stored on ice. 10µl of this was assayed by adding to 1ml of freshly mixed 50mM potassium phosphate buffer/0.333mM catechol and the rate of change in absorbance at 375nm measured over an appropriate time period.

2.21 PROTEIN CONCENTRATION ASSAY

The concentration of protein in the cell-free extracts was measured by use of the "Biorad" protein assay reagent. A 1:5 dilution of this was made and the solution filtered. 10µl of cell-free extract were added to 1ml of diluted reagent, mixed and the OD at 595nm measured against a "blank" of 1ml reagent + 10µl 75mM potassium phosphate buffer/10% acetone. A standard curve was constructed using BSA (1µg to 20µg per 1ml reaction). At high protein concentration (greater than 10µg) the graph departs from linearity and so extracts which had a value in this non-linear range were diluted 1:1 with buffer and the protein concentration of this diluted extract measured and the measurement corrected for the dilution.

2.22 SEQUENCING DNA FRAGMENTS

2.22.1 Preparation of Single-stranded M13 Templates for Sequencing

2ml L-broth were seeded with a 1:50 dilution of an overnight culture of JM101. A single M13 "plaque" was picked into this and the culture grown at 37°C with good shaking overnight (about 15 hours). 1.5ml of the culture were removed into a tube and the cells pelleted in a microfuge. The supernatant was placed in a fresh tube and recentrifuged. 1ml of this supernatant was then placed in a fresh tube and 250µl of 2.5M NaCl/20% PEG 6000 added, the mixture vortexed and stood at room temperature for 30 minutes. The precipitated bacteriophage was pelleted in the microfuge for 5 minutes, then the supernatant carefully removed from the pellet with a drawn-out pasteur pipette. The tube was recentrifuged to bring down any further liquid from the sides of the tube and this was aspirated off with a drawn-out pasteur pipette. Any remaining solution was removed with a tissue. The pellet was resuspended in 100µl of 1.1M sodium acetate pH 7 and phenol extracted with 100µl of phenol/chloroform. The aqueous layer was removed and extracted with 25:1 chloroform: isoamylalcohol. The top layer was removed into 2 vols ethanol and placed at -20°C for several hours. The DNA was pelleted in a minifuge for 10 minutes then washed with ethanol at -20°C, the supernatant removed and the pellet dried in a vacuum desiccator. The pellet was resuspended in 30µl TE, the DNA concentration checked by agarose gel electrophoresis of 2µl and the remainder stored at -20°C.

1x TE: 10mM Tris HCl pH 7.5
1mM EDTA pH 8.0

2.22.2 Sequencing Reactions

The procedure used was based on the dideoxy chain terminating method described by Sanger *et al* (1977). The enzyme used was "sequenase", a

modified T7 polymerase from Cambridge Bioscience. For annealing of the sequencing primer to the single-stranded template 7μl (about 1μg) of the single stranded DNA were combined with 0.5pmoles (1μl) of M13 universal primer and 2μl sequenase buffer. This was warmed to 65°C for 2 minutes then allowed to cool slowly to room temperature.

To the annealed template-primer reaction 1μl 0.1M DTT, 2μl diluted labelling mix (1:5 with water), 0.5μl [α-^{35}S]dATP (5μCi) and 2μl (3U) diluted enzyme (1:8 with ice-cold TE) were added and mixed well. This was incubated for 5 minutes at room temperature. 2.5μl of each of the dideoxynucleotide mixes were placed into separate tubes and prewarmed at 37°C for several minutes. 3μl of the labelled template-primer mix were then transferred into each of these using a clean tip each time. They were then centrifuged and incubated at 37°C for 5 minutes. 4μl of Formamide dye mix were added to each. They were stored on ice before running on the gel (see 2.22.3).

5x Labelling Mix: 7.5μM dGTP
7.5μM dCTP
7.5μM dTTP

Sequenase Buffer: 200mM Tris HCl pH 7.5
100mM $MgCl_2$
250mM NaCl

Dideoxynucleotide Mixes:

ddGTP Mix:	ddTTP Mix:
80μM dGTP	80μM dGTP
80μM dATP	80μM dATP
80μM dCTP	80μM dCTP
80μM dTTP	80μM dTTp
8μM ddGTP	8μM ddTTP
50mM NaCl	50mM NaCl

ddATP Mix:	80μM dGTP	ddCTP Mix:	80μM dGTP
	80μM dATP		80μM dATP
	80μM dCTP		80μM dCTP
	80μM dTTP		80μM dTTP
	8μM ddATP		8μM ddCTP
	50mM NaCl		50mM NaCl

Formamide dye mix: 10ml deionised formamide

10mg xylene cyanol

10mg bromophenol blue

200μl 0.5M EDTA pH 8.0

2.22.3 DENATURING POLYACRYLAMIDE GELS FOR ANALYSING SEQUENCING REACTIONS

50cmx20cm, 0.4mm thick 6% gradient polyacrylamide gels were used. The gel plates were thoroughly cleaned with ethanol and the "rabbit eared" plate was siliconised with Dimethyldichlorosilane solution, then rinsed thoroughly with water and dried. They were assembled with 0.4mm side spacers, yellow tape and bull dog clips. 40ml 0.5xTBE sequencing solution and 210μl 10% ammonium persulphate (freshly made) were pipetted into a glass beaker and 12ml 2.5xTBE sequencing solution and 60μl 10% ammonium persulphate into a second beaker. These solutions were degassed under vacuum. 14μl TEMED were added to the 0.5xTBE mix and and 5μl TEMED were added to the 2.5xTBE mix, swirling to disperse. 10ml of the 0.5xTBE mix were immediately taken up in a plastic 25ml pipette followed by all of the 2.5xTBE mix. 4 bubbles were introduced into the end of the pipette to generate a rough gradient, then this was poured down the side of the glass plate assembly which was held at an angle of about 45°, taking care not to trap any air bubbles. The rest of the 0.5xTBE mix was gently poured over this gradient and the gel laid horizontally and the straight edge of the sharkstooth comb inserted. Bull dog clips were fastened on the glass plates either side of the comb. This was left to polymerise, then the comb removed

and the slot rinsed well with distilled water. The yellow tape was then removed from the bottom edge and the plates assembled in the electrophoresis tank.

1xTBE was placed in the bottom buffer reservoir and 0.5xTBE in the upper reservoir. The gel was then prerun at 1800V for about 15 minutes to warm it, then the sample slot was rinsed well with buffer using a needle and syringe taking care to remove all bubbles. The comb (well size of 5.7mm) was inserted and 5μl of the samples (having been heated to 85°C for 2 minutes and stored on ice immediately before loading) were loaded using a Gilson P20 pipette. When one set of templates had been loaded they were run a little way into the gel before the next set were loaded to prevent diffusion. The gel was run at 1800V constant voltage usually until the bromophenol blue marker reached the bottom edge of the gel for reading closest to the primer. If bases were to be read further upstream than about 250bp, the gel was run until the xylene cyanol marker reached the bottom edge.

The comb and spacers were then removed and the upper plate eased away from the gel. The surface of the gel was flooded with a fixing solution of 10% acetic acid; 10% methanol and left for 15 minutes. This was then poured off and the gel rinsed with distilled water. The gel was removed from the glass plate by overlaying it with a sheet of Whatman 3MM paper and lifting it off. It was covered with a piece of "Saranwrap" and dried down on a Biorad gel drier at 80°C under vacuum. It was then exposed to X-ray film at room temperature overnight.

10xTBE: 0.89M Tris Base
0.89M Boric Acid
2mM EDTA pH 8.0

40% Acrylamide solution: 38g acrylamide (SERVA)

2g bis-acrylamide

Made up to 100ml, deionised by stirring with "amberlite" for 10 minutes, then filtered and stored in the dark at 4°C.

0.5xTBE sequencing solution: 215g urea

25ml 10xTBE

75ml 40% acrylamide mix

Made up to 500ml and stored in the dark at 4°C

2.5xTBE sequencing solution: 86g urea

50ml 10xTBE

30ml 40% acrylamide solution

10g sucrose

10mg bromophenol blue

Made up to 200ml and stored in the dark at 4°C

CHAPTER 3

PRODUCTION OF AN *IN VITRO* COUPLED TRANSCRIPTION/TRANSLATION SYSTEM BASED ON *PSEUDOMONAS*

3.1 THE CELL-FREE SYSTEM

Cell-free systems for protein synthesis have been in use for some 30 years and have been the basis of some important scientific discoveries. In 1961 Nirenberg and Mattaei, using a system based on an S105 extract supplemented with ribosomes and ATP, found that protein synthesis was dependent on the addition of template (messenger) RNA, and that a synthetic polynucleotide, polyuridylic acid, resulted in the incorporation of L-phenylalanine into poly-L-phenylalanine. This laid the foundation for the elucidation of the genetic code.

Such extracts were solely translational systems dependent on the addition of exogenous mRNA, usually viral RNA. Low activity of such systems when bacterial mRNA was used was probably due to the fact that transcription and translation in prokaryotes are closely linked *in vivo*. Zubay (1973) developed a coupled transcription/translation system for the synthesis of proteins from DNA templates. This was an S30 extract of *E.coli* supplemented with a multitude of components considered to be essential for coupled transcription and translation. This method has since been modified (Collins 1979, Pratt *et al* 1981) and is the system of choice for examination of the protein products of cloned genes in *E.coli in vitro*. Protein synthesising systems based on the reconstitution of purified components are considered to be less analogous to the *in vivo* situation than these S30 coupled transcription/translation systems.

An *in vitro* system has certain advantages over semi-*in vivo* systems (such as minicells or maxicells) in that it can be generated from different strains having particular properties. For example, Zubay (1973) made an extract from a strain of *E.coli* which lacked the *lac* operon in order that protein synthesis directed by exogenous DNA carrying the β-galactosidase gene could be measured. Also it is more controllable, the experimenter can directly add components to the system to stimulate or repress expression.

Since the production of an *in vitro* system based on *E.coli*, extracts supporting protein synthesis have been derived from other bacteria; Leventhal and Chambliss (1979) made an *in vitro* protein synthesising system for *Bacillus subtilis*: this was an S150 fraction supplemented with initiation factors and washed ribosomes from *B.subtilis*. Thompson *et al* (1984) generated an active S30 extract of *Streptomyces lividans*, exactly as the Zubay method for *E.coli*, but with an alteration in the assay mixture; cAMP and tRNA were found not to be required for protein synthesis and calcium ions were strongly inhibitory. Dusha *et al* (1986) prepared an active S30 extract of *Rhizobium meliloti* using the Zubay method, they too found that the addition of tRNA was not necessary for protein synthesis. Kartasova *et al* (1981) produced an active S30 extract of *Agrobacterium tumefaciens* by manual disruption of the cells using alcoa and glass powder.

3.2 PRODUCTION OF AN *IN VITRO* GENE EXPRESSION SYSTEM FOR *P.PUTIDA*

3.2.1 Production of a *P.putida* S30 extract by use of the French Press

The starting point for the generation of a coupled transcription/translation system based on *Pseudomonas putida* was the method described for *E.coli* by Pratt *et al* (1981). This comprises a cell-free extract generated by the breakage of the cells using a French Press and centrifugation to yield an S30 supernatant. This is incubated with ATP, phosphoenol pyruvate and pyruvate kinase as an energy-regenerating system and amino acids. During this time the mRNA present becomes degraded and the sheared chromosomal DNA is subjected to exonucleolytic attack. Dialysis is then used to remove components added for this preincubation step, and the extract is frozen in small aliqouts in liquid nitrogen. Assay of coupled transcription/translation is usually carried out by the addition to the thawed extract of magnesium acetate, components of the low molecular weight mix; DTT, ATP, CTP, UTP, GTP, phosphoenol pyruvate, 19 amino acids (excluding methionine), polyethylene glycol, folinic acid, cAMP, tRNA,

ammonium acetate, potassium acetate, calcium acetate and ^{35}S-labelled methionine and exogenous plasmid DNA. Although the role that all of these components play in protein synthesis *in vitro* is unclear, Zubay (1973) established that each contributed to the stimulation of incorporation of radiolabel in an S30 extract of *E.coli*. The mixture is incubated at 37°C for times ranging from 30 minutes to one hour, then a solution of "cold" methionine added as a chase and the incubation continued for a further 5 minutes. The amount of incorporation of ^{35}S-methionine into protein can be estimated by measurement of the radioactivity in TCA-precipitable products using a β-emission scintillation counter. Labelled products are analysed by SDS-PAGE and fluorography. All solutions used in the preparation of the extract and for the assay are prepared using water treated with diethyl pyrocarbonate for the removal of contaminating RNase which would totally inhibit protein synthesis.

Two strains were used for the production of S30 extracts; *E.coli* MRE600 and *P.putida* mt-2 KT2440. MRE600 (used by Pratt *et al* 1981) was chosen since it is deficient in RNase I activity and mRNA synthesised in the S30 system may be more stable and give rise to increased levels of incorporation of radiolabel into protein products. KT2440 was chosen since it is deficient in a restriction system. This may be important when using DNA prepared in an heterologous host, such as *E.coli*, so that DNA added to the extract is not degraded. It is more convenient to prepare plasmid DNA from *E.coli* as the recovery of plasmid DNA from *P.putida* is low. The DNA used to programme the S30 extracts was routinely purified by caesium chloride/ethidium bromide density gradient centrifugation and was free of RNase, phenol and ethidium bromide. Such impurities strongly inhibit protein synthesis in coupled transcription/translation systems of *E.coli*.

S30 extracts of the two strains were generated using the French Press for cell disruption. The amount of incorporation of ^{35}S-methionine into protein is known to be highly dependent on the concentration of magnesium

ions (Wetekam *et al* 1971) and therefore for every extract produced the optimum magnesium concentration must be determined. The two extracts were optimised for magnesium using an *E.coli* plasmid pJW252 (Walker *et al* 1987), which is pUC8 containing the cloned pneumolysin gene of *Streptococcus pneumoniae*, in the *E.coli* extract (E1) and the *E.coli* plasmid pGB123 (which encodes genes involved in capsule biosynthesis cloned in pUC8, see 4.3.1.1) and a broad host range plasmid pGSS15 (Sharpe 1984) based on R300B specifying resistance to ampicillin and tetracycline in the *P.putida* extract (P1).

Figures 3, 4 and 5 show how incorporation of ^{35}S-methionine in the different extracts depends on the magnesium concentration. Both the *E.coli* and *P.putida* extracts exhibit DNA-dependent incorporation of radiolabel and the effect of varying magnesium concentration is similar in both systems. Maximum incorporation occurs at a specific but different concentration of magnesium in the two extracts; the optimum magnesium concentrations for the two extracts at which the maximum radiolabel is incorporated lie in the range previously described for S30 extracts of *E.coli* (Pratt *et al* 1981) and *S.lividans* (Thompson *et al* 1984). The *E.coli* extract required a magnesium concentration of 12.0mM for optimal protein synthesis giving a maximum amount of incorporation of radiolabel in the sample of 31.3×10^6 cpm. The *P.putida* extract required a magnesium concentration of 6.7mM for optimal protein synthesis when programmed with pGB123 giving a maximum level of incorporation of radiolabel in the sample of 5.2×10^6 cpm, and an optimum magnesium concentration of 8.3mM when programmed with pGSS15 giving a maximum incorporation of radiolabel in the sample of 0.27×10^6 cpm.

The *E.coli* extract (E1) was fortuitously highly efficient at promoting protein synthesis; subsequent *E.coli* extracts were less efficient (see Figure 6). However, this *E.coli* extract (E1) incorporated radiolabel into TCA-precipitable material in the absence of exogenous DNA (background

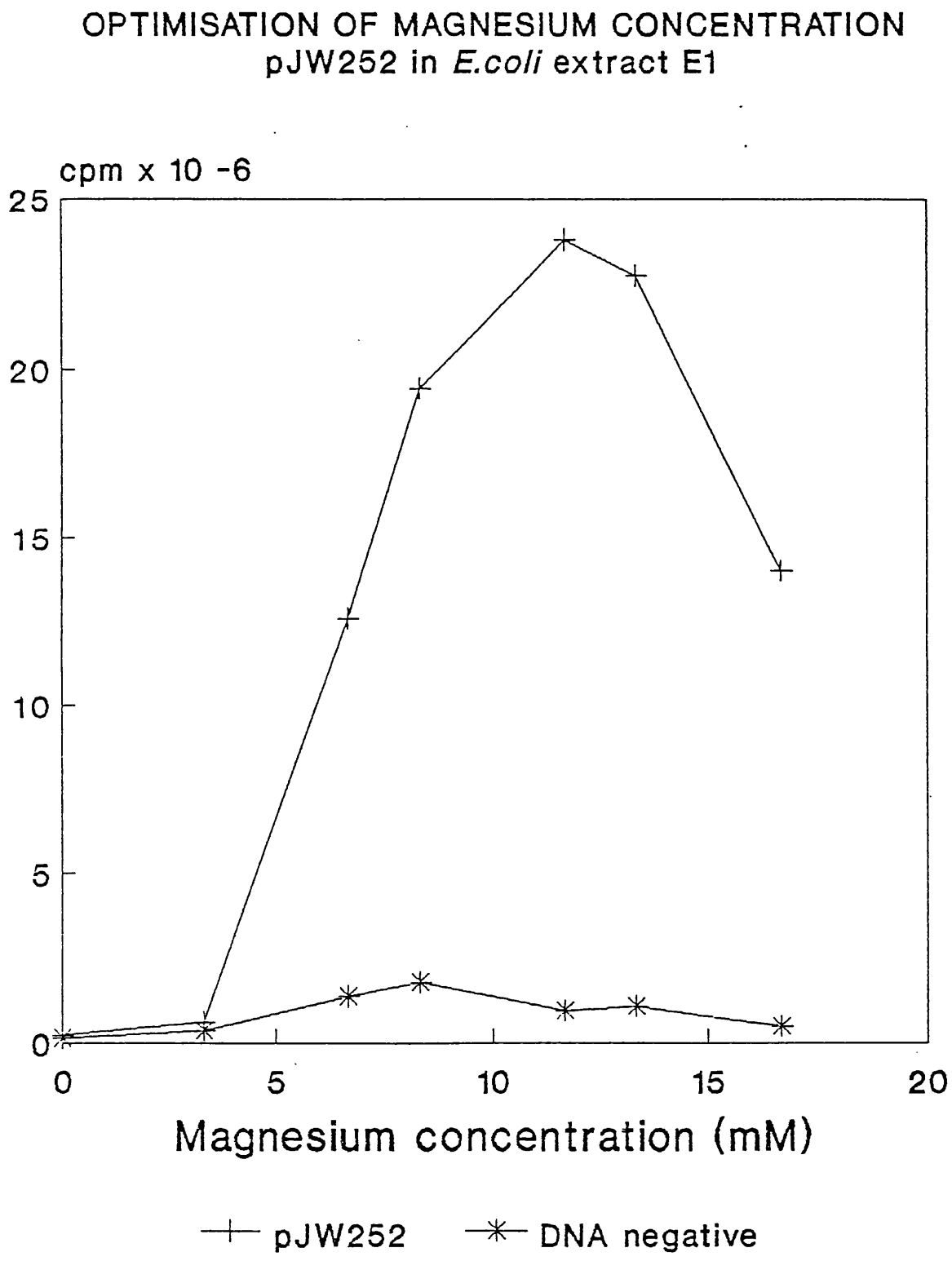

OPTIMISATION OF MAGNESIUM CONCENTRATION
pJW252 in E.coli extract E1
cpm x 10 -6
25
20
15
10
5
0
0
5
10
15
20
Magnesium concentration (mM)
pJW252
DNA negative

FIGURE 3

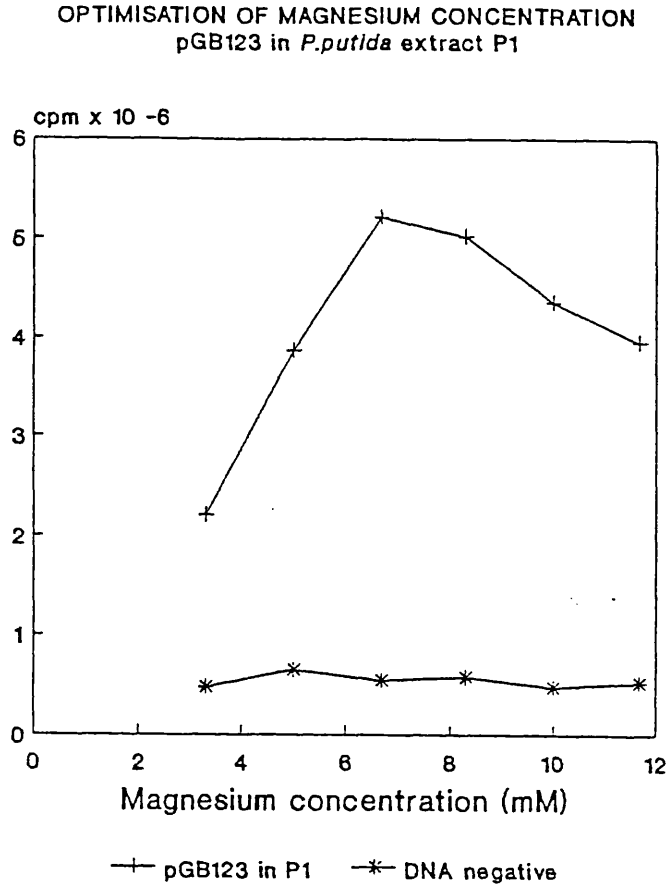


FIGURE 4

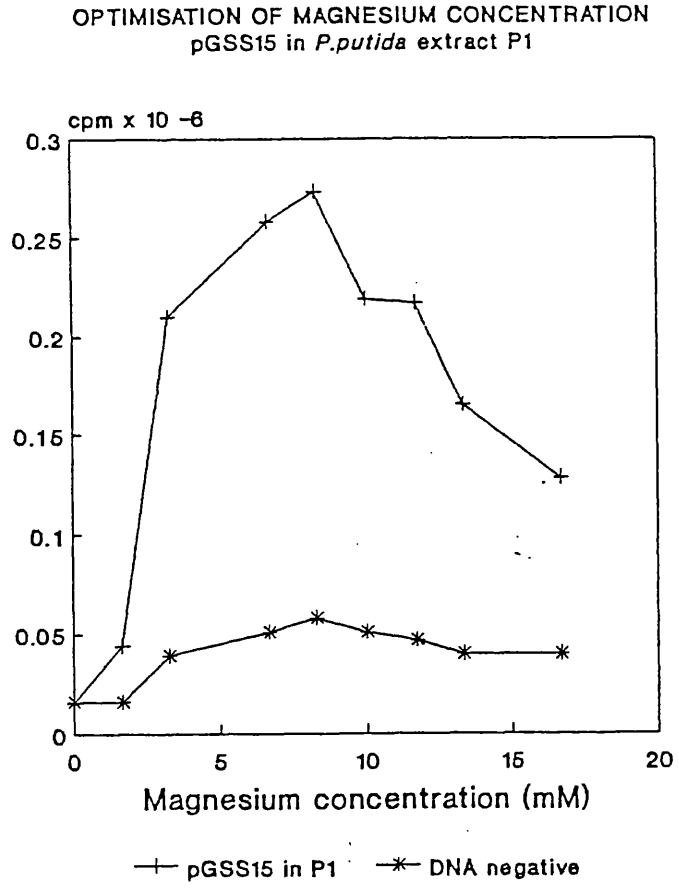


FIGURE 5

incorporation) to a higher degree than does the *P.putida* extract P1. Previous reported values were $1.8x10^6$cpm per sample for an extract of MRE600 (Jackson *et al* 1983). The difference between the maximum amount of incorporation into TCA-precipitable material for the *P.putida* extract programmed with the *E.coli* plasmid and the broad host range plasmid probably partly represents a difference between the DNA used to programme the extract, (a lower concentration of pGSS15 was used compared with pGB123), but also the specific activity of the radiolabel may have been lower. Alternatively, the LMM concentration was not quite optimal since the incorporation of ^{35}S-methionine into the sample incubated in the absence of exogenous DNA was significantly lower than that corresponding to the incubation of pGB123 in the *P.putida* extract (P1). It probably does not represent a difference due to the host range of the particular plasmids since pGSS15, based on the broad host range plasmid R300B, is known to express genes in *P.putida*. The optimum magnesium concentrations were used in subsequent assays.

In order to assess the accuracy of protein synthesis in this *E.coli* extract before it was itself used for comparison with the *P.putida* extract, it was programmed with pGB123 and the resultant sample was analysed by SDS-PAGE alongside ^{35}S-labelled *E.coli* minicells harboring the same plasmid. The results (Figure 7) indicated that the polypeptides synthesised in this *E.coli* extract were almost identical to those synthesised by *E.coli* minicells, except the β-lactamase polypeptide was present as the unprocessed form with apparent molecular weight of 30kD in the *in vitro* system. This reflects the absence of membranes which prevents processing. In the *in vivo* system the unprocessed β-lactamase was present, but the majority existed as the processed polypeptide of molecular weight 29kD.

A further *E.coli* extract (E1') was produced exactly as E1 and it was optimised for magnesium concentration using plasmid pGB123 (Figure 6). This second extract directs the synthesis of exactly the same polypeptides as

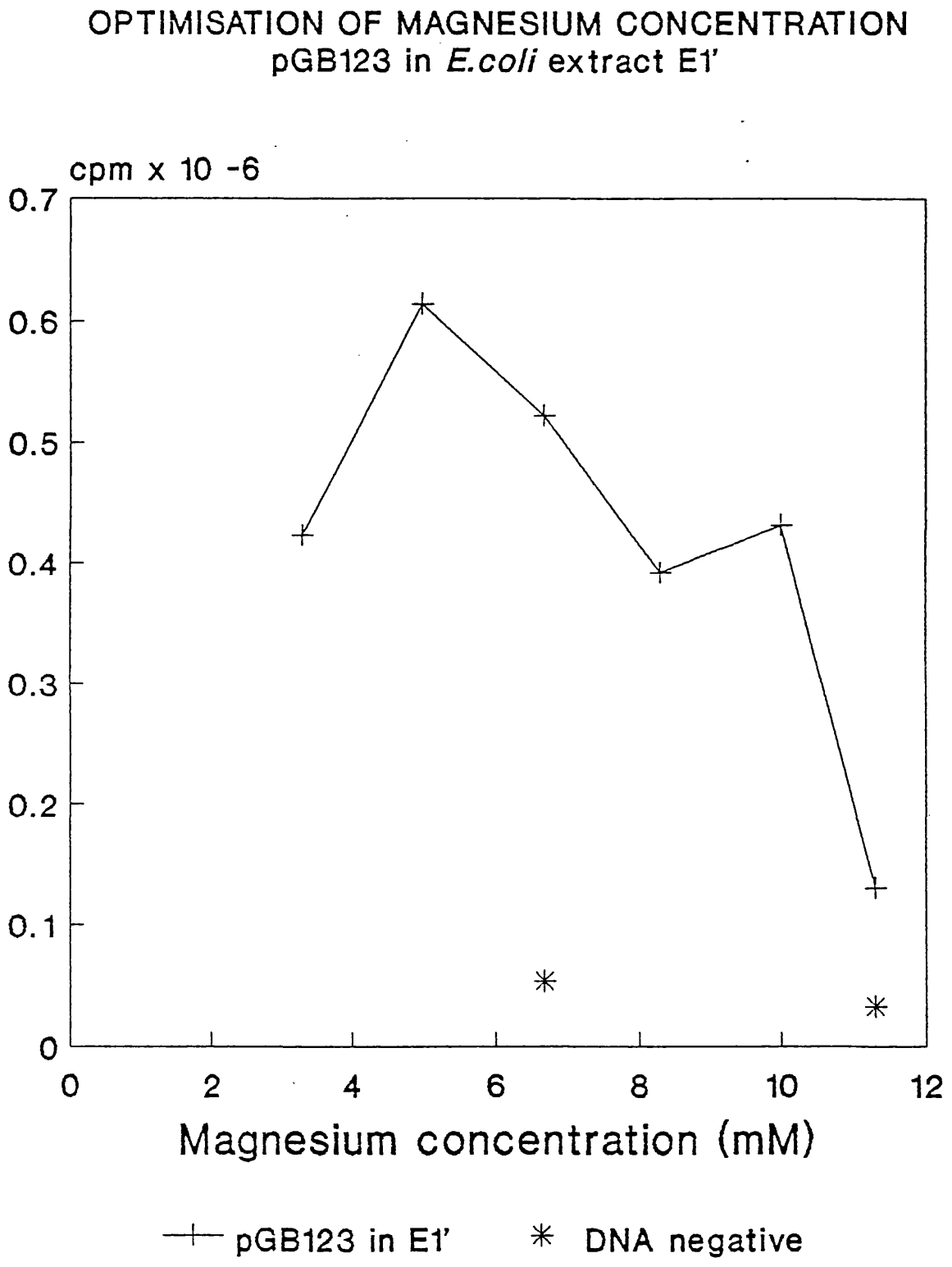

OPTIMISATION OF MAGNESIUM CONCENTRATION
pGB123 in E.coli extract E1'
cpm x 10 -6
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
Magnesium concentration (mM)
pGB123 in E1'
DNA negative

FIGURE 6

FIGURE 7 Comparison of polypeptides produced by pGB123 in the *E.coli* S30 extract and in *E.coli* minicells.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *E.coli* minicells harboring pGB123 were labelled with ^{35}S-methionine (Track b). *E.coli* S30 extract was incubated with pGB123 with magnesium concentrations of 10mM, 11.7mM and 13.3mM (Tracks c, d and e respectively). As a control the *E.coli* S30 extract was incubated without DNA with magnesium concentrations of 10mM, 11.7mM and 13.3mM (Tracks f, g and h respectively). The solid arrows indicate unprocessed β-lactamase and the open arrow indicates processed β-lactamase. Track a contains molecular weight standards.

FIGURE 8 Polypeptides directed by pGB123 in the *E.coli* and *P.putida* French Press-derived S30 extracts.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. The *P.putida* French Press-derived S30 extract (P1) was incubated with pGB123 at magnesium concentrations of 3.3 to 13.3mM (Tracks b-h respectively) and without DNA at magnesium concentrations of 3.3 to 13.3mM (Tracks k-o respectively). The *E.coli* French Press-derived S30 extract (E1) was incubated with pGB123 at optimal magnesium concentration (Tracks i and j). Open arrows indicate cloned gene products, solid arrow indicates β-lactamase and the brackets mark the *Pseudomonas*-specific DNA-independent product. Track a contains molecular weight standards.

FIGURE 7

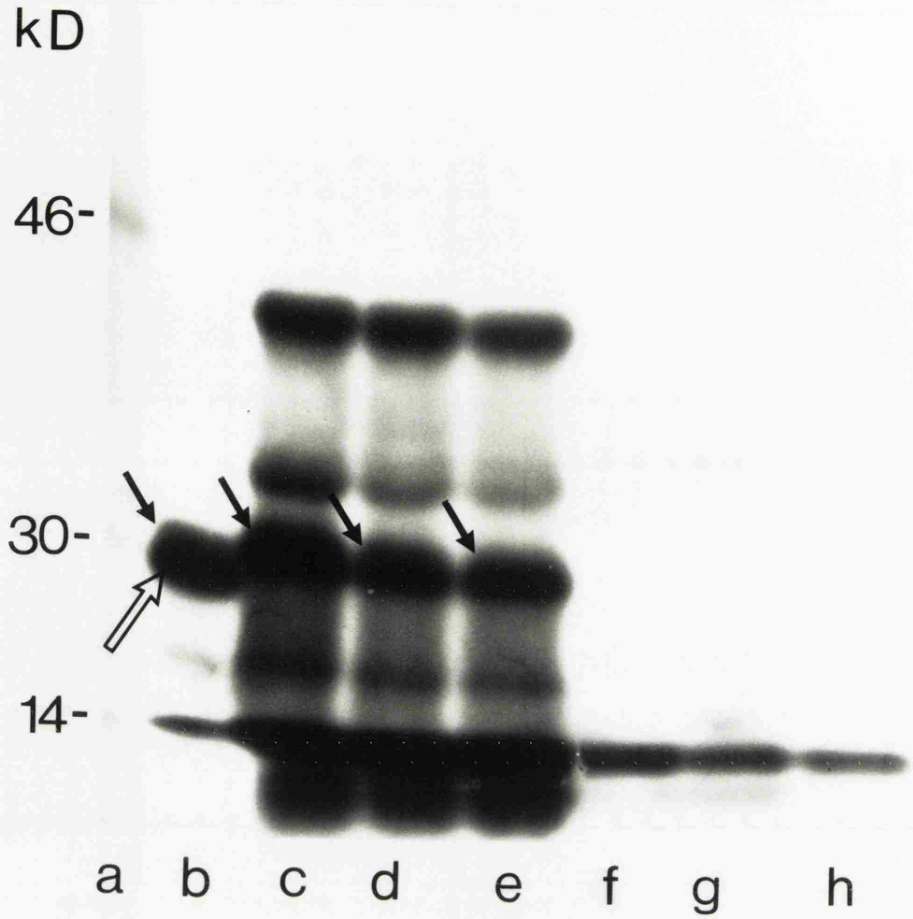
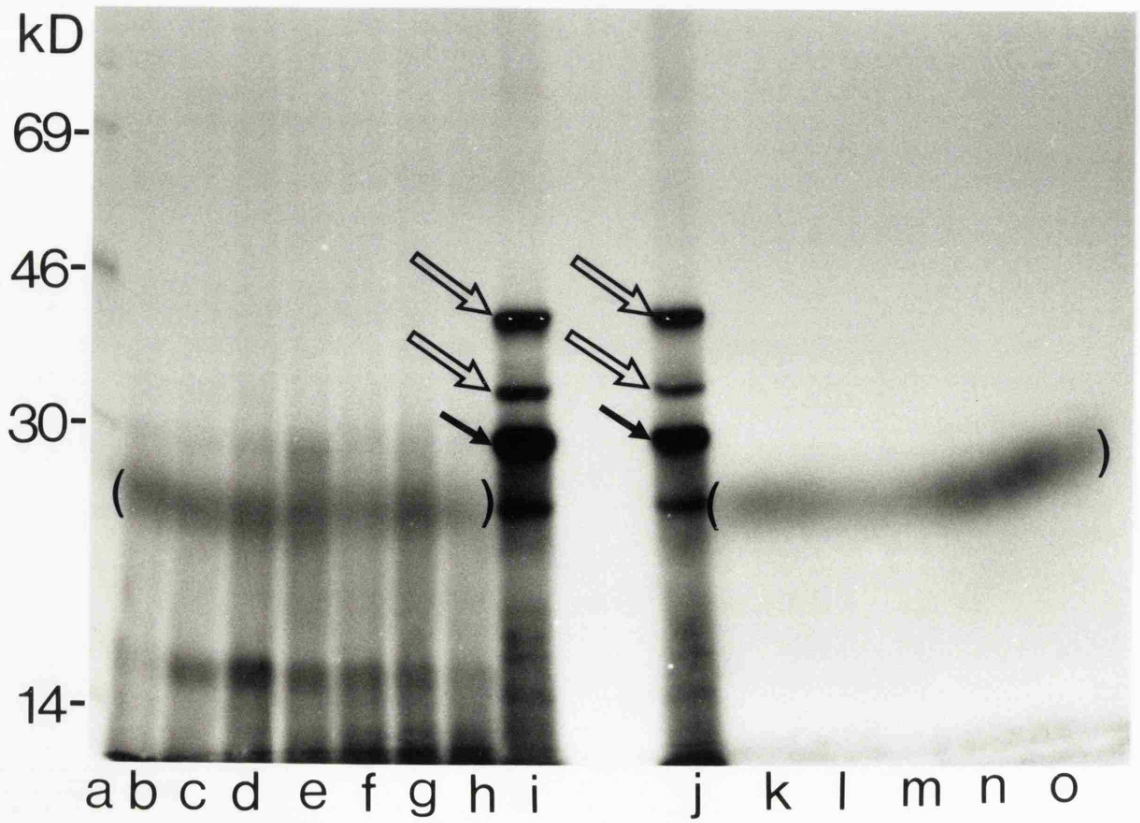


FIGURE 8



E1, but its maximum activity as measured by incorporation of radiolabel into TCA-precipitable material was lower than E1, but was of a similar level as that of the *P.putida* extract P1.

Having established that both of these *E.coli* extracts (hereafter referred to as the *E.coli* S30 extract) showed no obvious aberration in their protein synthesising abilities, the S30 extracts of *E.coli* and *P.putida* were programmed with the *E.coli* plasmid pGB123 and the products analysed by SDS-PAGE and fluorography. Discrete polypeptides were synthesised in the *E.coli* S30 extract, the β-lactamase and the cloned products are indicated on Figure 8. However in the *P.putida* extract no polypeptides in excess of 30kD were synthesised. The radioactive material was in the form of a smear down the gel from 30kD. This suggested that either the *P.putida* system was subject to abortive transcription or translation, or aberrant initiation or termination of transcription or translation was occurring. Alternatively, full length products were synthesised but then subsequently degraded, or that the exogenous plasmid DNA was being restricted.

Restriction of the added DNA should not occur since the extract was derived from a restriction deficient strain of *P.putida*. The possible instability of the polypeptides synthesised in both extracts was investigated. The assays were carried out by programming both the *E.coli* and *P.putida* extracts with pGB123, incubating for 20 minutes and then "chasing" with non-radioactive methionine for varying lengths of time from 5 minutes to 40 minutes and determining the amount of TCA-precipitable radioactivity for each. The results (Figure 9) indicate that for neither extract is there a significant breakdown of the products (into TCA-soluble material, however there may be degradation by endopeptidases which would not be detected). In fact, for the *E.coli* extract it is quite the reverse with the incorporation of radiolabel continuing beyond the usual 5 minute chase. For "chases" of 5 to 20 minutes the incorporation of radiolabel by

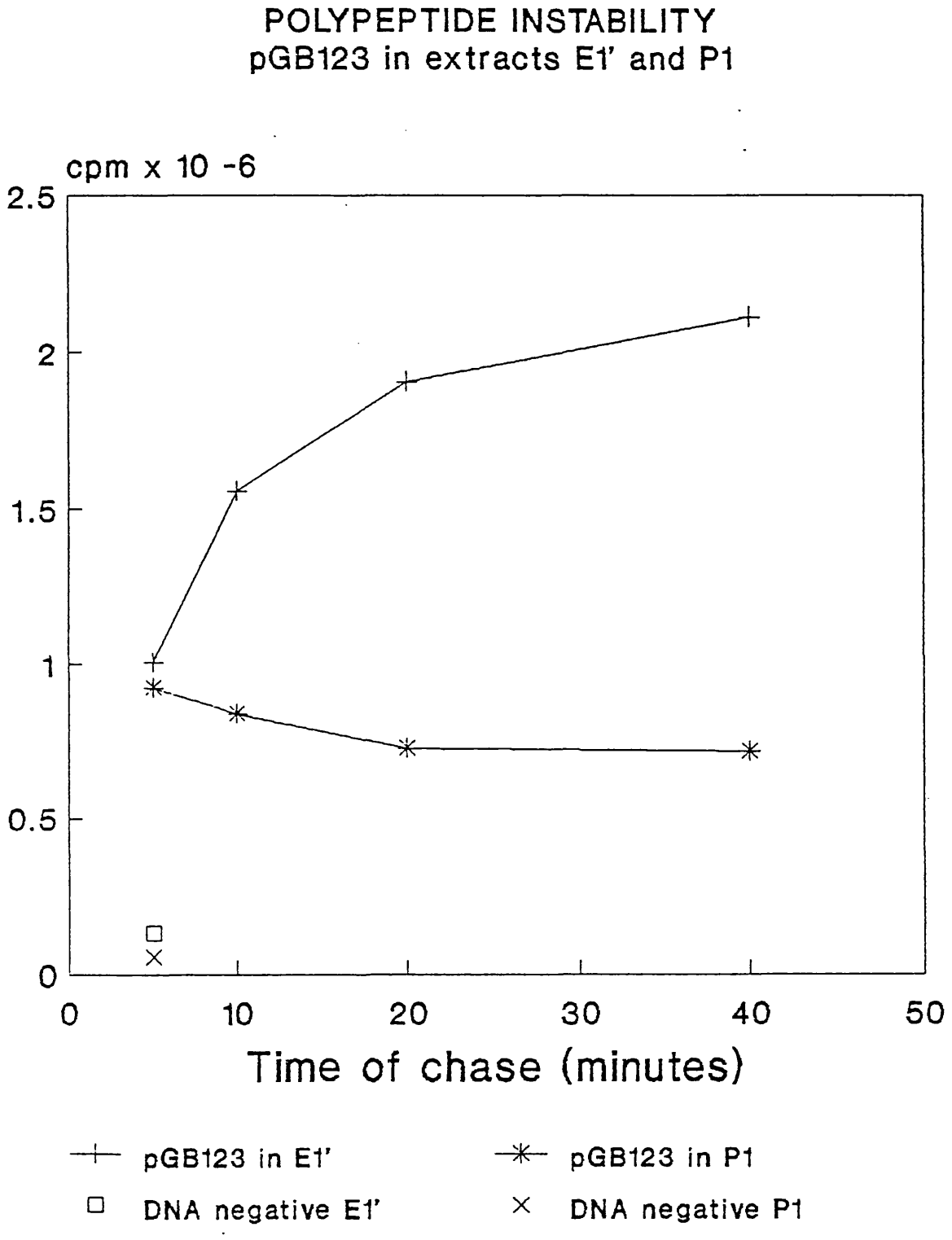

POLYPEPTIDE INSTABILITY
pGB123 in extracts E1' and P1
cpm x 10 -6
2.5
2
1.5
1
0.5
0
0
10
20
30
40
50
Time of chase (minutes)
pGB123 in E1'
pGB123 in P1
DNA negative E1'
DNA negative P1

FIGURE 9

the *P.putida* extract decreases (Figure 9) suggesting that there may be some proteolytic activity in this *P.putida* extract. Analysis by SDS-PAGE (Figure 10) indicates that this is the case. Incorporation into polypeptide products following a 40 minute "chase" in the *E.coli* S30 extract is greater than that following a 5 minute "chase". In contrast there is an obvious decrease in incorporation of label after a 40 minute "chase" compared with that following a 5 minute "chase" with the *P.putida* extract. Chen and Zubay (1983) found that for their *E.coli* S30 system the proteins were stable for 2 hours following removal of the ^{35}S-methionine.

It was conceivable that the synthesis of non-discrete polypeptides was the result of programming a *Pseudomonas* coupled transcription/translation system with an *E.coli* plasmid that would not naturally occur in *Pseudomonas*. Genes from such plasmids might be expected to be aberrantly expressed in this foreign background. This was investigated by programming the *P.putida* extract with a broad host range plasmid known to replicate and express genes in *Pseudomonas*. The plasmid used was pKT231 (Bagdasarian *et al* 1981), an RSF1010-based broad host range plasmid cloning vector specifying resistance to kanamycin and streptomycin. Similar results were obtained using this plasmid as DNA template in the *P.putida* coupled transcription/translation system as were obtained with pGB123 (Figure 11), indicating that aberrant expression was not an effect of plasmid host range. To determine whether the results were due to extensive proteolytic activity in the extract, the protease inhibitor PMSF (phenyl methyl sulphonyl fluoride) was included in the assay reaction at a concentration of 0.03-0.2mM. Although in one experiment PMSF enhanced the incorporation of radiolabel into TCA-precipitable material, this was not found to be repeatable and when the products were analysed by SDS-PAGE no discrete polypeptides were seen.

The solution to this problem of non-discrete polypeptides came about fortuitously. In an attempt to reduce DNA-independent background

FIGURE 10 Stability of polypeptides in French Press-derived S30 extracts of *E.coli* and *P.putida.*

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *E.coli* French Press-derived extract (E1') was incubated with pGB123 with "cold" methionine "chases" of 5, 10, 20 and 40 minutes (Tracks a-d respectively). *P.putida* French Press-derived extract (P1) was incubated with pGB123 with "cold" methionine "chases" of 5, 10, 20 and 40 minutes (Tracks e-h respectively).

FIGURE 11 *E.coli* and *P.putida* French Press-derived S30 extracts incubated with the broad host range plasmid pKT231.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was incubated with pKT231 at optimal magnesium concentration (Tracks b and c) and without DNA (Tracks d and e). *E.coli* French Press-derived S30 extract was incubated with pKT231 at optimal magnesium concentration (Tracks f and g) and without DNA (Tracks h and i). Track a contains molecular weight standards.

FIGURE 10

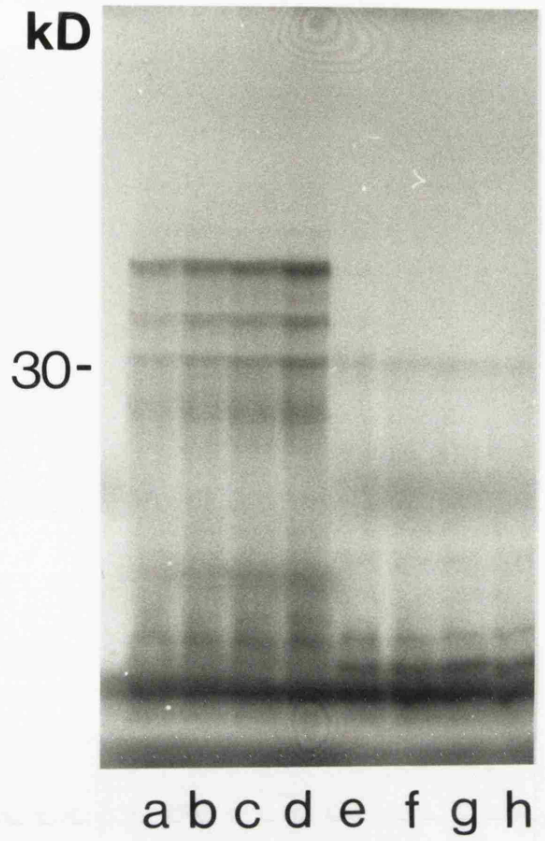
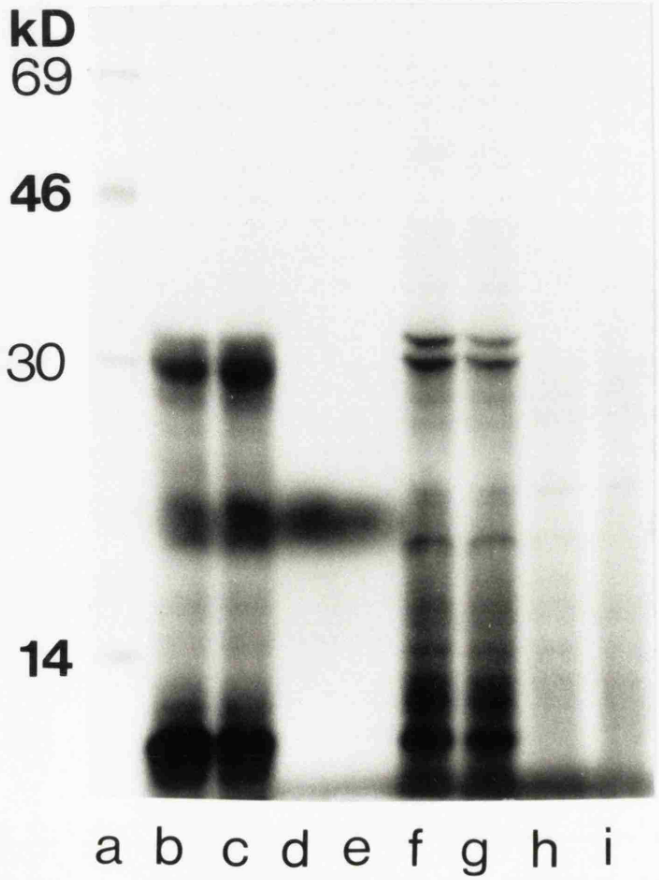


FIGURE 11



incorporation of ^{35}S-methionine the *P.putida* extract was preincubated for one hour at 37°C before the reaction was carried out, since it was possible that residual mRNA was present in the extract. This had some effect in this regard, but also on analysing the products, instead of the smearing, discrete bands were evident (Figures 12 and 13). Although PMSF (a generalised protease inhibitor) had had no effect in this regard the effect of preincubating the extract on the resultant appearance of polypeptide products may suggest that such a preincubation exhausts some protease activity of the extract. However still no polypeptides of greater than 30kD were synthesised by pGB123 even though polypeptides of greater molecular weight were synthesised by the same plasmid in the *E.coli* extract. These results imply that protein synthesis directed by either *E.coli* or broad host range plasmid DNA in this *P.putida* extract is not analogous to the expected *in vivo* situation.

3.2.1.1 Effect of temperature on incorporation of ^{35}S-methionine

Since the temperature optimum for growth of *P.putida* is 30°C, the effect on coupled transcription/translation of incubation of the reaction at 30°C and 37°C was investigated. The rate of incorporation of ^{35}S-methionine into protein was measured at these two temperatures using pKT231 as template in the *P.putida* extract (P1). The result (Figure 14) shows the rate of incorporation of radiolabel into TCA-precipitable material was decreased at the lower temperature and the majority of incorporation of radiolabel had taken place after 20-30 minutes incubation at both temperatures. On analysis of the products by SDS-PAGE (Figures 15 and 16), the polypeptides synthesised at the two temperatures are identical, but still all less than 30kD. However, there was an evident increase in the maximum amount of incorporation of radiolabel into those polypeptides synthesised at the higher temperature. For this reason all subsequent reactions involving both the *E.coli* and *P.putida* extracts were

FIGURE 12 Effect of preincubating the S30 extract on incorporation of ^{35}S-methionine into polypeptides directed by pGB123 in French Press-derived *P.putida* extract.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was preincubated before assaying coupled transcription/translation with pGB123 (Tracks c and d) or without DNA (Track e). The same extract was assayed for coupled transcription/translation without a preincubation with pGB123 (Tracks f and g) or without DNA (Track h). For comparison *E.coli* French Press-derived S30 extract (E1) was incubated with pGB123 (Track a) or without DNA (Track b).

FIGURE 13 Effect of preincubating the S30 extract on incorporation of ^{35}S-methionine into polypeptides directed by R300B in French Press-derived and alumina-derived *P.putida* extracts.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was preincubated before assaying coupled transcription/translation with R300B (Track d) or without DNA (Track h). The same extract was assayed for coupled transcription/translation without a preincubation with R300B (Tracks b and e) or without DNA (Track i). The alumina ground S30 extract of *P.putida* (P2) was preincubated before assaying coupled transcription/translation with R300B (Track f) or without DNA (Track j). The same extract was assayed for coupled transcription/translation without a preincubation with R300B (Tracks c and g) or without DNA (Track k). *E.coli* French Press-derived extract (E1') was incubated with R300B (Track a). Brackets indicate *Pseudomonas*-specific DNA-independent product.

FIGURE 12

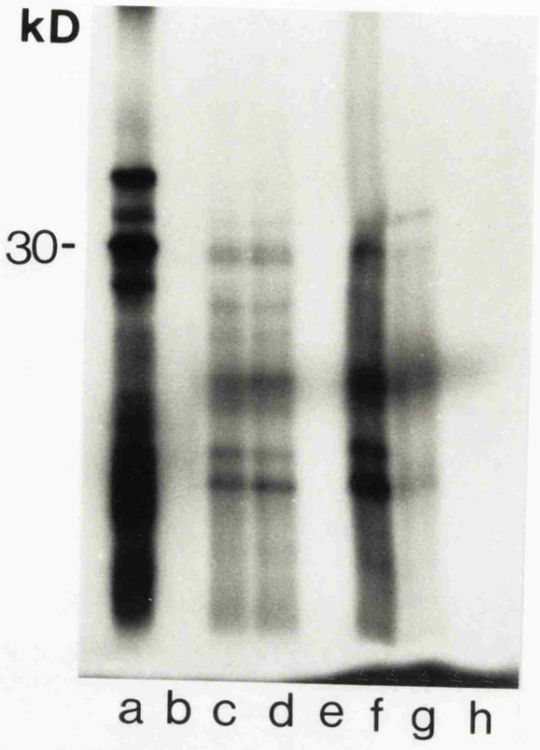
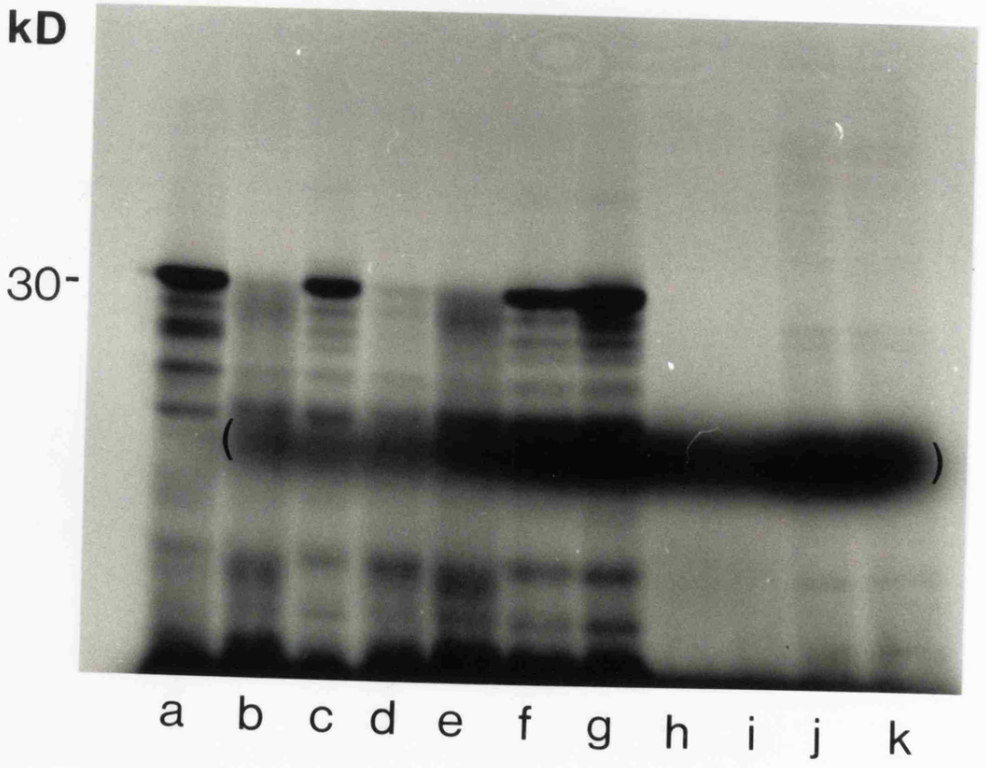


FIGURE 13



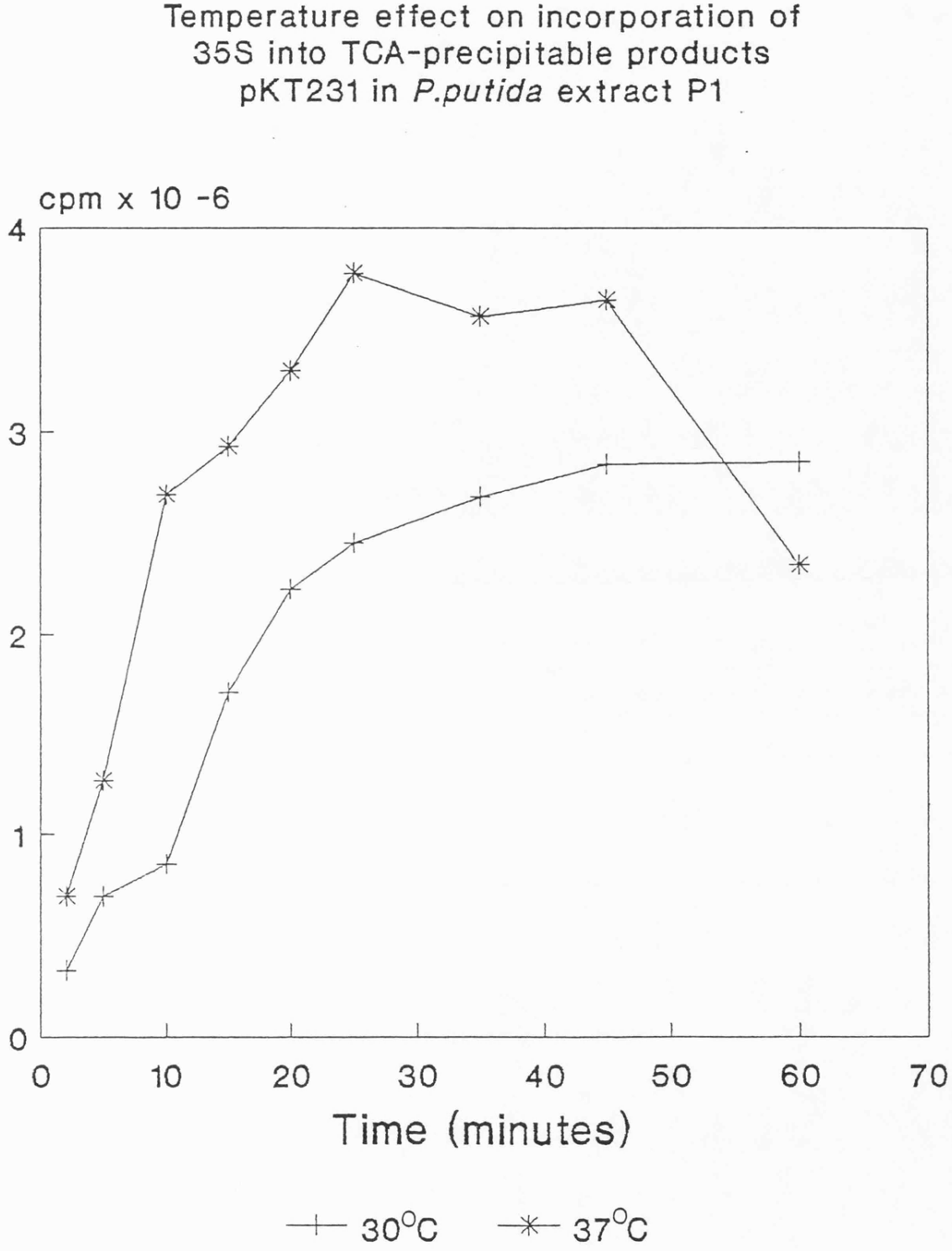

Temperature effect on incorporation of
35S into TCA-precipitable products
pKT231 in P.putida extract P1
cpm x 10 -6
4
3
2
1
0
0
10
20
30
40
50
60
70
Time (minutes)
30°C
37°C

FIGURE 14

FIGURE 15 Time course of incorporation of ^{35}S-methionine into polypeptide products in French Press-derived *P.putida* S30 extract at 30°C.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was incubated with pKT231 for 5, 10, 20 and 30 minutes at 30°C (Tracks b, c, d, e, respectively) or without DNA for 30 minutes at 30°C (Track f). Track a contains molecular weight standards.

FIGURE 16 Time course of incorporation of ^{35}S-methionine into polypeptide products in French Press-derived *P.putida* S30 extract at 37°C.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was incubated with pKT231 for 5, 10, 20 and 30 minutes at 37°C (Tracks b, c, d, e, repectively) or without DNA for 30 minutes at 37°C (Track f). Track a contains molecular weight standards.

FIGURE 15

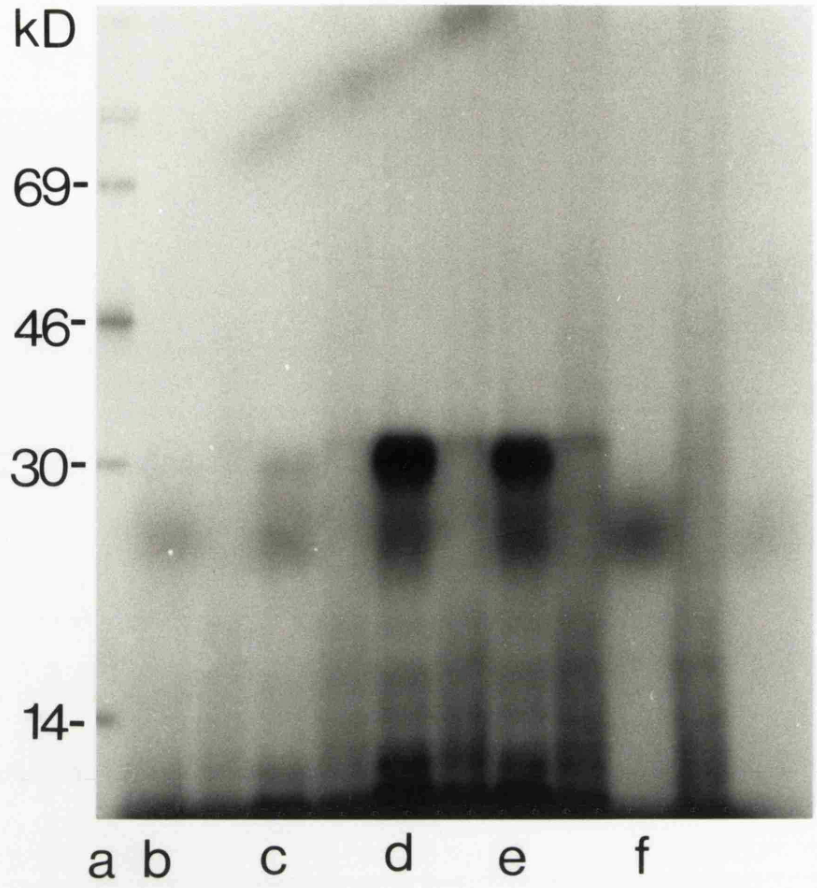
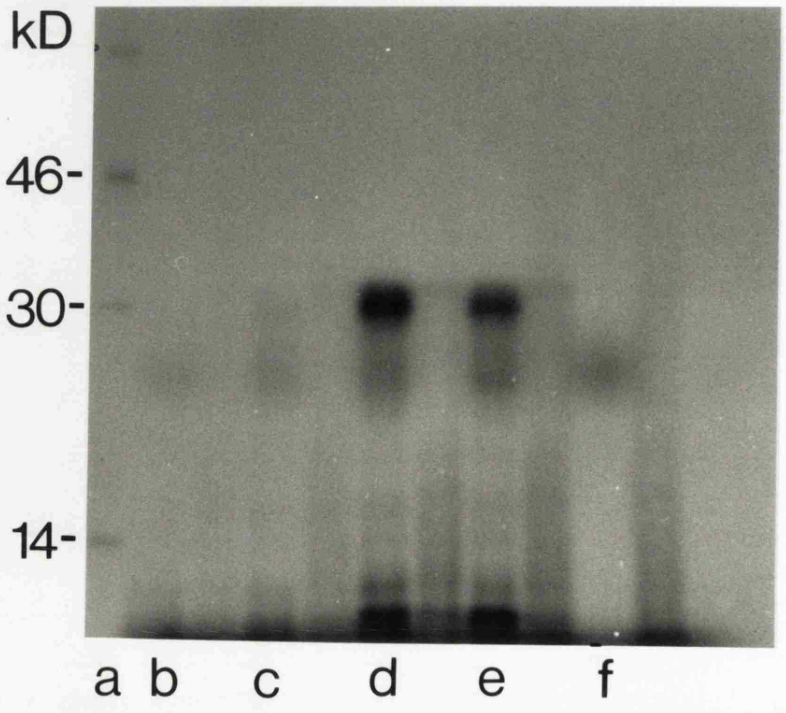


FIGURE 16



carried out at 37°C.

3.2.2 Production of a *P.putida* S30 extract by disruption of cells with alumina

The abnormal products of protein synthesis generated by the *P.putida* coupled transcription/translation system were potentially the result of physical damage to the cellular components of transcription or translation. It was plausible that the French Press was too severe in its disruption of the *P.putida* cells, perhaps damaging the ribosomes for instance. For this reason an alternative method for production of an S30 extract was employed based on that of Matthaei and Nirenberg (1961). This was essentially the same as the method described by Pratt *et al* (1981), but cell breakage was effected by grinding frozen cell pellets with alumina at 4°C using a pestle and mortar. ^{35}S-methionine incorporation into TCA-precipitable material by this extract, P2, was optimised for magnesium concentration using the broad host range plasmid RSF1010. This extract shows the typical dependence on magnesium for incorporation of ^{35}S-methionine into protein (Figure 17) exhibited by the French press-derived S30 extracts E1 and P1. The *P.putida* S30 extract generated by the method of grinding cells with alumina incorporated radiolabel into TCA-precipitable material as efficiently as both the *E.coli* and *P.putida* extracts made by the French Press method; the maximum amount of incorporation of ^{35}S-methionine into TCA-precipitable material at magnesium optimum of 8.3mM was 1.14×10^6 cpm per sample. An *E.coli* extract was synthesised in the same way and the optimum magnesium concentration for this *E.coli* extract (E2), using pGB123 as template was 6.7mM giving a maximum incorporation of radiolabel of 1.28×10^6 cpm per sample (Figure 18). Both of these extracts (E2 and P2) exhibit a higher background incorporation of radiolabel in the absence of DNA than the French Press-derived extracts (E1 and P1).

OPTIMISATION OF MAGNESIUM CONCENTRATION
RSF1010 in *P.putida* extract P2

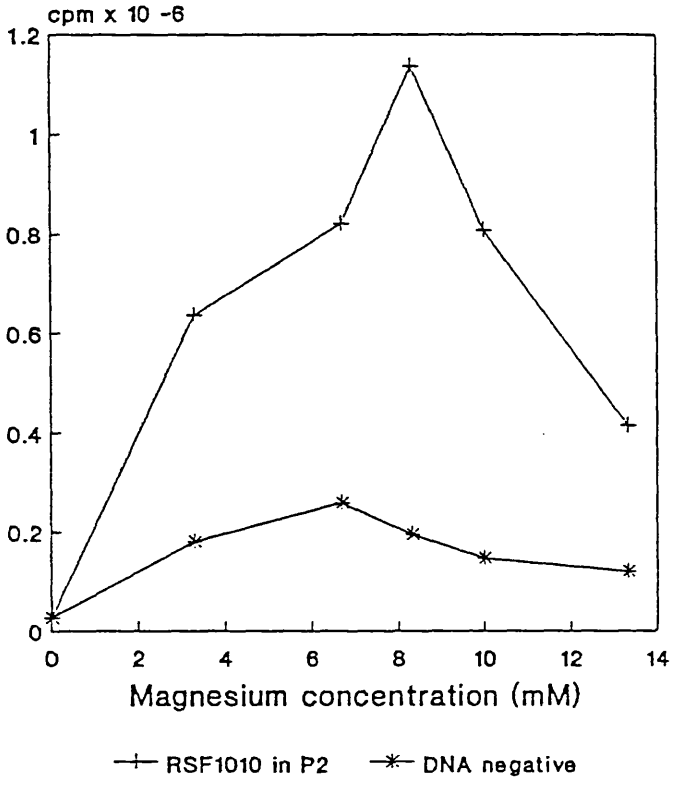


FIGURE 17

OPTIMISATION OF MAGNESIUM CONCENTRATION
pGB123 in *E.coli* extract E2

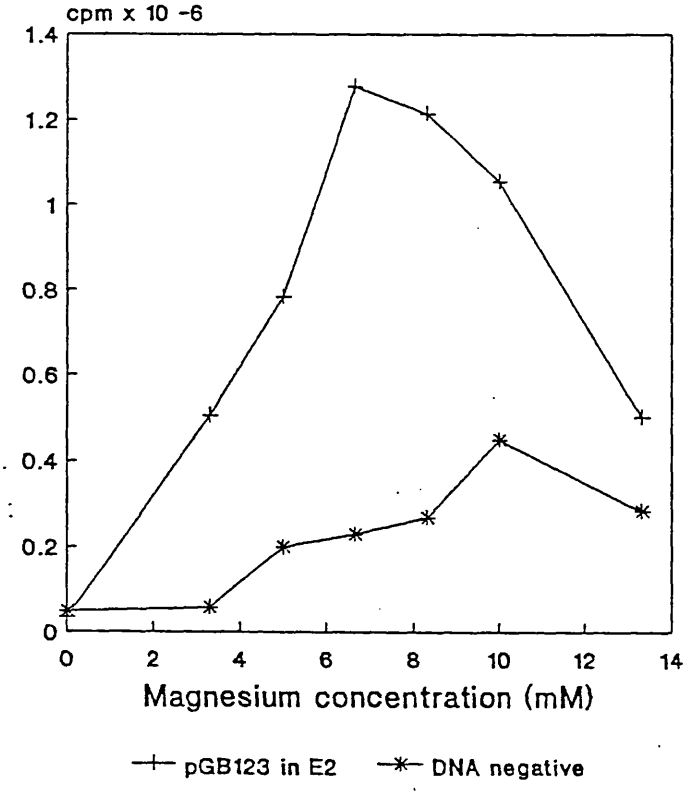


FIGURE 18

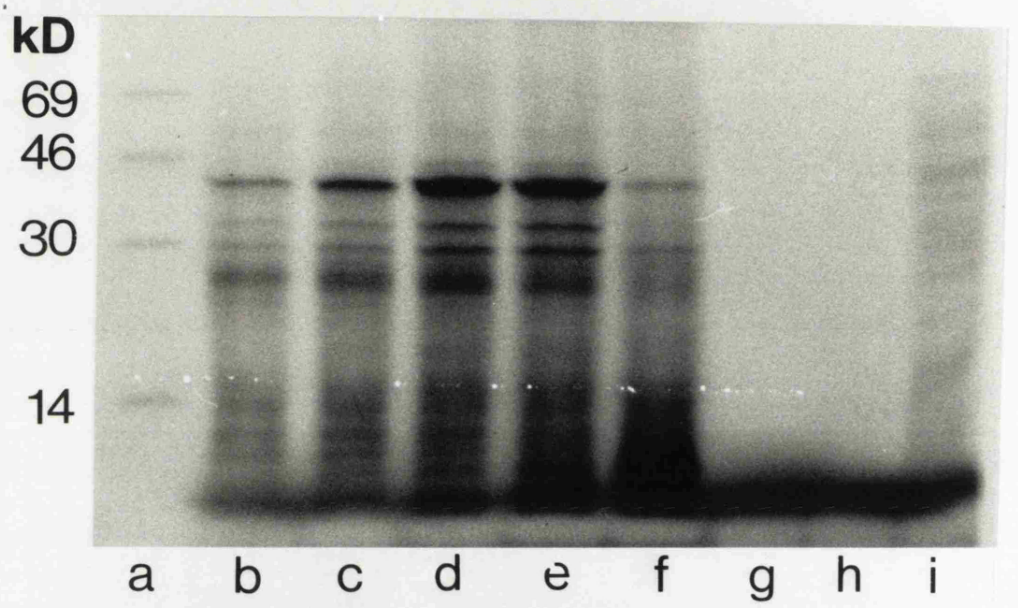
To assess the protein synthesising ability of these extracts, the alumina-derived *E.coli* extract (E2) was programmed with pGB123 and the polypeptide products compared with those synthesised by the French Press-derived *E.coli* extract (E1) with pGB123. Figure 19 shows the effect of varying magnesium concentration on polypeptide synthesis for the alumina-derived *E.coli* extract E2. At high magnesium concentrations, incorporation of ^{35}S-methionine into high molecular weight polypeptides was decreased (Track f). This is a typical result for all S30 extracts. Polypeptides synthesised in the alumina-derived *E.coli* extract (E2) were identical to those synthesised by the same plasmid in the *E.coli* extract (E1) which was generated by use of the French Press (Figures 8 and 19) indicating the applicability of this method for the production of S30 extracts for *in vitro* protein synthesis. However, the *P.putida* extract generated by grinding the cells with alumina (P2) exhibited the following differences compared with the French Press-derived *P.putida* extract (P1): discrete protein products were synthesised without the need for preincubation and polypeptide products of greater than 30kD were synthesised. On the basis of these results and the fact that the alumina-derived *E.coli* extract (E2) produces polypeptide products identical to those made in the French Press-derived *E.coli* extract (E1) by the same plasmid, the *P.putida* extract generated by alumina grinding (P2) was considered to be an *in vitro* protein synthesising system capable of accurate transcription and translation of exogenous DNA.

In order to determine whether the results of coupled transcription/translation were simply due to an irregularity by possible indeterminable experimental error during the production of the French Press-derived *P.putida* extract (P1), a further *P.putida* extract (P1') was made by the same method by which P1 was derived. The optimium magnesium concentration of this extract for protein synthesis was determined by programming the S30 extract (P1') with pGB123. Figure 20 shows that it has

FIGURE 19 Polypeptides directed by pGB123 in alumina-ground S30 extract of *E.coli.*

Labelled polypeptides were analysed by SDS-PAGE and fluorography. Alumina-ground S30 extract of *E.coli* (E2) was incubated with pGB123 at magnesium concentrations of 3.3, 5.0, 6.7, 8.3 and 10.0mM (Tracks b, c, d, e and f respectively) and without DNA at magnesium concentrations of 0, 3.3 and 5.0mM (Tracks g, h and i respectively). Track a contains molecular weight standards.

FIGURE 19



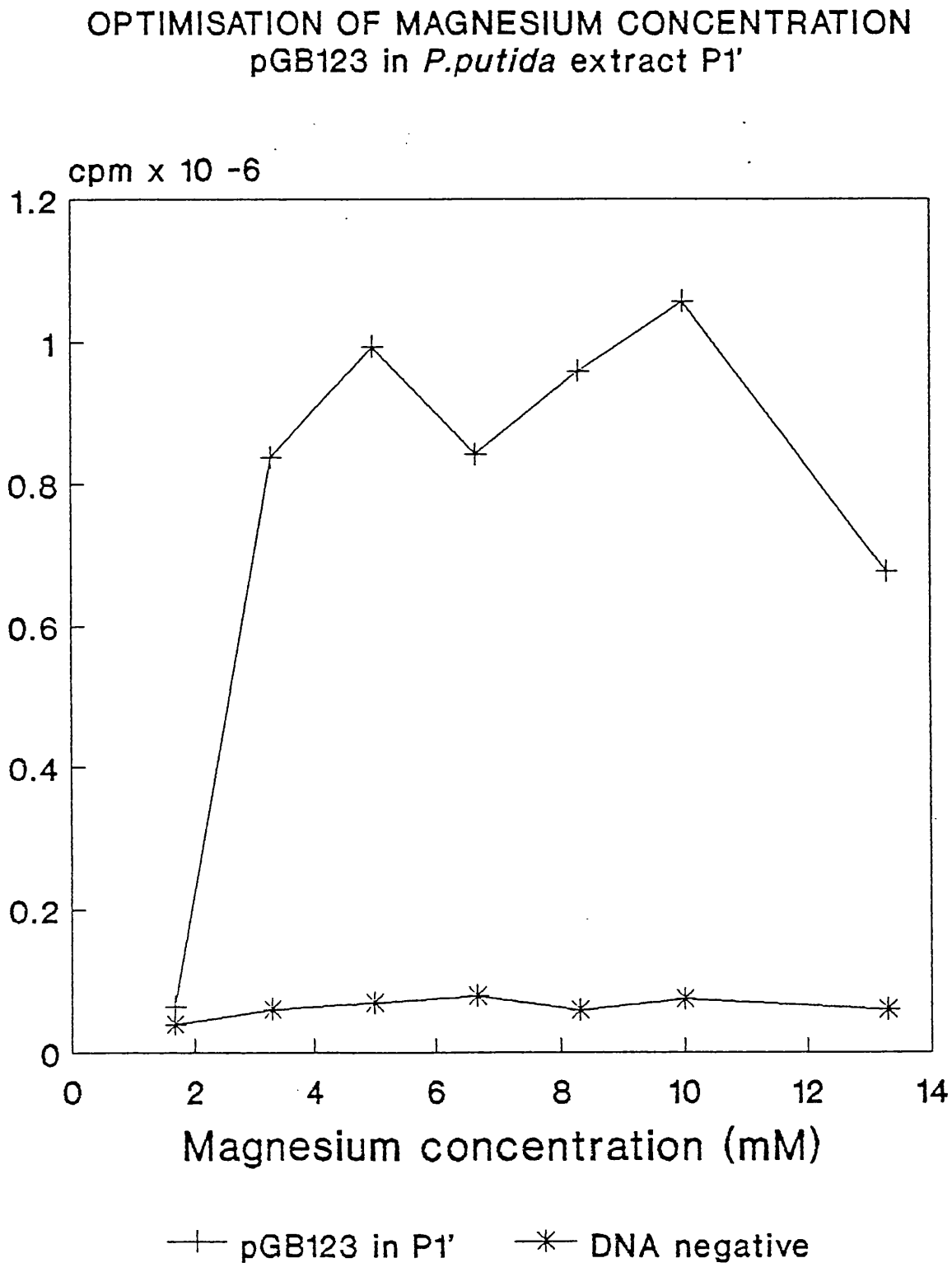

OPTIMISATION OF MAGNESIUM CONCENTRATION
pGB123 in P.putida extract P1'
cpm x 10 -6
1.2
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
12
14
Magnesium concentration (mM)
pGB123 in P1'
DNA negative

FIGURE 20

characteristics similar to that of the other French Press-derived *P.putida* extract (P1), with a magnesium optimum of 10mM giving a maximum incorporation of ^{35}S-methionine into TCA-precipitable material of 1.06×10^6 cpm per sample. Figure 21 shows extracts E1, P1, P1' and P2 programmed with pGB123, R300B and pGSS33 and Figure 29 shows a separate experiment of P1' programmed with pAT153. The French Press-derived extracts P1 and P1' show identical incorporation patterns; no discrete polypeptides were synthesised and there was no radiolabelled-material of molecular weight greater than 30kD. Thus aberrant expression of genes in a *P.putida* S30 system may be a general consequence of cell disruption using the French Press.

Hereafter all references to the "*E.coli* S30 extract" refer to the French Press-derived extracts (E1 or E1') and those references to the "*P.putida* S30 extract" denote the extract derived by alumina grinding (P2).

3.2.2.1 Effect of time on incorporation of ^{35}S-methionine

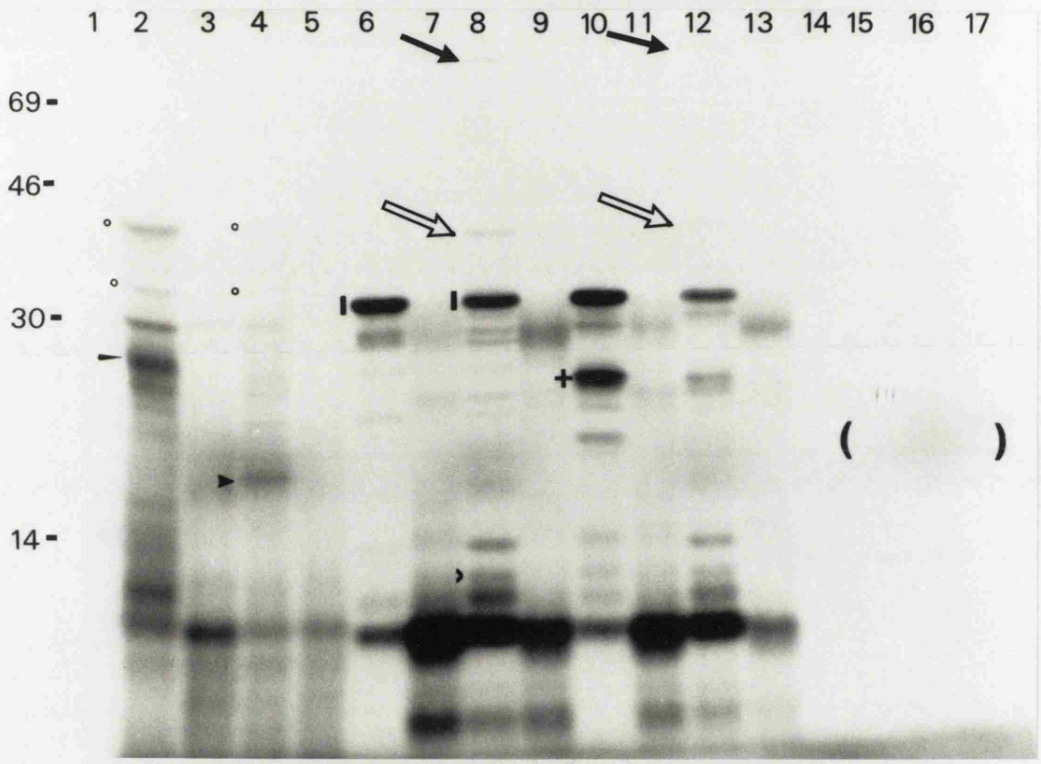
The rate of incorporation of ^{35}S-methionine into protein was determined for the *E.coli* and *P.putida* extracts. Figures 22 and 23 show that the extracts incorporated ^{35}S-methionine with similar kinetics, with the majority of incorporation of radiolabel taking place in the first 20 minutes of incubation, after which time the rate quickly diminished to zero. Whether this halt in protein synthesis is due to the condition of the DNA, for example it may become restricted, or whether the extract has lost the capacity for protein synthesis after this time is unclear. The possibility that the halt in protein synthesis was due to the inability to incorporate ^{35}S-methionine into protein after incubation of longer than 20 minutes was investigated using plasmids pKT231 and pGB123 in the S30 extracts of *E.coli* and *P.putida*. The S30 extracts were incubated without DNA for 30 minutes then either pKT231 or pGB123 was added and the incubation extended for a further 30 minutes. Alternatively, the extracts were incubated first with one plasmid then the other added and the

FIGURE 21 Polypeptides directed by *E.coli* and broad host range plasmid DNA in S30 extracts of *E.coli* and *P.putida* generated by French Press and alumina grinding.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. pGB123 was incubated with an *E.coli* French Press-derived S30 extract (Track 2), *P.putida* French Press-derived S30 extracts P1 and P1' (Tracks 3 and 5 respectively) and the alumina-ground *P.putida* S30 extract (P2) (Track 4). R300B was incubated with an *E.coli* French Press-derived S30 extract (Track 6), *P.putida* French Press-derived S30 extracts P1 and P1' (Tracks 7 and 9 respectively) and the alumina-ground *P.putida* S30 extract (P2) (Track 8). pGSS33 was incubated with an *E.coli* French Press-derived S30 extract (Track 10), *P.putida* French Press-derived S30 extracts P1 and P1' (Tracks 11 and 13 respectively) and the alumina-ground *P.putida* S30 extract (P2) (Track 12). The *E.coli* French Press-derived S30 extract, the *P.putida* French Press-derived S30 extracts P1, P1' and the alumina-ground S30 extract P2 were incubated without DNA (Tracks 14, 15, 17 and 16 respectively).

The open circles sre the cloned gene products of pGB123. The *P.putida*-specific 73kD polypeptide (RepB*) is denoted by the solid arrows and the 37kD polypeptide (RepB) is denoted by the open arrows. The arrowhead in Track2 indicates the 27kD polypeptide uniquely synthesises by pGB123 in the *E.coli* S30 extract. The arrowhead in Track 4 indicates the 17kD polypeptide uniquely synthesised by pGB123 in the *P.putida* alumina-ground S30 extract (P2). The arrowhead in Track 8 indicates the 11kD polypeptide synthesised by R300B. The cross in Track10 indicates the 23kD synthesised by pGSS33. The bars in Tracks 6 and 8 indicate the polypeptides apparently synthesised to the same level by R300B in both the *E.coli* extract and the alumina-ground *P.putida* extract (P2). The brackets indicate the *Pseudomonas*-specific DNA-independent product. Track 1 contains molecular weight standards.

FIGURE 21



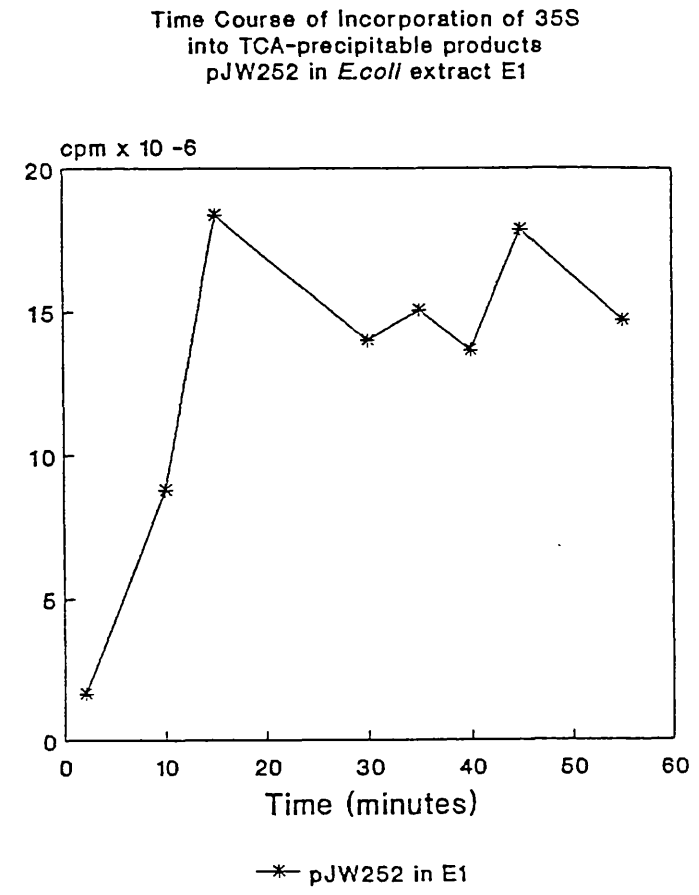


FIGURE 22

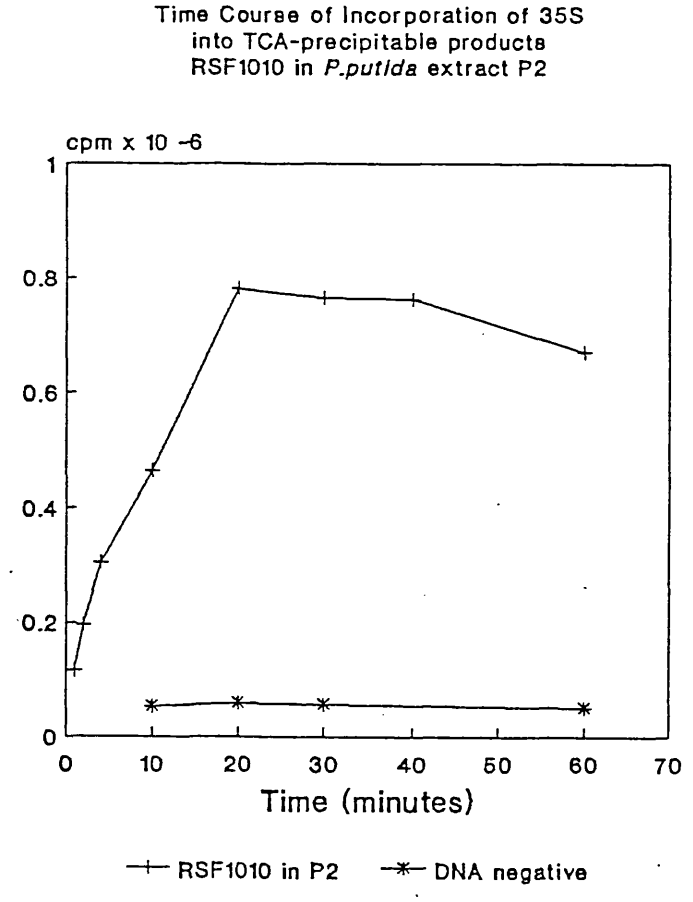


FIGURE 23

incubation extended. The incorporation of radiolabel into TCA-precipitable products is shown in Figure 24. The preparation of pGB123 directs protein synthesis to a much higher level than does that of pKT231. The preparation of pKT231 probably contained some inhibitor of protein synthesis, for example some residual ethidium bromide. When pGB123 was incubated with the *E.coli* extract, ^{35}S-methionine incorporation was high (Figure 24, Bar 1), whereas pKT231 incubated in the *E.coli* extract gave very low incorporation of radiolabel (Figure 24, Bar 2). When the *E.coli* extract was incubated first with pGB123 and then pKT231 and incubation continued, a level of incorporation of radiolabel was obtained (Figure 24, Bar 3) that was close to that of pGB123 incubated alone (Figure 24, Bar 1). When the *E.coli* extract was incubated first with pKT231 then pGB123 added and the incubation extended a level of incorporation of radiolabel was obtained (Figure 24, Bar 4) that was only approximately the same as for pKT231 alone (Figure 24, Bar 2). This suggests that the *E.coli* extract cannot promote protein synthesis from DNA after 30 minutes of incubation and this was further reinforced by the fact that when the extract was incubated in the absence of DNA for 30 minutes then plasmid DNA added and the incubation extended there was no incorporation either from pKT231 or from pGB123 (Figure 24, Bars 5 and 6 respectively)

When the products were analysed by SDS-PAGE (Figure 25), it was seen that polypeptide products were directed by pGB123 alone (Track b) or pKT231 alone (Track c) when incubated with the *E.coli* extract for the first 30 minutes. When the DNA was added after 30 minutes preincubation no polypeptides were synthesised (Tracks d and e). When the extract was incubated with one plasmid alone then a second one added only those polypeptides directed by the first plasmid were evident (Tracks f and g).

The *P.putida* extract gave similar results. When pGB123 was added to the assay reaction at the beginning of incubation the incorporation of ^{35}S-

FIGURE 24 Double plasmid experiment, incorporation into TCA-precipitable counts.

Results are shown as cpmx10^6 for each incubation reaction which are denoted as follows: *E.coli* French Press-derived S30 extract (E1') was incubated without DNA (Bar 9) or with pGB123 (Bar 1) or with pKT231 (Bar 2). The E1' assay mix was incubated with pGB123 for 30 mminutes then pKT231 added and the incubation extended for a further 30 minutes (Bar 3). The E1' assay mix was incubated with pKT231 for 30 minutes then pGB123 added and the incubation extended for a further 30 minutes (Bar 4). The E1' assay mix was incubated for 30 minutes then pGB123 added and the incubation extended for another 30 minutes (Bar 5). The E1' assay mix was incubated for 30 minutes then pKT231 added and the incubation extended for another 30 minutes (Bar 6).Extract E1' was incubated with pGB123 and pKT231 together (Bar 7). Extract E1' was incubated with pGB123 and pKT231 separately then the incubation reactions were mixed and the incubation continued for a further 30 minutes (Bar 8).

P.putida alumina ground S30 extract (P2) was incubated without DNA (Bar 18) or with pGB123 (Bar 10) or with pKT231 (Bar 11). The P2 assay mix was incubated with pGB123 for 30 minutes then pKT231 added and the incubation extended for a further 30 minutes (Bar 12). The P2 assay mix was incubated with pKT231 for 30 minutes then pGB123 added and the incubation extended for a further 30 minutes (Bar 13). The P2 assay mix was incubated for 30 minutes then pGB123 added and the incubation extended for another 30 minutes (Bar 14). The P2 assay mix was incubated for 30 minutes then pKT231 added and the incubation extended for another 30 minutes (Bar 15) Extract P2 was incubated with pGB123 and pKT231 together (Bar 16). Extract P2 was incubated with pGB123 and pKT231 separately then the incubation reactions were mixed and the incubation continued for a further 30 minutes (Bar 17).

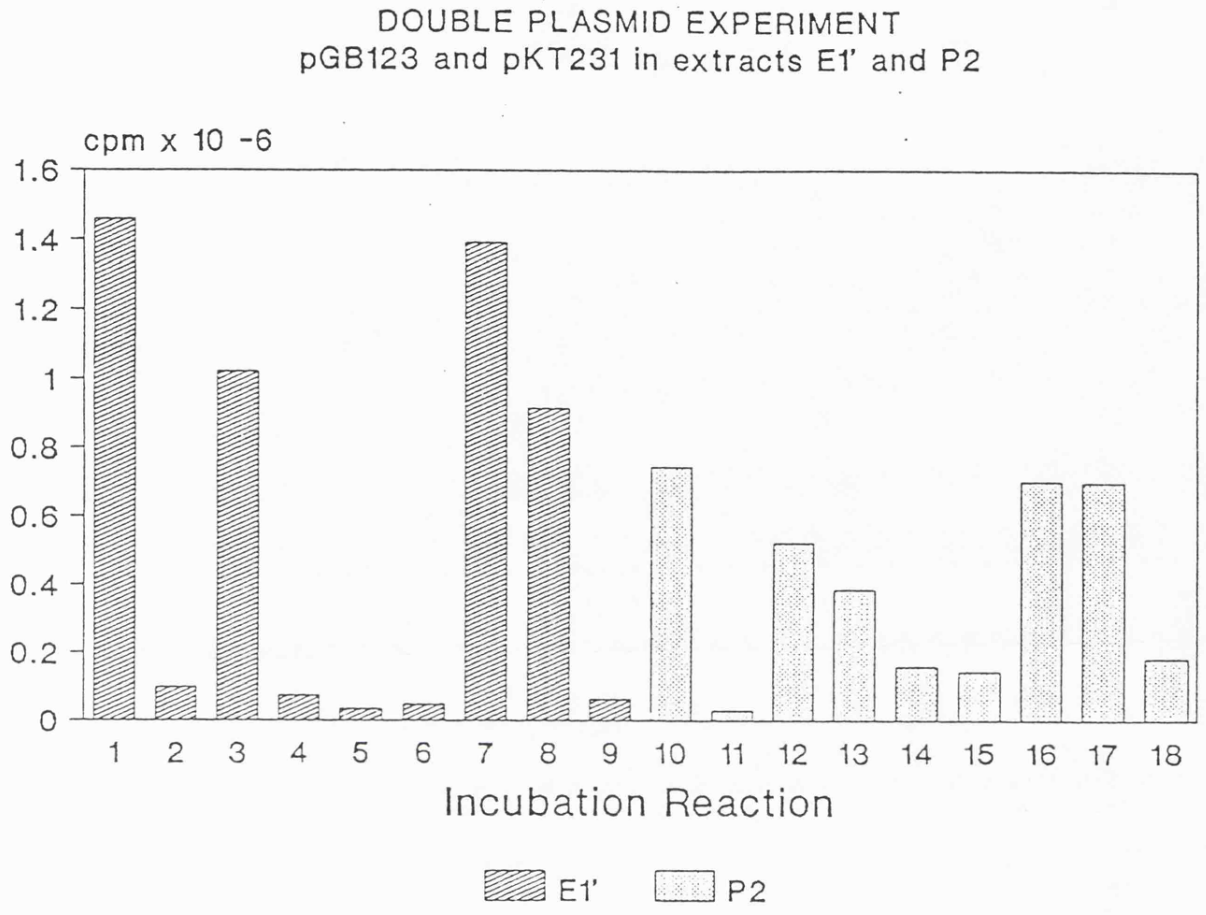

DOUBLE PLASMID EXPERIMENT
pGB123 and pKT231 in extracts E1' and P2
cpm x 10 -6
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Incubation Reaction
E1'
P2

FIGURE 24

FIGURE 25 Double plasmid experiment, analysis of labelled polypeptides by SDS-PAGE and fluorography.

E.coli French Press-derived S30 extract (E1') was incubated without DNA (Track a) or with pGB123 (Track b) or with pKT231 (Track c). The E1' assay mix was incubated for 30 minutes then pGB123 added and the incubation extended for another 30 minutes (Track d). The E1' assay mix was incubated for 30 minutes then pKT231 added and the incubation extended for another 30 minutes (Track e). The E1' assay mix was incubated with pGB123 for 30 minutes then pKT231 added and the incubation extended for a further 30 minutes (Track f). The E1' assay mix was incubated with pKT231 for 30 minutes then pGB123 added and the incubation extended for a further 30 minutes (Track g). Extract E1' was incubated with pGB123 and pKT231 together (Track h). Extract E1' was incubated with pGB123 and pKT231 separately then the incubation reactions were mixed and the incubation continued for a further 30 minutes (Track i).

P.putida alumina ground S30 extract (P2) was incubated without DNA (Track r) or with pGB123 (Track j) or with pKT231 (Track k). The P2 assay mix was incubated for 30 minutes then pGB123 added and the incubation extended for another 30 minutes (Track l). The P2 assay mix was incubated for 30 minutes then pKT231 added and the incubation extended for another 30 minutes (Track m). The P2 assay mix was incubated with pGB123 for 30 minutes then pKT231 added and the incubation extended for a further 30 minutes (Track n). The P2 assay mix was incubated with pKT231 for 30 minutes then pGB123 added and the incubation extended for a further 30 minutes (Track o). Extract P2 was incubated with pGB123 and pKT231 together (Track p). Extract P2 was incubated with pGB123 and pKT231 separately then the incubation reactions were mixed and the incubation continued for a further 30 minutes (Track q).

FIGURE 25

methionine into TCA-precipitable products was high (Figure 24, Bar 10). However, if the extract was preincubated before the addition of pGB123, the level of incorporation of ^{35}S-methionine (Figure 24, Bar 14) was close to that of the extract incubated in the absence of DNA (Figure 24, Bar 18). The results obtained from the incubation of pKT231 based on incorporation values initially appears rather anomalous. The incubation of pKT231 alone in the *P.putida* extract (Figure 24, Bar 11) gave a much lower level of incorporation of radiolabel than that of the DNA negative control (Figure 24, Bar 18). On analysis by SDS-PAGE no labelled material was evident (Figure 26, Track k) suggesting that an impurity in the incubation reaction had inhibited incorporation. When pGB123 was incubated with the *P.putida* extract for 30 minutes, pKT231 added and incubation continued, incorporation of ^{35}S-methionine (Figure 24, Bar 12) was close to that obtained for pGB123 incubated alone. However when pKT231 was incubated with the *P.putida* extract then pGB123 added and incubation continued, incorporation of radiolabel was much higher than expected (Figure 24, Bar 13) suggesting that pGB123 had directed protein synthesis after a 30 minute incubation of the extract. However when the products were analysed by SDS-PAGE (Figure 25), it was seen that the products in this case (Track o) were those directed by pKT231, not by pGB123. For some inexplicable reason pKT231 had directed protein synthesis to a high level in this particular incubation reaction. When the extract was preincubated then pGB123 added and the incubation continued, incorporation of radiolabel was close to that of the DNA negative control. When pKT231 was used instead of pGB123 incorporation was similarly low. This parallels the results from the *E.coli* extract.

The results show that after 30 minutes of incubation the protein synthesising machinery of both the *E.coli* and *P.putida* S30 extracts has shut down and addition of DNA after this time does not enhance protein synthesis. This suggests that the potential of the S30 extract to

synthesise protein was exhausted after about 30 minutes incubation at 37°C with the assay mixture.

3.2.2.2 Necessity for components of the Low Molecular Weight Mix

The components of the LMM and their concentrations were determined by Zubay (1973). However several authors subsequently noted that the addition of various components of the LMM as detailed in the original paper were not required in the *in vitro* systems developed for *E.coli* and other bacterial systems. In particular inclusion of tRNA was found not to be necessary in extracts of *B.subtilis* (Leventhal and Chambliss 1979), *S.lividans* (Thompson *et al* 1984) and *R.meliloti* (Dusha *et al* 1986) since presumably the extracts contained enough endogenous tRNA. The S30 extracts of *E.coli* and *P.putida* were investigated in this regard by incubating these extracts with plasmid DNA and LMM lacking various components and the products analysed by SDS-PAGE (Figures 26 and 27). Omission of tRNA from the assay mixture had no effect on either the level of incorporation of radiolabel or the nature and extent of polypeptide products synthesised by either the *E.coli* or *P.putida* extract. There is an absolute requirement for PEP and the CTP/GTP/UTP mix in the *P.putida* extract (Figure 26, Tracks e and d repectively). DTT, ATP, PEG and FA all have stimulatory effects on incorporation of radiolabel but do not alter the nature of the products synthesised (Figure 26, Tracks b, c, f and g respectively). The addition of Ca^{2+}, although apparently necessary for the *E.coli* extract, inhibited protein synthesis in the *P.putida* extract, such that its omission from the LMM caused an increase in incorporation of radiolabel into TCA-precipitable protein (Table 11 and Figure 27, Track i). A similar effect was observed for the *S.lividans* extract (Thompson *et al* 1984). Omission of K^+ and NH_4^+ salts or of cAMP have little effect (Figure 27, Tracks h, j and k respectively), although cAMP may be required for the activation of some genes. There may be other, as yet unknown, components which might

FIGURE 26 Necessity of LMM components for protein synthesis in *P.putida* S30 extract and effect on production of *Pseudomonas*-specific DNA-independent product of alteration in assay reactants.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* alumina-ground S30 extract was incubated with pCF32 and LMM (Track h) or LMM lacking DTT (Track b), ATP (Track c), CTP/GTP/UTP mix (Track d), PEP (Track e), PEG (Track f) or FA (Track g). The same extract was incubated without DNA with LMM (Track i), or without LMM (Track j) or without magnesium (Track k) or substituting ^{35}S-methionine with ^{35}SdATP (Track l). Track a contains molecular weight standards.

FIGURE 27 Necessity of LMM components for protein synthesis in *P.putida* S30 extract and for the production of the *Pseudomonas*-specific DNA-independent product.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* alumina-ground S30 extract was incubated with pCF32 and LMM (Track l) or LMM lacking K^+ (Track h), Ca^{2+} (Track i), NH_4^+ (Track j) or cAMP (Track k). The same extract was incubated without DNA with LMM (Track g) or with LMM lacking K^+ (Track b), Ca^{2+} (Track c), NH_4^+ (Track d), 19 amino acids (Track e) or cAMP (Track f). Track a contains molecular weight standards.

FIGURE 26

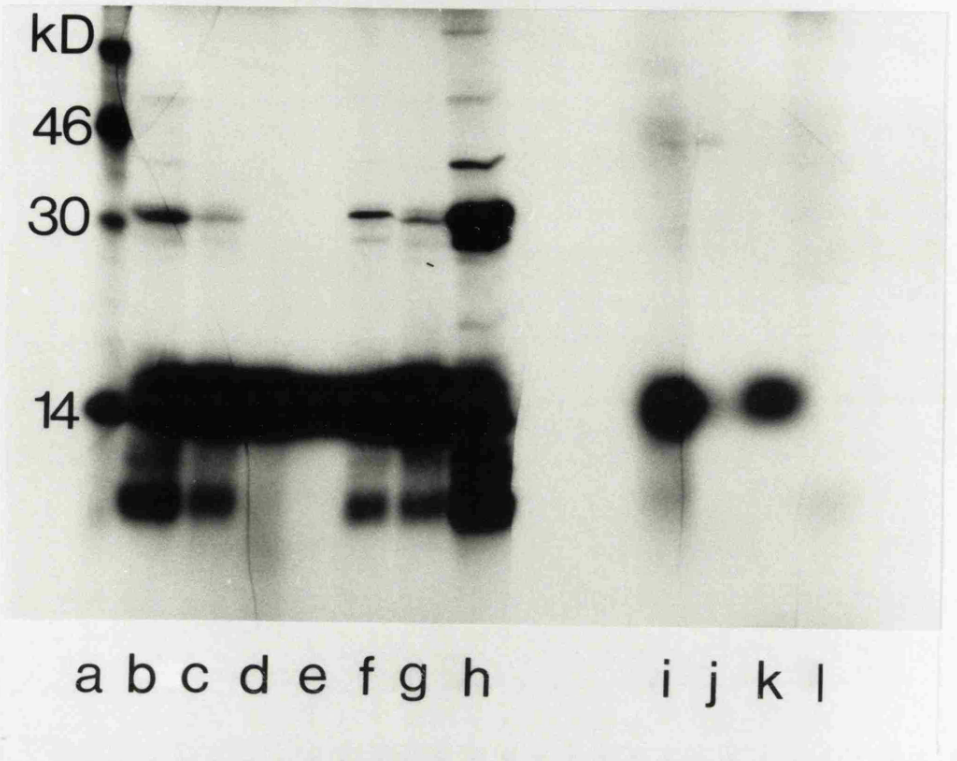


FIGURE 27

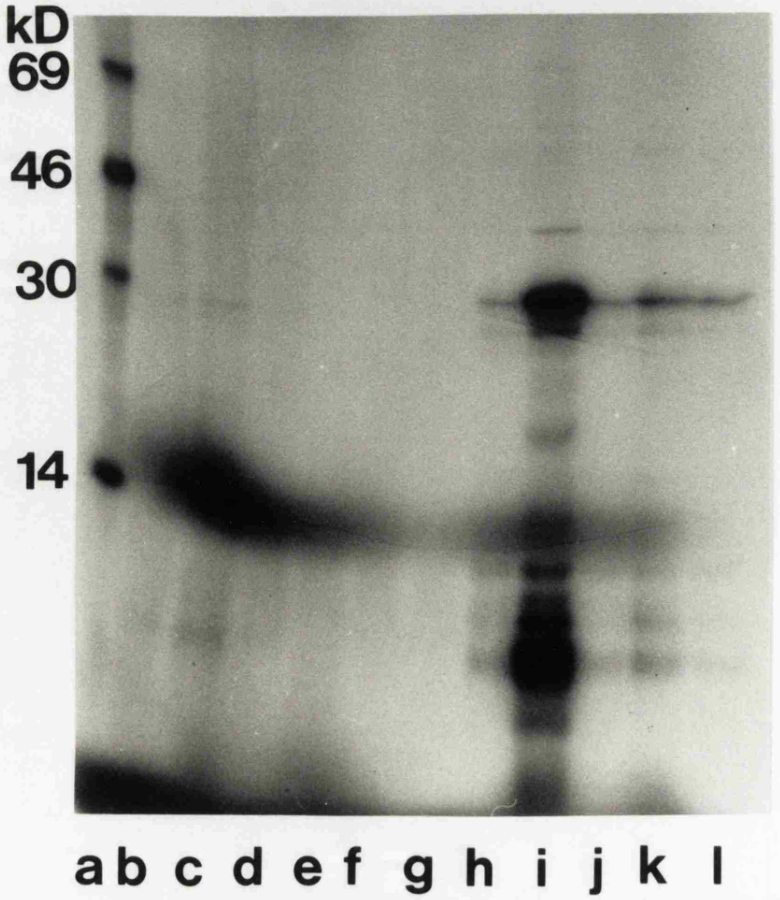


TABLE 11

EFFECT OF OMISSION OF CALCIUM FROM THE LOW MOLECULAR WEIGHT MIX ON INCORPORATION OF ^{35}S INTO TCA-PRECIPITABLE PRODUCTS IN THE *P.PUTIDA* COUPLED TRANSCRIPTION/TRANSLATION SYSTEM

PLASMID	CPMx10^{-5} per sample Mean of 3 replicates	
	$+Ca^{2+}$	$-Ca^{2+}$
EXPERIMENT I		
pGB123	5.8 ± 0.63	7.4 ± 0.28
RSF1010	0.31 ± 0.02	0.50 ± 0.03
DNA negative	0.24 ± 0.02	0.30 ± 0.01
EXPERIMENT II		
pCF32	0.56 ± 0.05	4.8 ± 0.32
DNA negative	0.42 ± 0.03	0.60 ± 0.05

specifically stimulate gene expression in the *P.putida* extract.

3.2.2.3 Protein synthesis directed by restriction fragments in the *P.putida* S30 extract

It has been reported (Pratt *et al* 1981) that S30 extracts of *E.coli* can be programmed with linear restriction fragments of DNA (rather than using supercoiled plasmid molecules) without apparent loss of sensitivity although significantly higher concentrations of exogenous linear DNA are required. Yang *et al* (1979) suggested that it was the lack of supercoiling in linear fragments which led to the observed decrease in activity, but Pratt *et al* (1981) found that although supercoiling may have had some effect it is not a major limiting factor and the reduction was more likely to be due to the exonucleolytic degradation of the exogenous DNA. Yang *et al* (1980) and Jackson *et al* (1983) have both reported that coupled transcription/translation systems generated from *E.coli* strains lacking exonuclease V activity (*recB* mutants) substantially improved the efficiency with which linear fragments of DNA promoted protein synthesis compared with systems based on *recB*$^+$ strains. This suggests that the reason for decreased activity with linear fragments compared with circular DNA was that they became rapidly degraded by exonuclease V on incubation with the extract. Chen and Zubay (1983) have used an S30 extract derived from a *recB* mutant strain of *E.coli* to analyse the coding capacity of restriction fragments of the ColE1 plasmid.

When the *P.putida* S30 extract was programmed with linearised R300B, incorporation of radiolabel into protein was only approximately 50% of the value obtained for circular DNA. The most efficient method proved to be isolation of restriction fragments of a plasmid from an agarose gel followed by electroelution of the DNA and removal of the ethidium bromide by extraction with water-saturated butanol (failure to do this results in total inhibition of protein synthesis). This method was used to assign

certain polypetides to particular regions on the R300B molecule (results are presented in Chapter 4).

3.2.3 Protein concentration and polypeptide synthesis

The protein concentrations of the S30 extracts (Table 12) indicated that protein concentration was not directly related to extract efficiency in terms of incorporation of radiolabel into TCA-precipitable polypeptides. The most active extract, the French Press-derived *E.coli* extract (E1), had a protein concentration of only 2.85 $mgml^{-1}$ compared with the *P.putida* extract (P2) a less active extract with a protein concentration of 5.74 $mgml^{-1}$. Dusha *et al* (1986) report that an active extract of R.meliloti had a protein concentration of 7-10 $mgml^{-1}$.

3.3 ATTEMPTS AT PRODUCTION OF AN ACTIVE *P.AERUGINOSA* S30 EXTRACT

The possibility that the method described above for the production of a protein synthesising extract of *P.putida* could be a general method for the generation of such systems based on other *Pseudomonas* species was investigated by an attempt to derive an active S30 extract from *P.aeruginosa*. Both methods of cell breakage used to derive similar extracts from *P.putida* were employed and two different strains were used. An extract of PAO1161 (*leu*$^-$, restriction$^-$) was prepared by the French Press method. This extract programmed with DNA template incorporated radiolabel in a magnesium-dependent manner, but the maximum incorporation was only 1.2x the background incorporation in the absence of exogenous DNA. This may reflect poor cell disruption. In an attempt to increase the efficiency of cell breakage an extract was derived from the same strain by two passages through the French Press. This, however, gave no incorporation over the background level. The added circular plasmid DNA could be recovered from both of these extracts following incubation confirming that it had not been restricted or degraded.

TABLE 12

PROTEIN CONCENTRATIONS OF S30 EXTRACTS

Extract	Protein Concentration ($mgml^{-1}$)
E.coli MRE600 E1	2.85
E.coli MRE600 E1'	2.81
E.coli MRE600 E2	3.57
P.putida KT2440 P1	4.85
P.putida KT2440 P2	5.74
P.putida KT2440 P1'	5.06

Several extracts were made from PAO1, a wild type strain of *P.aeruginosa*. The French Press method yielded an extract which showed DNA-dependent incorporation of radiolabel and gave a maximum incorporation of three fold over DNA negative background values (this was the highest activity obtained for any of the extracts) (Figure 28). The method of alumina grinding gave extracts which incorporated radiolabel to a lower level.

The most active extract was subjected to further investigation. Maximum incorporation of radiolabel at optimum magnesium concentration (Figure 28) was only about one tenth of that for the *P.putida* extracts and even though this was 3 times the background level it was not sufficient to allow analysis of the products by SDS-PAGE. It was conceivable that this *P.aeruginosa* extract (PA) was preventing protein synthesis by, for example, the added exogenous DNA being restricted or degraded or a component present specifically interfering with the coupled transcription/translation reaction. This was investigated by assaying coupled transcription/translation directed by pAT153 in the S30 extracts of *P.putida*, *P.aeruginosa* and *P.putida* mixed with *P.aeruginosa*. If the *P.aeruginosa* S30 extract was inhibitory to protein synthesis then the presence of the *P.putida* extract should not increase incorporation of ^{35}S-methionine. However the results (Figure 29) indicate that protein synthesis does take place in the mixed *P.aeruginosa* and *P.putida* extracts. Most polypeptides synthesised in this mixture of extracts were identical to those made in the *P.putida* extract alone, but there appears to be processing taking place, suggesting that the *P.aeruginosa* extract contains some membranous material. In addition to showing that the *P.aeruginosa* extract is not inhibitory to protein synthesis, it also indicates that the exogenous plasmid DNA is not being restricted, since incorporation of ^{35}S-methionine occurs to almost the same level in the mixture of extracts as in the *P.putida* extract. Omission of calcium from the LMM did not

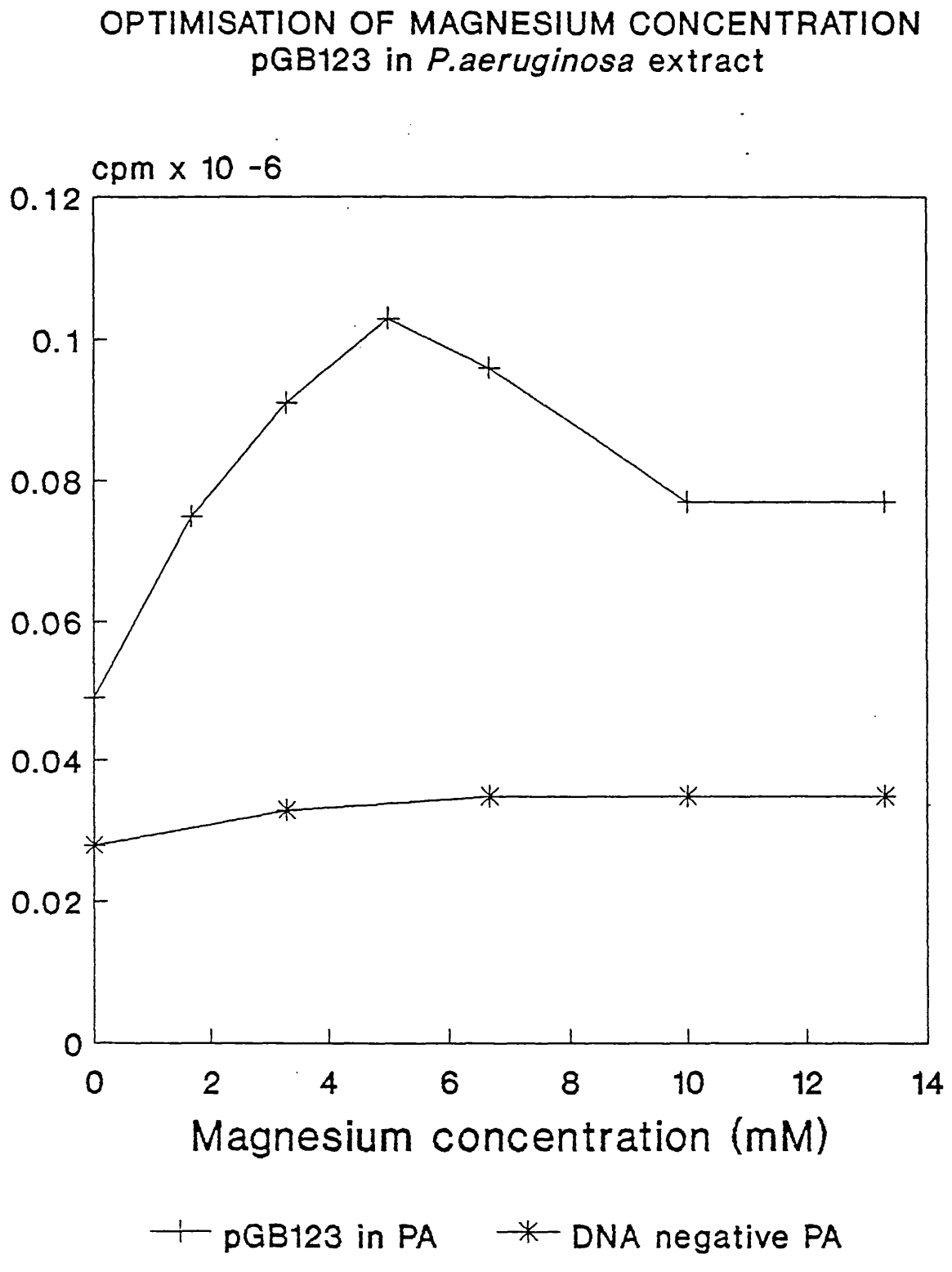

OPTIMISATION OF MAGNESIUM CONCENTRATION
pGB123 in P.aeruginosa extract
cpm x 10 -6
0.12
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
8
10
12
14
Magnesium concentration (mM)
pGB123 in PA
DNA negative PA

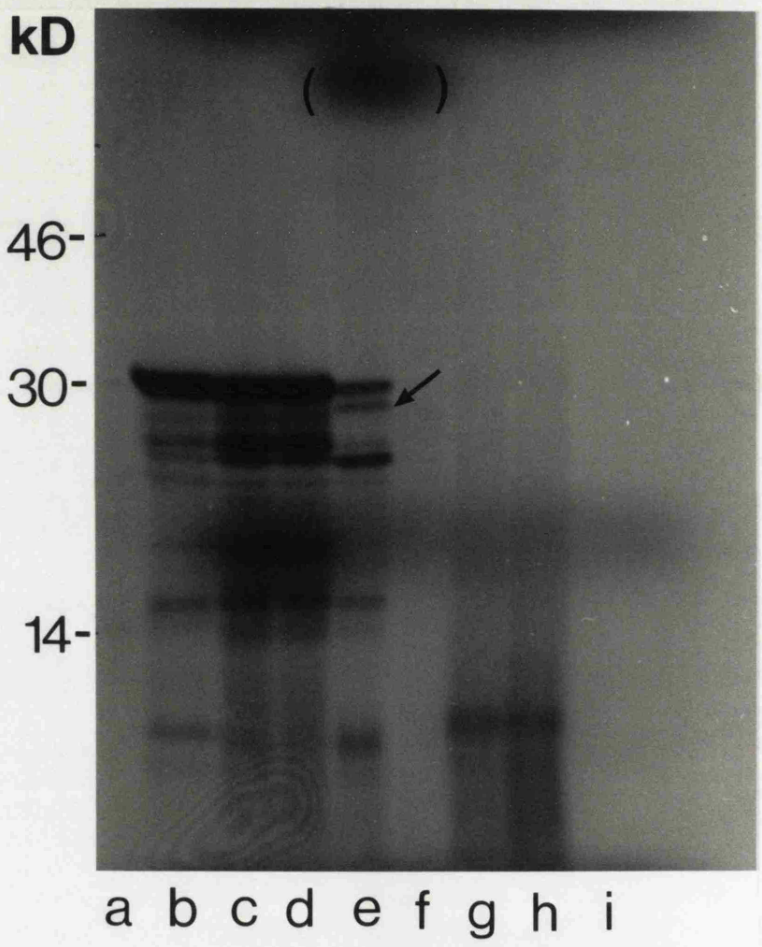
FIGURE 28

FIGURE 29 Polypeptides synthesised by pAT153 in S30 extracts of *E.coli*, *P.putida* and *P.aeruginosa* and polypeptides synthesised by pGB123 in French Press-derived S30 extract of *P.putida*.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. pAT153 was incubated with the *E.coli* S30 extract (Track b), the *P.putida* S30 extract (Tracks c and d) or with *P.putida* S30 extract mixed with *P.aeruginosa* S30 extract (Track e). The mixed extracts of *P.putida* and *P.aeruginosa* were also incubated without DNA (Track f). Track a contains molecular weight standards. The arrow indicates probable processed polypeptide in mixed extracts of *P.putida* and *P.aeruginosa* (Track e). Brackets indicate the *P.aeruginosa*-specific DNA-independent product.

P.putida French Press-derived S30 extract (Pl') was incubated at optimal magnesium concentration with pGB123 (Tracks g and h) and without DNA (Track i).

FIGURE 29



stimulate incorporation of radiolabel as was the case with the *P.putida* extract.

Reasons for the inability to produce an active protein synthesising extract from *P.aeruginosa* are unclear. The possibility that the cell wall of *P.aeruginosa* may be more refractory to breakage than either *E.coli* or *P.putida* was investigated by disrupting the *P.aeruginosa* cells at a higher pressure (14psi) with the French Press than adopted for the *E.coli* and *P.putida* cells (10psi). However this had no stimulatory effect on protein synthesis in the *P.aeruginosa* extract. The measurement of protein concentrations of the S30 extracts indicated that those of *P.aeruginosa* were within the same range as those of *E.coli* and *P.putida*. Alternatively some component of the LMM may be inhibitory or another unknown component may be required for maximal activity of a *P.aeruginosa* extract but not required for the *E.coli* or *P.putida* extracts. This needs further investigation by determination of the effects of omission of components from the LMM and alterations in the concentrations of the LMM components.

3.4 THE *PSEUDOMONAS*-SPECIFIC DNA-INDEPENDENT PRODUCT

P.putida extracts (both P1 and P2) incorporate ^{35}S into a product which forms a diffuse band on polyacrylamide gels (Figure 8, Tracks b-i and k-o). The nature of this product is unclear but its formation is not dependent upon the addition of exogenous DNA. It also exhibits varying electrophoretic mobility when gels of different concentrations of acrylamide are used. For example it had an apparent molecular weight of 17kD on a 14% gel and 40kD on an 11% gel. Its appearance is not diminished by preincubation of the extract prior to addition of ^{35}S-methionine indicating that it does not represent the presence of long-lived mRNA in the extract, nor is it RNA since it is not destroyed by addition of RNase to the incubation mix.

To test whether the appearance of this product was related to protein

synthesis, chloramphenicol (an inhibitor of protein synthesis) was included in the assay mix at a final concentration of $50\mu gml^{-1}$ and extracts incubated with and without exogenous DNA. Addition of pKT231 with and without Cm was carried out as a control. The results (Figure 30) show that Cm inhibited pKT231-dependent protein synthesis, but it did not prevent formation of the diffuse product. In addition, the formation of this product did not require the 19 amino acids normally used in the assay (Figure 27, Track e) thus its formation is not dependent on the incorporation of amino acids into protein.

In addition to occurring in the *P.putida* S30 extracts, the diffuse product was also formed in every *P.aeruginosa* extract made. In these *P.aeruginosa* extracts two other similar DNA-independent products were also synthesised. The product is not synthesised in the *E.coli* extract. An experiment was carried out to determine the effect of the presence of the *E.coli* extract on the production of this product by the *P.putida* extract. Reactions of the *E.coli* extract and *P.putida* French Press-derived extract (P1) were set up without the addition of any DNA (Figure 31) and incubated for 30 minutes or the *P.putida* extract (P1) was incubated alone for 30 minutes then the *E.coli* extract was added and the incubation extended for another 30 minutes (Track b), or the two extracts were incubated together for varying lengths of time (Tracks d-i). The results were assessed by analysing the presence or absence of the product by SDS-PAGE. When the *E.coli* extract was added to the *P.putida* extract, the product formed by the *P.putida* extract subsequently disappeared. The time course shows that the product was not formed when the *E.coli* extract was present from the start of the incubation indicating that the presence of the *E.coli* extract prevented its formation. These results show that the *E.coli* extract does not synthesise the product and it has the ability both to degrade the product once it is formed in the *P.putida* extract and also to prevent its formation.

FIGURE 30 The effect of chloramphenicol on synthesis of the *Pseudomonas*-specific DNA-independent product.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was incubated with pKT231 in the presence of Cm (Track b) or absence of Cm (Track d), or without DNA in the presence of Cm (Track c) or absence of Cm (Track e). Brackets indicate *Pseudomonas*-specific DNA-independent product. Track a contains molecular weight standards.

FIGURE 31 The effect of the presence of the *E.coli* S30 extract on the synthesis of the *Pseudomonas*-specific DNA-independent product by the *P.putida* S30 extract.

All incubations were carried out in the absence of the addition of plasmid DNA. Labelled material was analysed by SDS-PAGE and fluorography. *P.putida* French Press-derived S30 extract (P1) was incubated for 30 minutes then *E.coli* French Press-derived S30 extract was added and the incubation continued for a further 30 minutes (Track b) and as a control P1 was incubated for 30 minutes then DEPC-water added and the incubation extended (Track c). The *E.coli* and *P.putida* extracts were incubated together for times varying from 2 to 30 minutes (Tracks d-i). The *E.coli* extract was incubated alone for 30 minutes (Track j) and the *P.putida* extract was incubated alone for 30 minutes (Track k). Track a contains molecular weight standards.

FIGURE 30

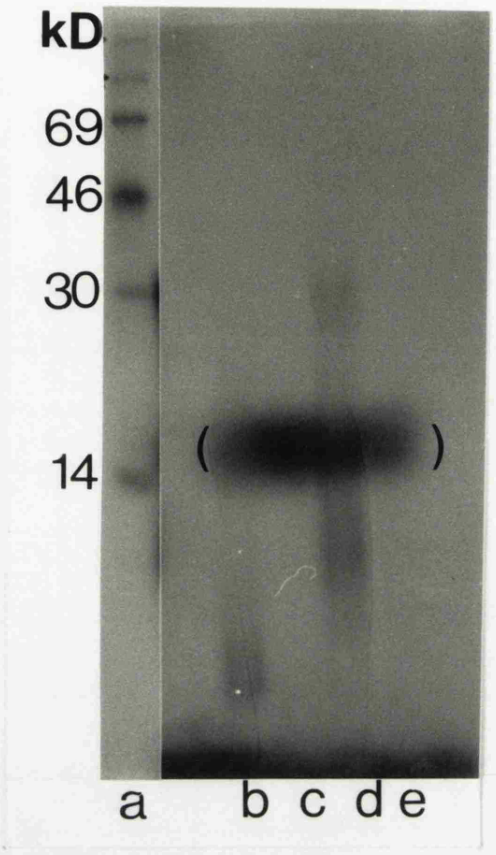
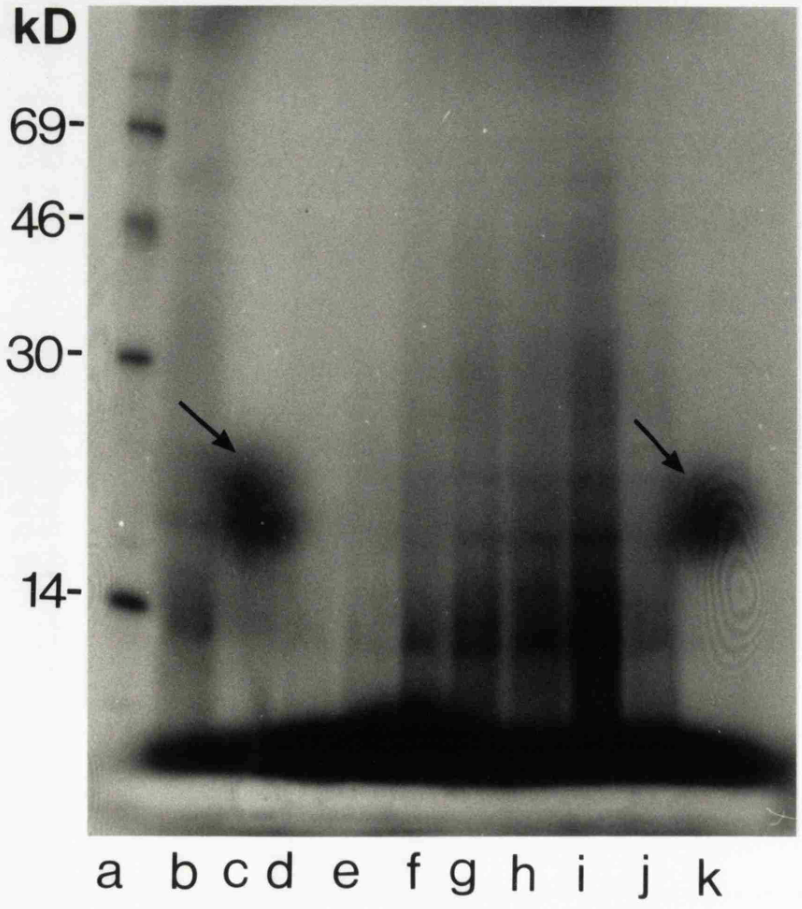


FIGURE 31



A further experiment was carried out to determine which of the components of the LMM were required for the formation of this diffuse product. Omission of K^+, Ca^{2+}, NH_4^+, 19 amino acids or cAMP from the LMM had no effect on the formation of the product (Figure 27, Tracks b, c, d, e and f respectively). Omission of Ca^{2+} appeared to stimulate its formation. (Omission of Ca^{2+} had been found to stimulate protein synthesis when the *P.putida* extract was programmed with template plasmid DNA.) However omitting the LMM entirely inhibits its formation (Figure 26, Track j) and omitting magnesium gives lower incorporation of ^{35}S-methionine into the diffuse product (Figure 26, Track k). If ^{35}S-dATP is used in place of ^{35}S-methionine at approximately the same specific activity (chasing with cold ATP instead of methionine), there is no formation of the product (Figure 26, Track l). This implies that the ^{35}S-labelled group from the methionine or the whole labelled methionine molecule becomes attached to an amorphous substance in the extract. The *E.coli* extract appears to possess an activity, not present in the *P.putida* extract, which can remove that labelled group from the diffuse product.

DNA-independent incorporation of radiolabel into TCA-precipitable material in the *E.coli* system had been noted by Pratt *et al* (1981). These had molecular weights of 60kD, 18kD and 16kD. The appearance of the largest was strain-dependent and its formation disappeared following storage of the extract. This is in contrast to the anomalies encountered here, since the diffuse product was not formed in the S30 extracts of *E.coli* strain MRE600 but was synthesised in both the *P.putida* and *P.aeruginosa* S30 extracts.

3.5 SUMMARY OF RESULTS

Production of an *in vitro* system for coupled transcription and translation from *P.putida* was unsuccessful using the method for production of an S30 extract described for *E.coli*. Breakage of the cells by the French Press lead to the synthesis of products, some of which may have been

degraded. Other products may have been the result of abortive transcription or translation. No products greater than 30kD were synthesised by this extract, although such products were synthesised in the comparable *E.coli* system.

An *in vitro* system of *P.putida* in which apparent faithful transcription and translation takes place has been derived by disruption of cells with alumina using a pestle and mortar. No extensive proteolysis was observed and products of greater than 30kD were synthesised. An *E.coli* extract made by this method synthesises identical polypeptides as in the French Press-derived *E.coli* extract, indicating the applicability of this method for production of a coupled transcription/translation system. Stimulation of protein synthesising activity in the *P.putida* extract was achieved by the omission of calcium ions from the LMM.

CHAPTER 4

FURTHER ANALYSIS OF THE *P.PUTIDA* COUPLED TRANSCRIPTION/TRANSLATION SYSTEM; PROGRAMMING WITH *E.COLI* AND BROAD HOST RANGE PLASMID DNA.

4.1 BROAD HOST RANGE

Plasmids are usually classified according to their incompatibility group; two plasmids which cannot be maintained in the same cell are said to be incompatible and thus belong to the same incompatibility group. This is considered to be a natural classification dependent on the evolution of replication systems since those plasmids which have identical replication systems will be incompatible. The plasmids of Gram-negative bacteria have been assigned to about 20 incompatibility groups. Incompatibility is related to host range and the ability to replicate in specific hosts. For example the F plasmid of *E.coli* incompatibility group FI will only replicate in the Enterobacteria and the FP2 plasmid of *Pseudomonas* incompatibility group P-8 will only replicate in *Pseudomonas*. Such plasmids are said to have a narrow host range. Conversely, some plasmids have a broad host range specificity particularly those of *E.coli* incompatibility groups P and Q, and can be maintained in several unrelated hosts.

For a native plasmid to exhibit a broad host range it must possess functions for some element of mobility; either it must be self-transmissible or it must be able to be mobilised from one host to another if the appropriate conjugation functions are provided *in trans*. The nature of broad host range is not yet understood, but there are suggestions that it may be determined by the replication system rather than by the transfer system, since it has been found (Guiney 1982) that the narrow host range plasmid F_{lac} can transfer (albeit at a low frequency compared with RK2) to *P.aeruginosa*, but its inability to be detected in this host is due to the inability of the plasmid to replicate.

The broad host range plasmids whose replication systems are being

analysed are the IncQ plasmid, RSF1010, the IncP plasmid, RK2, the IncX plasmid, R6K and the IncW plasmid, pSa. Although the replication proteins (all of which are about 30-40kD) from these different plasmids show little homology (Tait *et al* 1983b), the directly repeated hexanucleotides present at the origin of replication of R6K, RK2 and RSF1010 are highly homologous (Persson and Nordstrom 1986). These sequences may be the binding sites for the replication proteins.

4.1.1 RSF1010/R300B/R1162

RSF1010 (Guerry *et al* 1974) and R300B (Barth and Grinter 1974) were isolated from *Salmonella* and R1162 (Meyer *et al* 1979) from *P.aeruginosa* but they are probably identical. They belong to the IncQ incompatibility group and encode resistance to streptomycin and sulphonamide. They can replicate and be maintained in more than 50 Gram negative species (Barth *et al* 1981, Bagdasarian and Timmis 1982). The copy number of the plasmid in *E.coli* is 8-12 (Barth and Grinter 1974). The RSF1010 molecule is fairly well characterised having been completely sequenced (8685 bp) and 11 of the ORF's identified as functional protein products (E.Scherzinger, Personal Communication). The major interests of investigators of this plasmid are that of the mechanism of mobilisation and its replication functions. Unlike the narrow host range plasmids, the replication functions of RSF1010 cover a large sector of the genome. Scherzinger *et al* (1984) showed that the replication of RSF1010 *in vitro* depends on three replication proteins encoded by the plasmid and is independent of rifampicin-sensitive RNA polymerase (encoded by *rpoB*), one of the host proteins involved in the initiation of replication of several narrow host range plasmids. Further to this Scholz *et al* (1985) have shown that other host replication functions encoded by *dnaB*, *dnaC*, and *dnaG* are not required and also that the RSF1010 *rep* proteins cannot substitute for the host replication functions.

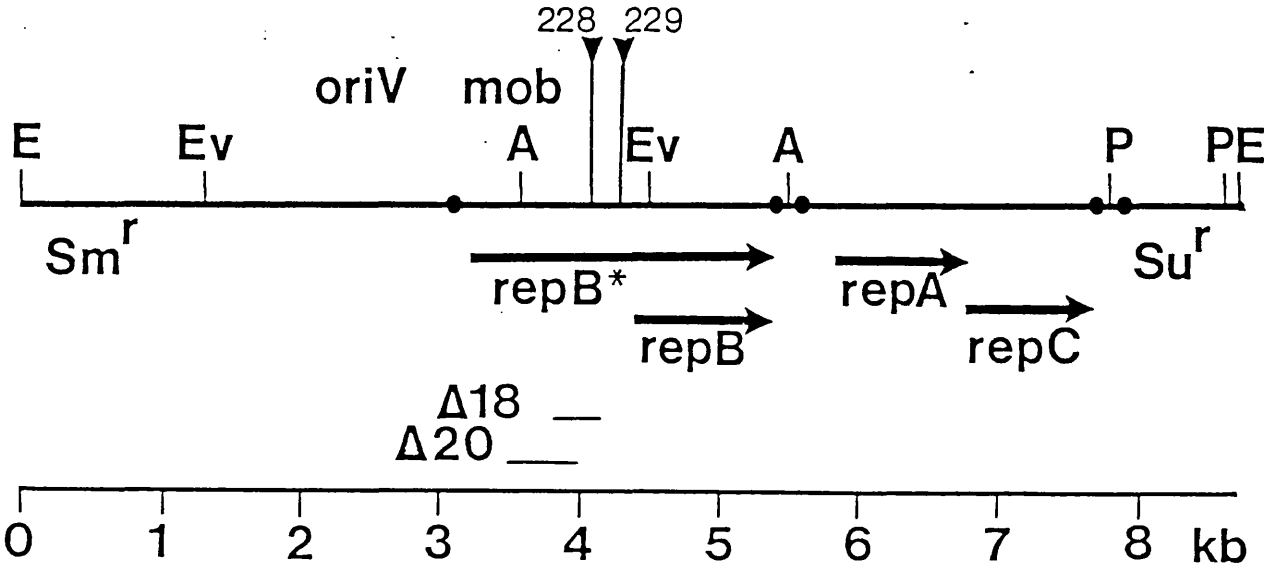
The replication proteins of RSF1010 have been denoted as *repA*, *repB* and

repC (Figure 32) and by cloning them on controlled expression vectors polypeptides of 29, 38 and 27kD have been directed from these genes respectively (Haring *et al* 1985). RepA is thought to be active as a hexamer. It has ATPase activity associated with single stranded DNA and it has sequence homology to phage P22 gene 12. Thus it may substitute for DnaB and act as a helicase. RepC has been shown to bind to double stranded DNA as a dimer specifically to the *ori*V region. *repC* controls copy number of the plasmid (Haring *et al* 1985) by regulating the frequency with which RSF1010 replication is initiated. The C-terminus of RepB is thought to be the primer for replication (E.Scherzinger, Personal Communication).

The replication genes are controlled by autoregulation at their promoter sites. The *rep*B gene is situated in the same translational reading frame of a larger gene, *rep*B*, and they are both transcribed from a promoter upstream of *rep*B*. One RNA polymerase binding site was identified here (Bagdasarian *et al* 1986), but in this region three overlapping promoters have been determined by nucleotide sequence analysis (Derbyshire *et al* 1987). This region also encodes *ori*T, the origin of plasmid transfer. Two of the promoters (p1 and p3) give rise to transcription in the same direction as *rep*B*/B, and one of these, p1, which has good similarity to the *E.coli* concensus sequence is the promoter which allows transcription these genes. Activity from p3 was only detected in mobilisation-deficient RSF1010 mutants which have mutations in the N-terminal region of the *mob*A (*rep*B*) or *mob*C genes. Presumably the products of *mob*A and *mob*C bind to the *ori*T region to promote mobilisation. Mutations in the N-terminus of the coding region must prevent their ability to bind to *ori*T thus allowing transcription from p3. The third promoter, p2, acts in the reverse orientation and leads to transcription of the *mob*C gene which encodes a 10kD polypeptide.

From S1 nuclease mapping data (Bagdasarian *et al* 1986) the *rep*A and *repC* genes are known to be expressed both from the promoter region upstream

FIGURE 32



RSF1010/R300B

FIGURE 32 Restriction map of RSF1010/R300B linearised at the unique *Eco*RI site.

Replication genes are denoted *rep*A, *rep*B, *rep*B* and *rep*C and their direction of transcription is indicated. Filled circles represent RNA polymerase binding sites. The location of the TSF1010 Tn3 insertions in pKT228 and pKT229 are shown, 228 and 229 respectively. The location and approximate size of the deletions in RSF1010Δ18 and RSF1010Δ20 are shown. Positions of the streptomycin (Sm) and sulphonamide (Su) resistance genes are indicated. Restriction sites are denoted E, *Eco*RI; Ev, *Eco*RV; A, *Acc*I; P, *Pst*I. The origin of vegetative replication is denoted by oriV.

of *rep*B* as a polycistronic message and as a shorter transcript originating from a promoter upstream of *rep*A. The activity of this promoter is governed by a repressor which is the product of a gene located just upstream of *rep*A. This protein can bind specifically to the sequences of the *rep*A/C promoter.

4.1.2 RK2/RP1/RP4

RK2, RP1, RP4, R18 and R68 are thought to be identical. They are members of incompatibility group P, are all 60kb and express resistance to ampicillin, kanamycin and tetracycline and they were all isolated as the causal agent transferring carbenicillin resistance in the same hospital burns unit. RK2 is the plasmid which has been most extensively studied. It is self-mobilisable and can replicate in a wide variety of Gram negative bacterial species (Olsen and Shipley 1973; Windass *et al* 1980). It has a low copy number, 3-5 in *E.coli* (Figurski *et al* 1979). The large size of these plasmids, compared with RSF1010, is due to the transfer functions which they carry. It has a complicated replication system although only 2 loci are essential for replication of RK2 (*trf*A and *ori*V) (Figurski *et al* 1979, Thomas *et al* 1980). Replication from *ori*V is dependent both on host functions and on transacting functions encoded by the gene *trf*A on RK2. There is a series of RK2 genes which are potentially lethal to the *E.coli* cell (Figurski *et al* 1982). These *kil* genes are normally suppressed by the presence of the *kor* (kill-override) genes, each *kil* gene having a corresponding specific *kor* gene. *kil* A, B and C, when not controlled by *kor,* are host lethal, whereas *kil*D is inhibitory to plasmid maintenance. The function of these genes is unknown but it is thought that they may function in plasmid maintenance (Figurski *et al* 1982).

In RK2 one of the replication genes, *trf*A, has the capacity for production of two polypeptides, of 32kD and 43kD (Shingler and Thomas 1984, Smith and Thomas 1983). These proteins are encoded by two inphase over-

lapping ORF's. Deletion of parts of the coding region for the larger polypeptide, but not the smaller, substantially affected plasmid maintenance in *P.aeruginosa*, but not in *E.coli*, *P.putida*, *R.meliloti*, *A.tumefaciens* or *Azotobacter vinelandii* (Durland and Helinski 1987).

4.2 BROAD HOST RANGE PROMOTER SIGNALS

The finding that many *Pseudomonas* genes are poorly expressed in the heterologous host suggests that broad host range plasmids must possess some novel expression signals which allow them to replicate in such a wide range of species. Possibilities for the nature of broad host range include:

(a) a different set of genes may be expressed in different hosts, each set having its own promoter sequences that are recognised by the host RNA polymerase. This seems quite unlikely in view of the large number of hosts that can support replication particularly of the IncP and IncQ plasmids. However there is some indirect evidence which supports this notion since there are an unusually large number of overlapping genes in the replication/mobilisation region of RSF1010 (Derbyshire *et al* 1987). This lays open the possibility that different products may be required in different hosts.

(b) the same set of genes are expressed in different hosts but there are different promoters recognised by the different host polymerases.

(c) a further possibility is that there are generalised promoters adapted to function in a variety of hosts.

There is little published data on broad host range promoter signals. The two plasmids which have been investigated most in this regard are RK2 and RSF1010. The promoters of several of the genes involved in replication and maintenance in RK2 have been located by RNA polymerase experiments and sequenced (Smith *et al* 1984). It appears that *trfA*, *trfB* and *kilB* possess typical *E.coli* promoter sequences. The promoter sequence of *trfA* has been further investigated to determine whether this sequence is responsible for

initiation of transcription in *E.coli, P.aeruginosa* and *P.putida* (Pinkney *et al* 1987). They found that the 5' end of the *trf*A mRNA is identical in all three species, indicating that the same region of DNA must encode the promoter in all of these species. To determine whether different base sequences within this promoter sequence are important for the level of transcription initiation in the different species, a series of plasmids with base substitutions in this region were introduced into each species and the level of promoter activity measured. The base substitutions had the same effect in all three species. The T to C transition in one of the highly conserved T's in the -10 region greatly reduced promoter activity, as would be expected for a functional *E.coli* promoter. A further T to C transition outside the -10 region had no additional effect. This implies that this promoter region is active in all three species.

On RSF1010 5 promoters have been identified as RNA polymerase-binding sites (Bagdasarian *et al* 1981) (Figure 32) and nucleotide sequence data has confirmed this in terms of the similarity to the *E.coli* concensus promoter sequence (Scholz *et al* 1985; Bagdasarian *et al* 1986). Within the replication region three RNA polymerase-binding sites were identified. One of the promoters (p1) upstream of *rep*B*/B, identified by nucleotide sequence analysis (Derbyshire *et al* 1987) has the ability to promote transcription from all the replication genes, but there appears also to be initiation from the promoter upstream of *rep*A (Bagdasarian *et al* 1986). Both these promoters have *E.coli* promoter-like nucleotide sequences.

This limited data indicates that there are *E.coli*-like promoter signals on broad host range plasmids and that these may be active, at least for the *trf*A gene on RK2, in both *E.coli* and *Pseudomonas spp.* However it does not totally preclude the possibility that for other genes there may be another overlapping promoter sequence more specific for *Pseudomonas.*

4.3 FURTHER ANALYSIS OF THE *P.PUTIDA* COUPLED TRANSCRIPTION/TRANSLATION SYSTEM

4.3.1 Comparison of polypeptides synthesised in the *E.coli* and *P.putida* coupled transcription/translation systems by *E.coli* and broad host range plasmid DNA

If the hypothesis that broad host range plasmids express a different set of genes in different hosts to allow replication therein is correct, then one would expect (if *in vivo* differences are reflected in the *in vitro* situation) to see some differences in the polypeptide products synthesised by broad host range plasmids in the *P.putida* coupled transcription/translation system compared with those synthesised in the *E.coli* system. To this end a comparison of the polypeptide products directed by broad host range and *E.coli* plasmid DNA in the two *in vitro* gene expression systems was made. A marked difference in the pattern and relative amounts of polypeptides produced in the *E.coli* and *P.putida* protein-synthesising extracts for both the broad host range and *E.coli* plasmids was observed when the coupled transcription/translation systems of *E.coli* and *P.putida* were programmed with either the *E.coli* plasmid pGB123 or with the broad host range plasmids R300B and the R300B-based vector pGSS33 (Figure 21, Chapter 3).

4.3.1.1 *E.coli* plasmid DNA

pGB123 is a 2.7kb *Eco*RI fragment of DNA from the human pathogenic *E.coli* K1 strain cloned in pUC8 (I.S.Roberts and G.J.Boulnois, unpublished). Vann *et al* (1987) investigated the coding capacity of a 3.3kb region spanning this 2.7kb fragment. By using minicells they identified two polypeptides of molecular weight 50kD and 45kD. The 50kD polypeptide was identified as CMP-NeuAc synthetase and the 45kD polypeptide was involved in the synthesis of N-acetyl neuraminic acid. When either the *E.coli* or the *P.putida* coupled transcription/translation systems were programmed with

pGB123 two polypeptides of approximate molecular weights 33kD and 42kD were produced (Figure 21, Chapter 3). Vann *et al* (1987) report that the ATG translational start codon is only 20bp downstream of the *Eco*RI site, which in pGB123 is the limit of the fragment of cloned DNA. Thus the discrepancy between the polypeptide sizes observed by Vann *et al* and those synthesised in the *in vitro* gene expression systems by pGB123 can be explained by the fact that the promoter for the CMP-NeuAc synthetase gene is probably not present in pGB123, although there may still be low level expression due to readthrough transcription from other promoters on the plasmid. The polypeptide of 42kD is probably analogous to the 45kD polypeptide observed by Vann *et al*, considering inaccuracies in determining molecular weights by SDS-PAGE. The polypeptide of 33kD synthesised in the coupled transcription/translation systems may be a truncated form of the 42kD polypeptide, but at present there is no evidence to support this.

The products encoded by the *E.coli* plasmid pGB123 are almost identical in the coupled transcription/translation systems of *E.coli* and *P.putida* (Figure 21, Chapter 3, Tracks 2 and 4) indicating that the cloned gene products were synthesised in both extracts. Most of the vector-encoded polypeptides are common to both extracts, however there are a few differences. For example the polypeptide of 27kD is uniquely synthesised in the *E.coli* extract, whilst the polypeptide of 17kD is uniquely synthesised in the *P.putida* extract. This close similarity between polypeptides made in the extracts directed by the *E.coli* plasmid pGB123 is not confined to this pUC8-based clone. With the pBR322-derived vector pAT153 (Figure 29, Chapter 3) almost all polypeptides appear to be synthesised in both extracts. It is possible that some of the differences observed between the two systems may be due to minor proteolytic activity.

4.3.1.2 Broad Host Range plasmid DNA

With R300B as DNA template a polypeptide of approximately 73kD is

strongly expressed in the *P.putida* extract (Figure 21, Chapter 3, Track 8) but not at all in the *E.coli* extract (Figure 21, Chapter 3, Track 6). In the same way as R300B, pGSS33, an R300B-based broad host range cloning vector with multiple antibiotic resisitance, directs the production of a 73kD polypeptide in the *P.putida* extract (Track 12) and not in the *E.coli* extract (Track 10). This product was investigated further (see 4.4). The polypeptide product of 37kD directed by R300B appears (on Figure 21) to be specific to the *P.putida* extract. However the 37kD polypeptide has been seen to be produced in the *E.coli* extract when the film is exposed longer to the gel. Hence it appears to be synthesised to a lower level in the *E.coli* extract than the *P.putida* extract. This polypeptide is probably the product of *repB*, since it has been reported to be 38kD in size (Haring *et al* 1985) and a polypeptide of 38kD (i.e. RepB) is the second largest polypeptide encoded by RSF1010, the largest being RepB* (E.Scherzinger, Personal Communication). There is also a 23kD polypeptide which is more intense in the *E.coli* extract than in the *P.putida* extract.

Although there are polypeptides produced with equal intensity in both extracts, there are also polypeptides of different intensities. For example, the 11KD polypeptide band is more intense in the *P.putida* extract than in the *E.coli* extract. Whilst caution must be exercised in the interpretation of levels of polypeptides synthesised as judged by band intensity, it does suggest that there is differential expression of genes in the two extracts.

Although R300B and RSF1010 were independently isolated they are thought to be identical. To determine the similarity of polypeptides synthesised in the *in vitro* gene expression systems by R300B compared with those made by RSF1010, the *E.coli* and *P.putida* coupled transcription/translation systems were programmed with both plasmids and the polypeptides compared. The results show that polypeptide products directed by R300B are identical to those synthesised by RSF1010 in both the *E.coli* and *P.putida* extracts

(Figure 33).

4.3.2 Programming the coupled transcription/translation systems with restriction fragments for the location of polypeptide products to coding sequences

To determine whether restriction fragments could be used to programme a coupled transcription/translation system of *P.putida*, as has been reported for the equivalent *E.coli* system (Pratt *et al* 1981), an attempt was made to locate the polypeptide products synthesised in the two extracts on restriction fragments of R300B. R300B DNA was digested with *Eco*RV, *Pst*I or *Acc*I. The samples were analysed by agarose gel electrophoresis and each fragment isolated by electroelution. The fragments were then extracted several times with water-saturated butan-2-ol to remove any ethidium bromide (which inhibits protein synthesis). These fragments were then used to programme the *E.coli* and *P.putida* coupled transcription/translation systems. Table 13 shows the total amount of incorporation of ^{35}S-methionine into TCA-precipitable counts per reaction. The difference in the amount of radiolabel incorporated by the fragment compared with that incorporated by the supercoiled molecule varies from 74% to 60% for the *P.putida* extract and 59% to 30% for the *E.coli* extract. It may be possible to achieve higher incorporation levels by the use of greater concentrations of the restriction fragments, this was not however checked. Pratt *et al* (1981) noted a similar level of reduction (up to 60%) in overall incorporation of radiolabel using linear fragments compared with supercoiled template.

There are known to be three antibiotic resistance genes on R300B; one Su^r and two Sm^r genes. The expression of the Su^r and Sm^r genes in *E.coli* is governed by a single transcriptional unit from a promoter upstream of the Su^r gene (Rubens *et al* 1976). From RNA polymerase binding studies (Bagdasarian *et al* 1981) it was found that there is a tandem promoter at coordinate 7.8kb (Figure 32) which is responsible for the transcription of

FIGURE 33 Polypeptides directed by R300B and RSF1010 in *E.coli* and *P.putida* S30 extracts.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. RSF1010 was incubated with *E.coli* S30 extract (Track b) and with *P.putida* S30 extract (Track d). R300B was incubated with *E.coli* S30 extract (Track c) and with *P.putida* S30 extract (Track e). As controls the *E.coli* extract and *P.putida* extracts were incubated without DNA (Tracks f and g respectively). Track a contains molecular weight standards.

FIGURE 34 Polypeptides directed by RSF1010 and mobilisation deficient derivatives of RSF1010 in S30 extracts of *E.coli* and *P.putida*.

Labelled polypeptides were analysed by SDS-PAGE and fluorography. *E.coli* S30 extract was programmed with RSF1010 (Track f), Mob^+ Tn3 derivatives (Tracks a and c), Mob^- Tn3 derivatives, pKT228 (Track h), pKT229 (Track i) and Mob^- deletion derivatives RSF1010Δ18 (Track 11) and RSF1010Δ20 (Track n). *P.putida* S30 extract was programmed with RSF1010 (Track g), Mob^+ Tn3 derivatives (Tracks b and d), Mob^- Tn3 derivatives, pKT228 (Track j), pKT229 (Track k) and Mob^- deletion derivatives RSF1010Δ18 (Track m) and RSF1010Δ20 (Track o). *E.coli* and *P.putida* extracts were incubated without DNA (Tracks p and q respectively). The arrows with solid arrowheads indicate the *rep*B product (37kD). The arrows with open arrowheads indicate the *rep*B* product (73kD). Track e contains molecular weight standards.

FIGURE 33

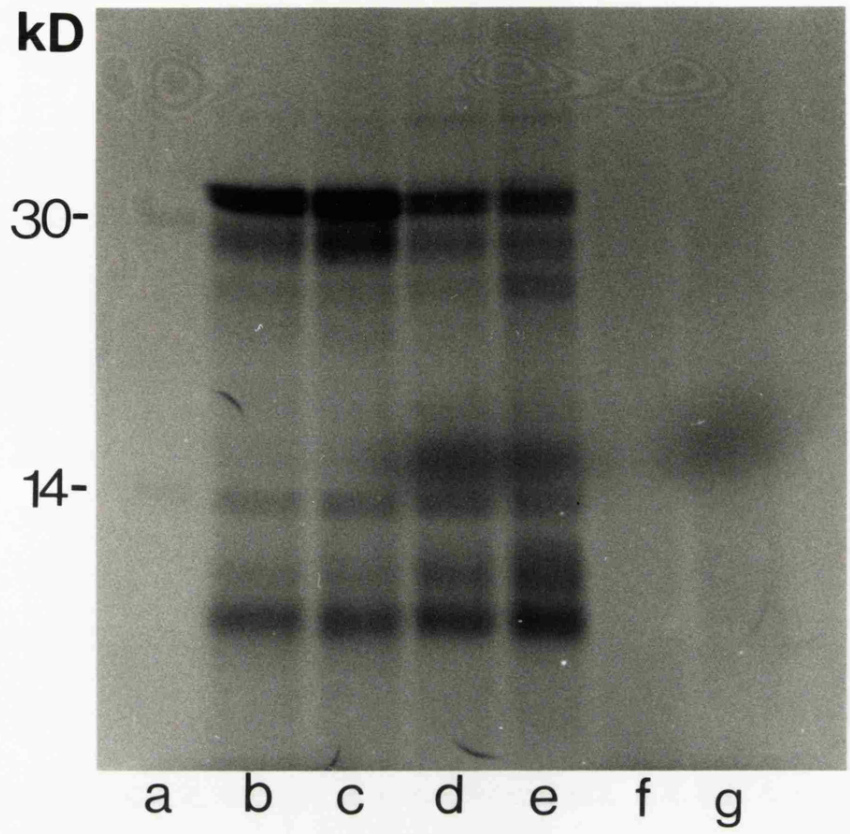


FIGURE 34

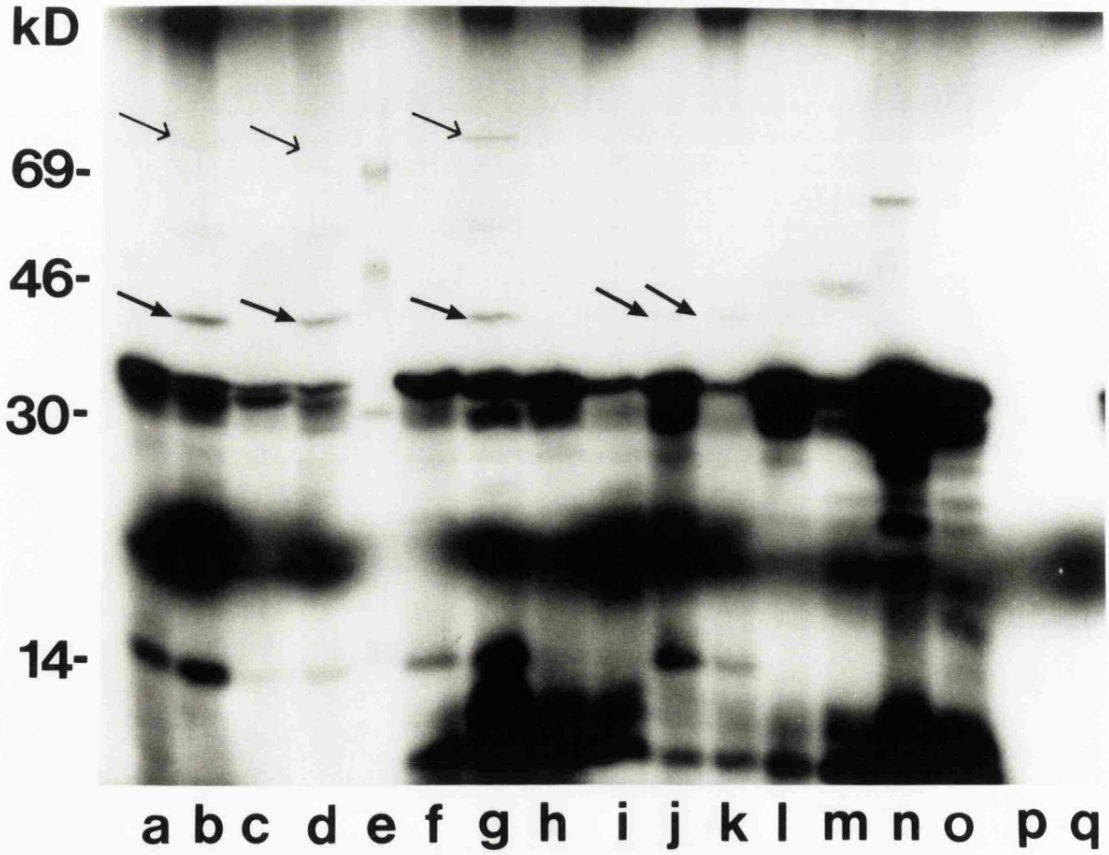


TABLE 13

INCORPORATION OF ^{35}S-METHIONINE INTO TCA-PRECIPITABLE COUNTS WITH COUPLED TRANSCRIPTION/TRANSLATION SYSTEMS PROGRAMMED WITH R300B AND RESTRICTION FRAGMENTS OF R300B

R300B/ RESTRICTION FRAGMENT OF R300B	S30 EXTRACT	TOTAL INCORPORATION INTO TCA-PRECIPITABLE COUNTS (CPM x 10^4) Mean of 3 replicates	LEVEL OVER DNA NEGATIVE BACKGROUND
R300B supercoiled	*E.coli*	10.33 ± 0.70	2.1
R300B supercoiled	*P.putida*	37.45 ± 1.93	3.9
*Eco*RV 3.2kb	*E.coli*	4.73 ± 0.18	1.0
*Eco*RV 3.2kb	*P.putida*	9.8 ± 0.53	1.0
*Eco*RV 5.5kb	*E.coli*	7.18 ± 0.35	1.5
*Eco*RV 5.5kb	*P.putida*	14.35 ± 0.70	1.5
*Pst*I 0.8kb	*E.coli*	4.20 ± 0.35	0.9
*Pst*I 0.8kb	*P.putida*	11.73 ± 1.23	1.2
*Pst*I 7.9kb	*E.coli*	6.30 ± 0.88	1.3
*Pst*I 7.9kb	*P.putida*	14.88 ± 0.53	1.5
*Acc*I 1.9kb	*E.coli*	6.48 ± 0.35	1.3
*Acc*I 1.9kb	*P.putida*	12.78 ± 0.18	1.3
*Acc*I 6.8kb	*E.coli*	4.90 ± 0.35	1.0
*Acc*I 6.8kb	*P.putida*	13.65 ± 0.18	1.4
DNA negative	*E.coli*	4.90 ± 0.18	
DNA negative	*P.putida*	9.63 ± 0.53	

the Sm^r and Su^r. The polypeptides of 30kD and 33kD were seen to be expressed by R300B in both the *E.coli* and *P.putida* coupled transcription/translation systems and by the 5.5kb *Eco*RV fragment and the 7.9kb *Pst*I fragment in both extracts. The 33kD polypeptide was directed by the 6.8kb *Acc*I fragment only in the *P.putida* extract. Incorporation of the radiolabel in the *E.coli* extract was generally lower than incorporation in the *P.putida* extract. Hence, the 33kD polypeptide was probably not detectable in the *E.coli* extract since the level of radiolabel incorporated by the 6.8kb *Acc*I fragment was too low. This indicates that the 30kD and 33kD polypeptides must be encoded by DNA lying between the *Acc*I site at coordinate 5.5kb and the *Pst*I site at coordinate 7.8kb or between the *Pst*I site at coordinate 8.6kb and the *Eco*RV site at coordinate 1.3kb (Figure 32). Therefore, based on their position on the RSF1010 molecule, either they are polypeptides from the replication region of the plasmid, or the Sm^r polypeptide. It is unlikely that either are a Sm^r gene product as, in *E.coli* at least, these genes would not be expressed since they are separated from the promoters which lie either side of the *Pst*I site at coordinate 7.8kb. The replication genes are known to be transcribed as a polycistronic message initiated from the promoter which lies just upstream of the *repB** ORF at 3.1kb (Bagdasarian *et al* 1986). The *repA* and *repC* genes are also transcribed as a separate message from a promoter upstream of *repA*. Experiments based on RNA polymerase binding studies in *E.coli* there is apparently no promoter which would allow transcription of *repC* alone. Within the region of the *Acc*I site at coordinate 5.5kb and the *Pst*I site at coordinate 7.8kb lie the two replication genes *repA* and *repC*. There is a promoter either side of the *Acc*I site at coordinate 5.5kb (Bagdasarian *et al* 1986) and although the majority of transcription derives from the promoter outside of the fragment examined here (Bagdasarian *et al* 1986), it is conceivable that this other promoter may be responsible for the transcription of *repA* and *repC* on this fragment. The RepA and RepC poly-

peptides were reported to be 29kD and 27kD respectively (Haring *et al* 1985). The fact that two polypeptides of approximately this size were directed by this fragment suggests that the polypeptides synthesised in the *E.coli* and *P.putida* coupled transcription/translation systems are likely to be the products of *rep*A and *rep*C.

A 29kD polypeptide is directed by R300B in both coupled transcription/translation systems and by the 6kb *Eco*RV fragment and by the 0.8kb *Pst*I fragment in both systems. This suggests it may be the product of the Su^r polypeptide since the small *Pst*I 0.8kb fragment is known to encode Su^r; one of the promoters will also be included on this fragment. It is known that when the *Pst*I site is cleaved the Su^r and Sm^r genes are still expressed (Barth *et al* 1981).

Although incorporation of label into TCA precipitable material for linear fragments was substantially lower than for supercoiled molecules, presumably due to exonucleolytic degradation of the linear DNA, it was possible to locate the positions of some genes encoding the replication polypeptides. The polypeptide of 73kD was apparently not synthesised by any of the restriction fragments. Even though some of the fragments carried the gene for the 73kD polypeptide, it was probably not detected since the incorporation level was so low that only the products from those genes that were expressed to a high level were visible by SDS-PAGE and fluorography.

4.4 IDENTIFICATION OF THE 73KD POLYPEPTIDE

The nature of the 73kD polypeptide encoded by RSF1010, but synthesised only in the *P.putida* coupled transcription/translation system, was investigated. Available literature on polypeptide products of RSF1010/R300B indicated that an open reading frame (ORF) existed encoding *rep*B* of the required size to encode a polypeptide of approximately 73kD. This open reading frame lies in the region of RSF1010 responsible for the replication

(*rep*) and mobilisation (*mob*) functions. This region has multiple overlapping open reading frames and one of these has been identified as *rep*B* by Bagdasarian (*et al* 1986) and as ORFA by Derbyshire *et al* (1987). Recent nucleotide sequence data of the entire plasmid has shown that this is the only ORF of a size large enough to encode a 73kD polypeptide (E.Scherzinger, Personal Communication). The expression of genes from the native plasmid has not been investigated by the use of *E.coli* gene expression systems such as minicells or maxicells or a coupled transcription/translation system. Only recombinant plasmids containing the cloned genes have been examined in this way. Haring *et al* (1985) cloned the replication genes onto an expression vector such that their expression was under control of the *tac* promoter and the products were detected by SDS-PAGE of whole cell lysates.

Derbyshire *et al* (1987) sequenced the entire mobilisation region of RSF1010 and identified several ORFs. By deriving mobilisation deficient mutants of RSF1010 and by locating these mutations to particular ORFs they have identified at least three genes which seem to be involved in mobilisation. These ORFs, which correspond to genes *mob*A, *mob*B and *mob*C, suggest that they would encode polypeptides of >65kD, 16kD and 9kD respectively. The products have been identified for the *mob*A and *mob*B genes by *E.coli* minicell analysis of the genes cloned onto pBR322-derived plasmids such that expression may be under control of vector-encoded promoter sequences. A peculiarity of the MobA polypeptide is that only the N-terminal portion appears to be required for effective mobilisation of RSF1010.

Buchanon-Wollaston *et al* (1987) cloned a region containing the replication genes and examined the products in an *E.coli* coupled transcription/translation system. They report the synthesis of a polypeptide of 72kD, but the fragment of DNA cloned was not sufficient for the whole *rep*B* ORF to be present. They present no data which supports this

observation.

The possibility that the 73kD polypeptide synthesised in the *P.putida* extract is RepB* was examined by programming the *P.putida* and *E.coli* coupled transcription/translation systems with 4 RSF1010 mutants having a Mob^- phenotype. All the mutations map within the open reading frame of *rep*B* but do not extend into the open reading frame of *rep*B (Figure 32). The mutants used were two deletion mutants RSF1010Δ18 and RSF1010Δ20 and two Tn3 insertion mutants pKT228 and pKT229 (Bagdasarian *et al* 1982). Control experiments were performed by programming the extracts with 2 other Tn3 insertion mutants which had Mob^+ phenotypes.

Figure 34 shows the results from these experiments. The 73kD polypeptide is not synthesised in the *E.coli* extract by any of the mobilisation-deficient mutants, nor by RSF1010 or the control (Mob^+) mutants. However, it is synthesised by RSF1010 in the *P.putida* extract and by the control (Mob^+) mutants, but not by the mobilisation-deficient mutants. This strongly suggests that the 73kD polypeptide is encoded by *rep*B* and ORFA and is involved in mobilisation of RSF1010. One might expect that the mutations in RSF1010, in addition to affecting expression of *rep*B*, may also affect transcription of *rep*B. However for the transposon insertion mutants pKT228 and pKT229 the *rep*B product was synthesised in both the *E.coli* and *P.putida* coupled transcription/translation systems (Figure 34). Possibly transcription across *rep*B was due to the presence of a strong promoter in the Tn3 region. The *rep*B product was not evident in either of the coupled transcription/translation systems for the deletion mutants, but two novel polypeptides were synthesised by these plasmids. In the *P.putida* coupled transcription/translation system a novel polypeptide of 43kD was synthesised by RSF1010Δ18 and in the *E.coli* coupled transcription/translation system a novel polypeptide of 60kD was synthesised by RSF1010Δ20. These are probably due to transcriptional/translational readthrough of the normal transcription/

translation terminators.

To my knowledge this is the first data on the expression of the *rep*B* gene in *E.coli* and *P.putida* under control of its own gene promoter on the native molecule. Other workers have examined the expression of *rep*B* only on recombinant plasmids under the control of vector promoters.

4.5 SUMMARY OF RESULTS

The *P.putida* coupled transcription/translation system is a useful gene expression system since it has the ability to synthesise cloned gene products from *E.coli*-based plasmid vectors. The products from *E.coli* vector plasmids are approximately the same in the two coupled transcription/translation systems, the observed differences may be due to proteolytic degradation or aberrations in transcription or translation, but this is not extensive in view of the discrete products as visualised by SDS-PAGE. In addition, a polypeptide of greater than twice the size of the largest polypeptide made in the *E.coli* extract is directed by R300B in the *P.putida* extract. It is possible to direct protein synthesis from restriction fragments of DNA in the *P.putida* coupled transcription/translation system (as has been described for the equivalent *E.coli* system) albeit with decreased efficiency compared with supercoiled DNA.

There are interesting differences in the polypeptides directed from broad host range plasmids in the two coupled transcription/translation systems. A large molecular weight polypeptide (73kD) was uniquely synthesised by R300B in the *P.putida* coupled transcription/translation system and this was found to be a protein involved in mobilisation of the plasmid. If these *in vitro* differences represent true *in vivo* differences then there must be some differences at the transcriptional or translational level in the expression of broad host range plasmid DNA in the two hosts.

RepB* is produced in *E.coli* when expression of the gene is directed by

the *tac* promoter, but it is not synthesised in an *E.coli* coupled transcription/translation system under control of its own promoter although RepB is synthesised in such a system. Both RepB and RepB* are synthesised in a *P.putida* coupled transcription/translation system. *repB* and *repB** are transcribed from the same promoter which has high homology to the *E.coli* concensus promoter sequence. These results suggest that either there is instability of the *repB** mRNA or the RepB* protein compared with the *repB* mRNA or RepB protein in *E.coli*. Alternatively there may be translational difficulties in the production of RepB* in *E.coli*. For example *repB* may have a good ribosome binding site in both species, but *repB** may have a poor ribosome binding site in *E.coli*. There is a precedent for such an effect. It has been reported (Goldfarb *et al* 1982) that *B.subtilis* cannot utilise the ribosome binding site of an *E.coli* Cm^r gene. The putative ribosome binding site for *repB** (Bagdasarian *et al* 1986) has some homology to the proposed ribosome binding sites of both *E.coli* and *P.aeruginosa* (Shine and Dalgarno 1975).

Conceivably host functions could substitute for some of the plasmid-encoded proteins across the spectrum of Gram-negative organisms able to support the replication of these IncQ plasmids and therefore it is possible that different proteins or regions of proteins involved in replication or mobilisation could be synthesised in different hosts. Evidence for the involvement of different protein regions in replication in different hosts comes from the finding that for the broad host range plasmid RK2 the replication gene *trfA* encodes two polypeptides via two translational start signals. One protein appears to be required for replication in *E.coli* and *P.putida*, the other may be required for replication in *P.aeruginosa*. A similar situation exists in the plasmid ColIb whose host range is restricted to the enterobacteria. The ColI plasmid directs the production of two forms of the *Sog* protein (of 180kD and 240kD) involved in plasmid transfer which are antigenically related and

the larger one has primase activity. They are apparently transcribed from the same promoter, but translated from two translational start signals (Boulnois *et al* 1982).

Since there was a high degree of similarity between protein products of R300B in an *E.coli* and a *P.putida* coupled transcription/translation system it seems reasonable to suggest that broad host range plasmids do not express an entirely different set of genes in different hosts. The results provide no evidence for the existence of different promoter signals active in the two hosts. The differences observed in the expression of R300B genes in the *E.coli* and *P.putida* coupled transcription/translation systems may be due to translational differences in the two species.

CHAPTER 5

THE CONSTRUCTION OF A BROAD HOST RANGE PROMOTER-PROBE VECTOR FOR THE CLONING OF *P.PUTIDA* PROMOTER SEQUENCES

5.1 INTRODUCTION

For the shotgun cloning of promoter regions from the *P.putida* chromosome a suitable plasmid cloning vector was required which would allow detection of promoter strength *in vivo*; such a vector is termed a promoter-probe vector.

5.2 DEFINITION OF A PROMOTER-PROBE VECTOR

A promoter-probe vector is a plasmid into which fragments of DNA can be inserted and their ability to promote transcription can be monitored. Thus, a good vector should have a multiple cloning site with several different and unique restriction sites for the cloning of a variety of restriction fragments of DNA. Downstream of this should be a promoterless gene with a readily assayable product and the plasmid should have an antibiotic resistance marker for the selection of bacteria harbouring the vector. Many such vectors are available for use in *E.coli*, such as those based on the assay of drug resistance due to the presence of a promoterless antibiotic resistance gene (West *et al* 1979), or those based on an enzyme-encoded gene the product of which can easily be assayed for, for example β-galactosidase, galactokinase or alkaline phosphatase (Chak and James 1985, Schneider and Beck 1986). As described in Chapter 1 only a few promoter-probe vectors exist for use in *Pseudomonas*, but this number is growing. During the course of this work several more have been described and one has been constructed (Konyecsni and Deretic 1988) which is essentially the same as that described here.

5.3 REQUIREMENTS OF THE VECTOR

The aim of the experiment was to detect clones with promoter activity in *P.putida*. The initial selection of such clones therefore had to be done in this host. Thus, it was necessary, not only that the promoter-probe

vehicle had a broad host range replicon, but also that it was easily mobilisable from *E.coli* to *P.putida*. Although *P.putida* can be transformed it is rather refractory to the process and it is obviously desirable to maximise the number of clones obtained in each experiment. The plasmid which was adopted for the construction of the promoter-probe vector, pCF32 (Spooner *et al* 1987), is 15.1kb and has RSF1010 replication (*rep*) and mobilisation (*mob*) functions (Figure 35). It has a promoterless *xylE* gene and expression of *xylE* after insertion of a promoter can easily be assayed by spraying colonies with a solution of catechol. *xylE*, one of the genes of the TOL pathway, specifies the enzyme catechol 2,3-oxygenase (C230) which cleaves catechol (a colourless substance) to form 2-hydroxymuconic semialdehyde (2-HMS, a yellow compound). Thus, when colonies are sprayed with catechol, promoter activity is indicated by a rapid production of this bright yellow compound. This is a sensitive assay involving the use of a cheap substrate. The plasmid also specifies resistance to kanamycin, streptomycin and carbenicillin and can be mobilised to *Pseudomonas* by IncIα, IncM, IncX and IncP plasmids with the IncP plasmids being the most efficient (Willetts and Crowther 1981). There is a sequence upstream of *xylE* which acts as an inducible promoter in *P.putida*, but does not have promoter activity in *E.coli*. This activity seems to involve a chromosomally-located element (Franklin *et al* 1983) and induction is by benzoate or *m*-toluate. *E.coli* cells harbouring pCF32 which are sprayed with catechol rapidly turn an intense yellow colour indicating high promoter activity. This is presumably due to read-through transcription from the strong kanamycin gene promoter upstream of the *xylE* gene. pCF32 can itself be used as a promoter-probe vector (Spooner *et al* 1987) by replacement of the *Hin*dIII/*Eco*RI fragment (containing the Km^r promoter) with DNA fragments. However, for the purposes of this work it was more desirable to have greater flexibility for cloning DNA fragments; a multiple cloning site was therefore inserted into the vector.

FIGURE 35

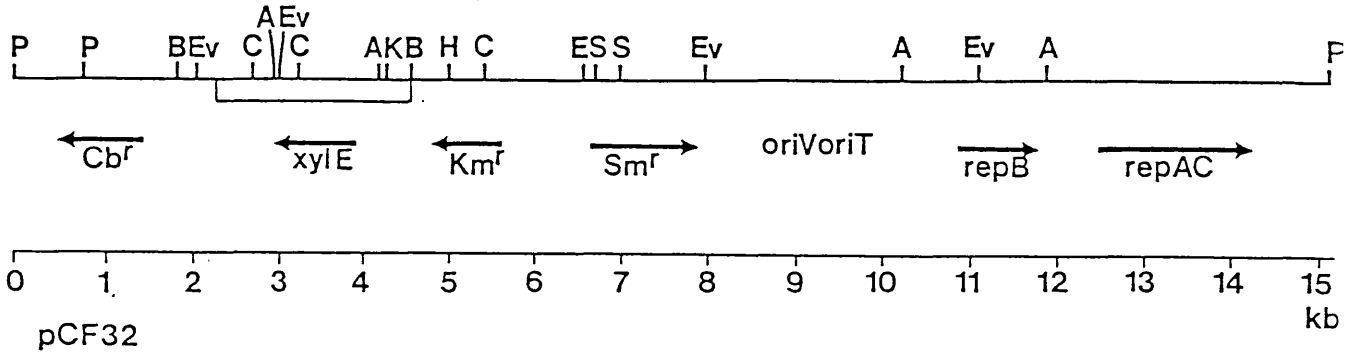


FIGURE 35 Restriction map of pCF32 linearised at a *Pst*I site.

Replication genes are denoted *rep*B and *rep*AC and their direction of transcription is indicated. The cloned C230 gene is indicated by *xyl*E and its direction of transcription is shown. The open box illustrates the fragment of DNA cloned from the TOL plasmid pWWO. pCF32 was constructed by cloning this region downstream of the Km^r determinant in pKT240. Resistance genes for kanamycin (Km), streptomycin (Sm) and carbenicillin (Cb) are indicated and their direction of transcription is shown. Origins of vegetative replication and transfer are denoted by oriV and oriT respectively.

5.4 CONSTRUCTION OF THE PROMOTER-PROBE VECTOR

The steps for generation of a promoter-probe vector from pCF32 were to delete the kanamycin resistance promoter and introduce a multiple cloning site upstream of the promoterless *xylE* gene. Since the vector possesses carbenicillin resistance in addition to resistance to kanamycin and streptomycin, it was decided to delete the DNA between the *Hin*dIII site at 13.5kb and the *Sph*I site at 0.5kb relative to the *Eco*RI site, so that in addition to removing the kanamycin resistance promoter the size of the vector was reduced quite considerably. This would also inactivate the streptomycin resistance gene. It is not possible to reduce the size of the vector much more since the mobilisation and replication functions cover about 5.5kb of the molecule.

A small multiple cloning site was designed (Figure 36) with one *Hin*dIII end and one *Sph*I end, and with restriction sites for *Bgl*II, *Xho*I and *Eco*RI for ligation into the plasmid. Ligation of this oligonucleotide into the large *Hin*dIII/*Sph*I fragment of pCF32 would create a molecule with unique restriction sites for *Hin*dIII, *Bgl*II, *Xho*I, *Eco*RI and *Sph*I. Two complementary oligodeoxynucleotides of 25 and 17 base pairs were synthesised using an automated DNA synthesiser (Figure 36). Equimolar amounts of the oligonucleotides were annealed together at 10°C overnight. Each single oligonucleotide and the annealed oligonucleotides were examined for purity by kinase-labelling 5μg of each with [γ^{-32}P]ATP (Forward reaction) and analysing the products on a 20% denaturing acrylamide gel. The gel was exposed to X-ray film for one hour and the result is shown in Figure 37. Relative to the positions of the dye fronts on the gel the major species in the preparation of the 25-mer were a fragment of 25 base pairs with some contaminating high molecular weight species. The preparation of the 17-mer contained a 17 base pair species also with some contaminating higher molecular weight species. The annealed oligonucleotide mixture shows the presence of all species. The two oligonucleotides were of sufficiently

FIGURE 36

THE MULTIPLE CLONING SITE

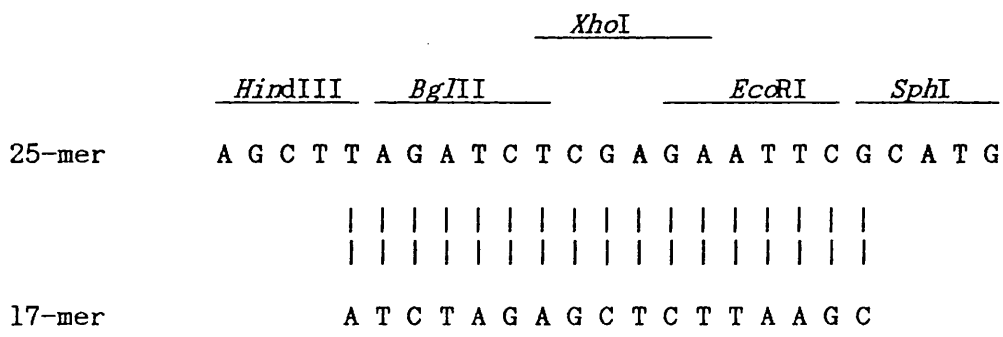


FIGURE 37 End-labelled preparations of the 17bp and 25bp oligonucleotides.

The 25bp and 17bp oligonucleotides and the annealed preparation were end-labelled with [γ-^{32}P]ATP and analysed on a 20% denaturing polyacrylamide gel (Tracks A, B and C respectively)

FIGURE 37

high purity for use directly in cloning into the vector.

pCF32 was digested sequentially with *Hin*dIII and *Sph*I (since different digestion buffers are required) and the products analysed by agarose gel electrophoresis. The DNA fragments obtained were of 350bp, 1.7kb and 13.1kb. The band of 13.1kb was isolated. 15pmoles (of 5' ends) of the annealed oligonuceotide were phosphorylated using T4 kinase (Forward reaction) and the 13.1kb fragment of pCF32 (1μg) was ligated to the oligonucleotide using a ratio of 5' ends of oligonucleotide:vector of 1:1 since a great excess of oligonucleotides (and hence ends) increases the probability of cloning in more than one oligonucleotide per vector molecule. *E.coli* strain LE392 was transformed with half of the products of this reaction and plated on Luria agar plates containing ampicillin. 62 transformants were obtained and each of these was inoculated onto three L-agar plates containing ampicillin, kanamycin or streptomycin and reincubated to determine whether they had lost the *Hin*dIII/*Sph*I fragment and hence the Km^r and Sm^r determinants. Of the 62 transformants, 60 were Ap^r, Km^s, Sm^s (the required phenotype). Of these 4 were white on spraying with catechol and 56 were pale yellow. 2 transformants were Ap^r, Km^r, Sm^s and on spraying with catechol were as bright yellow as LE392(pCF32). These were probably derived due to a partial digest of pCF32 such that some molecules were not cut with *Hin*dIII, but were cut with *Sph*I at both sites with the concomitant loss of the 350bp *Sph*I/*Sph*I fragment (and thus the Sm^r). This would be followed by religation of the molecule either incorporating no polylinker molecules or multiples of two ligated at their *Hin*dIII sites. Since these clones were of no interest they were not investigated further.

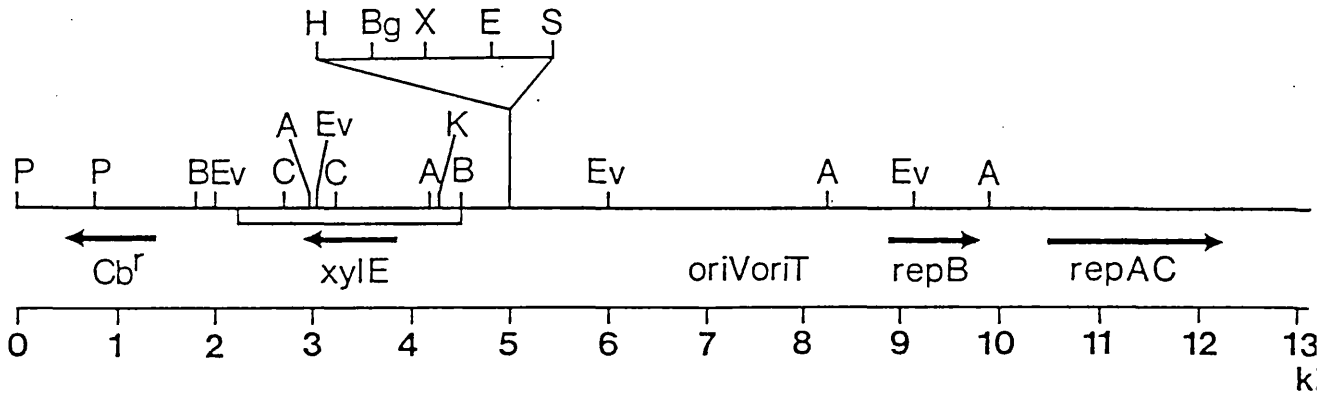
To check that the Ap^r, Km^s, Sm^s transformants contained one (or more) multiple cloning sites, plasmid DNA was prepared (by the miniprep method) from 24 of these colonies and digested with *Bgl*II which cleaves a unique site in the polylinker. This enzyme linearised all the plasmids indicating

that each had gained at least one *Bgl*II site. Some of the plasmids, when linearised with *Bgl*II, were larger than others, possibly indicating that more than one multiple cloning site had been cloned. To check this possibility, all the 24 clones were digested with *Hin*dIII and *Sph*I; this would cut out all the DNA cloned in, leaving a fragment which should be of identical size (13.1kb) for all clones. For 21 of the 24 clones this was the case; the other 3 were not digested. It is possible that in these latter three the junction sites may have been perturbed such that they could not be recleaved. The four plasmids which caused *E.coli* cells harbouring them to remain white on spraying with catechol tended to be those which could not be digested with some restriction enzymes. These were probably spontaneous deletion mutants (see 5.6).

To identify a plasmid which possessed a single multiple cloning site, 2 plasmids, pRLA1 and pRLA5, were chosen which appeared to have the smallest size on digestion with *Bgl*II and which were recleaved by *Hin*dIII+*Sph*I. Large scale preparations of the plasmid DNA were made from these. Plasmids pRLA1 and pRLA5 gave rise to a pale yellow colour on spraying cells harbouring them with catechol. Both plasmids, pRLA1 and pRLA5, were digested with *Hin*dIII, *Eco*RI, *Bgl*II, *Xho*I or *Sph*I and each of these enzymes linearised the plasmids. Hence, these plasmids had acquired unique restriction sites not present in the parent plasmid pCF32. Both clones were mapped by restriction analysis and appeared to be identical (Figure 38).

To confirm the presence of the polylinker in pRLA1 and pRLA5, the single stranded 25 base pair oligonucleotide was end-labelled with [γ-^{32}P]ATP (Forward reaction) and used to probe a DNA dot blot of pRLA1, pRLA5, pCF32 and the double-stranded oligonucleotide. The result of the dot blot (Figure 39) shows there is hybridisation of the labelled oligonucleotide to pRLA1 and pRLA5, but not to pCF32. A Southern blot of pCF32 and pRLA5 was also carried out using the end-labelled 25 base pair oligonucleotide as a probe. For this Southern blot, pRLA5 was digested with

FIGURE 38



pRLA5

FIGURE 38 Restriction map of pRLA5 linearised at a *Pst*I site.

pRLA5 was constructed from pCF32 by ligating the 13.1kb *Hin*dIII/*Sal*I fragment to a *Hin*dIII/*Sal*I bounded polylinker which has unique sites for *Hin*dIII (H), *Bgl*II (Bg), *Xho*I (X), *Eco*RI (E) and *Sph*I (S). Other restriction sites are denoted by P, *Pst*I; B, *Bam*HI; C, *Cla*I; A, *Acc*I; Ev, *Eco*RV; K, *Kpn*I. The origins of vegetative replication and transfer are denoted by oriV and oriT respectively. The position and direction of transcription of the carbenicillin resistance gene (Cb^r) and the promoterless C230 gene (*xylE*) are shown.

FIGURE 39 Dot blot using the end-labelled oligonucleotide.

pCF32 (Dots B), pRLA1 (Dot C), pRLA5 (Dot D) and the double-stranded oligonucleotide (Dot A) were probed with the end-labelled single-stranded 25bp oligonucleotide.

FIGURE 40 Southern blot of restricted pRLA5 and pCF32 probed with the end-labelled oligonucleotide.

The agarose gel which was probed with the end-labelled single-stranded 25bp oligonucleotide is shown on the left hand side of the Figure and the resultant hybridisation on the right hand side. pRLa5 was digested with *Eco*RI (Track b) or *Bam*HI (Track c). pCF32 was digested with *EcoRI* (Track d), *Bam*HI (Track e) or *Acc*I (Track f). Tracks a and g contain 1kb ladder as molecular weight markers.

FIGURE 39

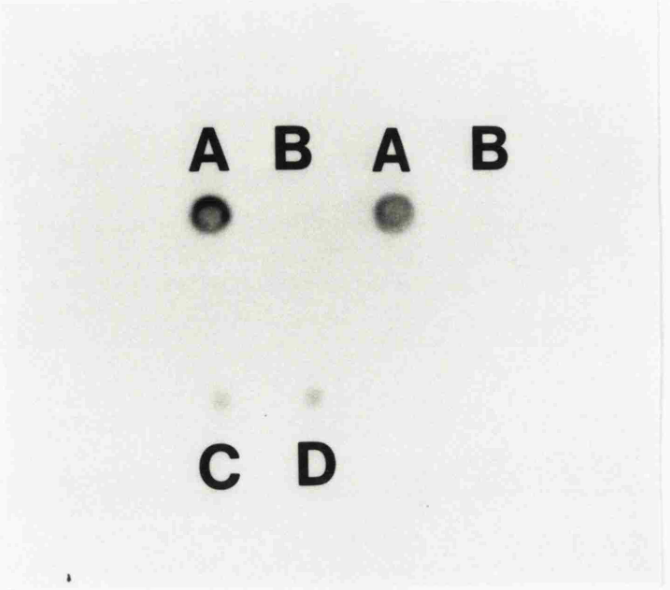
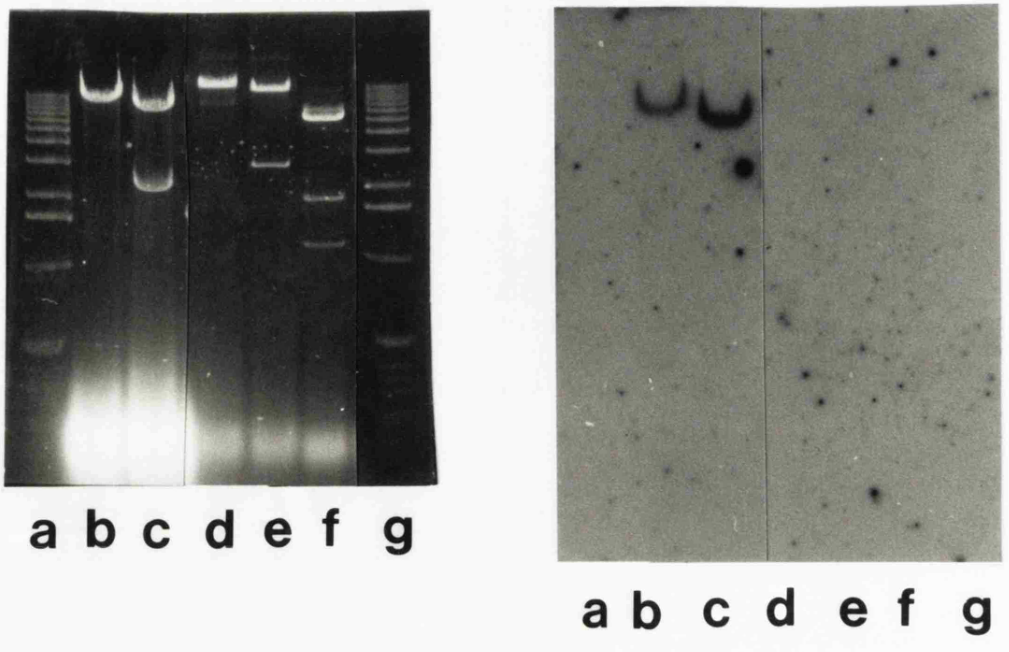


FIGURE 40



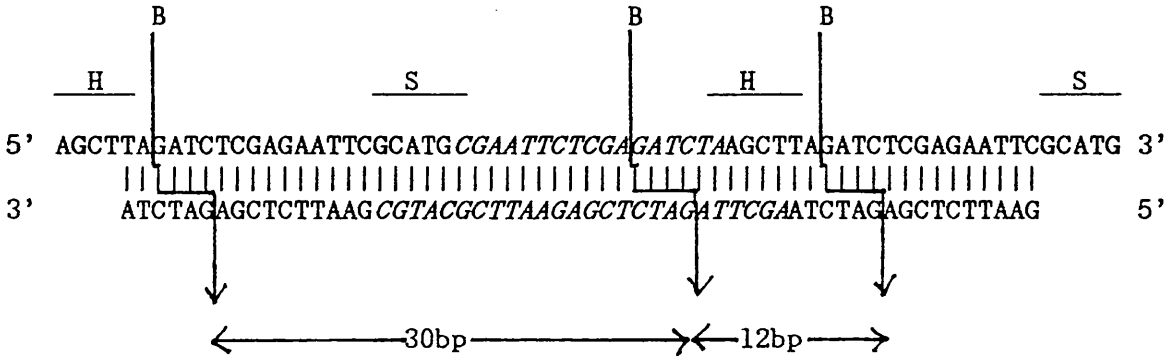
*Eco*RI or *Bam*HI and pCF32 was digested with *Eco*RI, *Bam*HI or *Acc*I. The result of hybridisation of this Southern blot to the [γ-^{32}P]ATP-labelled oligonucleotide is shown in Figure 40; there was no hybridisation to any of the pCF32 fragments, but hybridisation was observed to pRLA5 linearised with *Eco*RI and the *Bam*HI fragment of 10.4kb (the expected fragment surrounding the position of insertion of the polylinker). These results show that pRLA5 contains a region of DNA specifically complementary to the oligonucleotide, proof that pRLA5 contains at least one polylinker region.

In order to determine how many oligonucleotides had been cloned in pRLA1 and pRLA5, the plasmid DNA was isolated from contaminating tRNA by excising the undigested DNA band from an agarose gel, digested with *Bgl*II and labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Exchange reaction). For plasmids possessing just one oligonucleotide there would be just one *Bgl*II site and only the linear molecule of 13.1kb would be end-labelled. Plasmids with multiple copies of the polylinkers of 3, 5, 7 etc. would yield 3 species of labelled DNA fragments of sizes 13.1kb, 12bp, and 30bp (Figure 41). The labelled DNA fragments were analysed electrophoretically on a 20% non-denaturing acrylamide gel. Figure 42 shows that for pRLA1 three bands were present and on their location relative to the positions of the dye fronts these probably represent species of 12bp, 30bp and 42bp indicating that this plasmid possesses at least two oligonucleotides (the band of 42bp is assumed to be a partial digest). No such fragments were evident for pRLA5 indicating that it probably contained just one oligonucleotide molecule.

pRLA5 is Ap^r, Km^s, Sm^s, is linearised by all the enzymes which have unique sites in the polylinker and contains one polylinker. On this basis it was chosen as the promoter-probe vector for the shot-gun cloning of promoters from *P.putida*.

FIGURE 41

DIGESTION OF MULTIPLE POLYLINKERS WITH *Bgl*II



Restriction sites denoted by;

B *Bgl*II

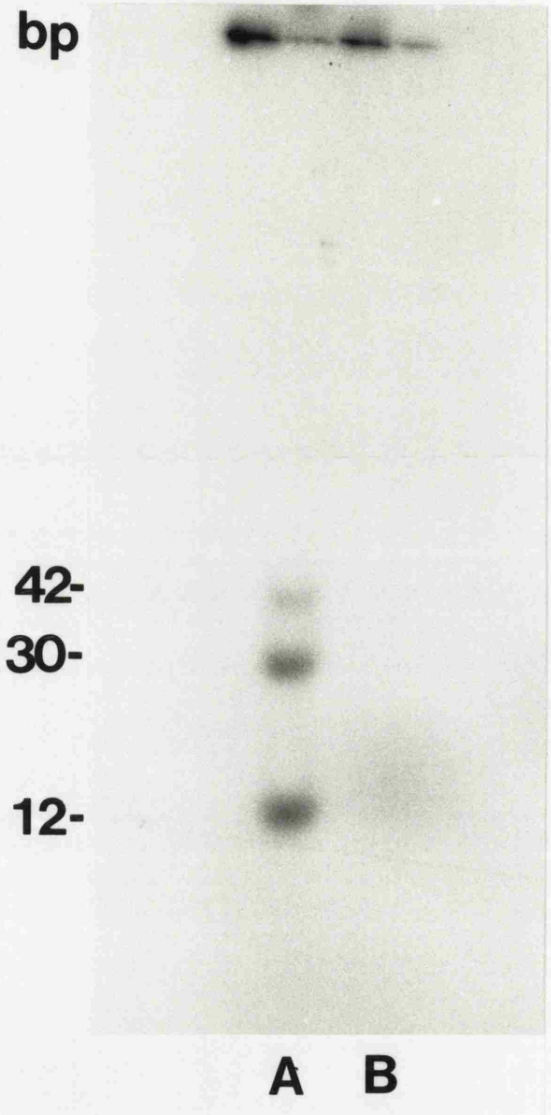
H *Hin*dIII

S *Sph*I

FIGURE 42 pRLA1 and pRLA5 digested with *Bgl*II and end-labelled.

Plasmids pRLA1 and pRLA5 were digested with *Bgl*II and the resultant fragments end-labelled with [γ-^{32}P]ATP and analysed on a 20% non-denaturing gel (Tacks A and B, repectively). Approximate molecular weights are indicated.

FIGURE 42



5.5 CATECHOL 2,3,-OXYGENASE ACTIVITY OF pRLA5

When colonies of *E.coli* harboring pRLA5 are sprayed with catechol they turn a paler yellow colour than the same bacteria carrying pCF32, indicating a much lower enzyme activity. This probably reflects a reduced activity across the *xylE* gene than that exhibited by pCF32. Quantitative measurements of C230 levels were made for JM83 and 2440 alone or harboring pRLA5 and pCF32. The results are shown in Table 14. There was no C230 activity detectable for JM83 or 2440.

There is an approximate 12 fold increase in C230 activity of 2440(pRLA5) compared with the activity of C230 of JM83(pRLA5). There is an error of 30% on both the value of C230 activity of pRLA5 in JM83 and the value of C230 activity in 2440. Although this is high, measurements made on the same day generally showed a 12 fold difference in enzyme activity in *E.coli* compared with *P.putida*. A possible reasons for the difference in C230 activity between the two hosts is that readthrough transcription from a promoter upstream of the C230 gene occurring to a higher degree in 2440 than in JM83. Alternatively, the promoter adjacent to *xylE* is responsible for the difference and is active in *P.putida* but not active in *E.coli*, or there is a pronounced difference in copy number of the plasmid in the two hosts.

Due to the practical problems in the determination of plasmid copy number particularly where comparison between different genera is involved, this possibility was not examined experimentally. In *P.putida* pCF32 produces approximately the same C230 activity as does pRLA5 in 2440, whereas there is a 14 fold difference in C230 activity from pCF32 compared with that from pRLA5 in JM83 (this was however only measured once). If it is reasonable to assume that the high level of C230 activity of pCF32 in *E.coli* is due to the activity of the Km^r gene promoter, this implies that the Km^r gene promoter is less strong in *P.putida* compared with its activity in *E.coli*, such that there is not so much readthrough transcription in

TABLE 14

C23O ACTIVITY OF pRLA5 AND pCF32 IN *P.PUTIDA* AND *E.COLI*

STRAIN	C23O ACTIVITY AT 37°C (OD_{375nm} units min^{-1} mg^{-1} protein)	NO. OF ESTIMATES
P.putida 2440	0.00	1
E.coli JM83	0.00	1
P.putida 2440(pRLA5)	7.8 ± 2.45	19
E.coli JM83(pRLA5)	0.63 ± 0.21	19
P.putida 2440(pCF32)	6.5	1
E.coli JM83(pCF32)	11.0	1

P.putida. There is no evidence in the literature for such a difference in activity of the Km^r gene promoter between these two species.

5.6 DETECTION OF PROMOTER ACTIVITY

The effect of varying the concentration of catechol used for spraying colonies was investigated. With *E.coli* as host low concentrations of catechol (10mM) resulted in the colour developing more slowly and it reached a lower final intensity than at higher concentrations (500mM) where there was a rapid change to a final more intense colour. An intermediate concentration of 100mM for spraying *E.coli* colonies was used throughout this study, since this gave a reasonable endpoint of colour intensity produced by those cells harbouring the vector and those cells harbouring plasmids with a putative promoter were easily discernible.

The detection of promoter activity in *E.coli* using pRLA5 as a promoter-probe vector is easy when colonies are sprayed with a 100mM solution of catechol, since the background colour (vector alone) is a pale yellow. The situation with the same vector in *P.putida* is more difficult; the change of colour of *P.putida* colonies following spraying with catechol is slower than that for *E.coli* colonies. This may suggest that catechol permeates *P.putida* more slowly than it does *E.coli*. In addition pRLA5 gives rise to 12 fold higher C230 activity in *P.putida* than in *E.coli*, therefore the background colour is considerably higher. A catechol concentration of 200mM was chosen since this caused a more rapid colour change than when 100mM catechol was used. The isolation of cells harbouring plasmids with a cloned putative promoter was carried out by picking the colonies which became yellow almost immediately on spraying with 200mM catechol. Colonies exhibiting lower C230 activity tended to develop the colour more slowly.

Occasionally, at an approximate frequency of 10^{-2}, colonies of both *E.coli* and *P.putida* containing pRLA5 arise which are totally white on spraying with catechol. Bacteria in these colonies were found to harbour a

plasmid which carried a deletion of about 9kb resulting in the loss of the *xylE* gene. These spontaneous deletion mutations indicate some instability of pRLA5, but it was not of a high enough frequency to compromise its use as a promoter-probe vector. Recombinant molecules, with DNA cloned in the multiple cloning site, did not undergo deletion. Interestingly, pCF32 also exhibits the same instability as pRLA5 (R.A. Spooner, Personal Communication).

5.7 SUMMARY OF RESULTS

A promoter-probe vector, pRLA5, was constructed from pCF32 by the replacement of a 2kb fragment containing the Km^r and Sm^r gene promoters with a multiple cloning site. This vector was shown to contain a single multiple cloning site. Bacteria harbouring pRLA5 possess some background C230 activity, which is 12 fold higher in *P.putida* than in *E.coli* making detection of promoter activity easier in *E.coli* than it is in *P.putida*. However, by the use of an appropriate concentration of catechol it is possible to distinguish between background and enhanced activity in *P.putida*.

CHAPTER 6

CLONING AND NUCLEOTIDE SEQUENCE ANALYSIS OF *PSEUDOMONAS*-SPECIFIC PROMOTERS

6.1 SHOTGUN CLONING OF *P.PUTIDA* DNA FRAGMENTS USING THE PROMOTER-PROBE VECTOR pRLA5

The aim of the experiment was to clone small (100-500bp) restriction fragments of *P.putida* chromosomal DNA in pRLA5 upstream of the promoterless *xylE* gene in order to determine their relative promoter strengths; first in *P.putida* and then to compare the activities in *P.putida* and *E.coli*; isolate those which exhibited promoter activity specific to *P.putida* and determine the nucleotide sequence of the cloned insert. If enough such clones were isolated, then a possible *Pseudomonas* concensus promoter sequence could be constructed.

The multiple cloning site of the promoter-probe vector pRLA5 was constructed in such a way that it would be possible to clone various restriction fragments, however problems were encountered with restricting 2440 chromosomal DNA with several enzymes. These were partially alleviated by purifying the DNA on a caesium chloride gradient, but since fragments of less than 500bp were required (so that they could be readily sequenced) total digestions of chromosomal DNA with *Sau*3A were carried out since no problems were encountered in the digestion using this enzyme, and it yields more fragments of the required size compared with enzymes whose recognition sequence is 6bp. It was not an attempt to produce a representative library of promoters, but rather to isolate any which appeared specific to *P.putida*. There was, therefore, no theoretical problem in using *Sau*3A alone as there would be, for example, in the construction of a representative gene library of a species.

P.putida 2440 DNA was digested with *Sau*3A and fragments of 500bp or less were size-selected by excising the relevant region from an agarose gel. These fragments were ligated into the pRLA5 vector which had been restricted with *Bgl*II and dephosphorylated to prevent self-ligation. An approximate ratio of dephosphorylated vector to insert of 1:1 was used for

maximum ligation efficiency.

The *E.coli* strain JM83(pLG221) was transformed with this reaction mixture and with the ligated vector alone. Since promoter activity was to be selected in *P.putida*, JM83(pLG221) was employed to provide a helper plasmid for the subsequent mobilisation of recombinants into 2440 since transformation efficiency of 2440 was too low for recombinant selection. pLG221 was a desirable helper plasmid because it possesses *tra* functions, has resistance to kanamycin for the selection of the plasmid in *E.coli* and it is not maintained in *P.putida* on transfer (Boulnois *et al* 1985). Thus any *Pseudomonas* colonies obtained after mobilisation by selection for resistance to carbenicillin would be 2440 harbouring the recombinant plasmid (providing the vector dephosphorylation had worked efficiently).

For an approximate estimate of numbers of recombinants and the efficiency of dephosphorylation, a tenth of each of the transformation reactions was plated out on Luria agar plates containing Ap and Km. The rest of the transformants were mobilised *en masse* into 2440 by filter matings, then plated out on Isosensitest agar containing carbenicillin and trimethoprim. Ampicillin resistance cannot be selected in *P.putida* since it is resistant to this antibiotic by virtue of the inability of the compound to enter the cell. Instead β-lactamase activity was assessed by using carbenicillin at a high concentration of $2mgml^{-1}$. The trimethoprim was incorporated into the selection medium to counter-select *E.coli* cells. *P.putida* is resistant to Tp since it is not taken up by these bacteria. JM83, however will grow on Luria agar containing Tp, since L-agar contains antagonists of Tp, thus Isosensitest agar, a medium specifically low in antagonists of Tp, was employed. JM83 will not grow on Isosensitest agar containing Tp at a concentration of $100\mu gml^{-1}$.

The plates were incubated at 30°C for 36 hours as this length of time produced colonies of a large enough size to detect the colour change on spraying. The colonies were sprayed with 200mM catechol and those colonies

with apparently higher activity (assessed on the intensity of the yellow colour) were picked and patched onto Isosensitest agar containing Tp and Cb, reincubated, restreaked onto the same medium containing Tp and Cb, reincubated, resprayed and one colony streaked onto L-agar containing Cb to purify. These were then grown up in 10ml cultures of L-broth containing Cb and assayed for C230 activity. The approximate total numbers of recombinants mobilised and assessed for activity by spraying the colonies was 35 000.

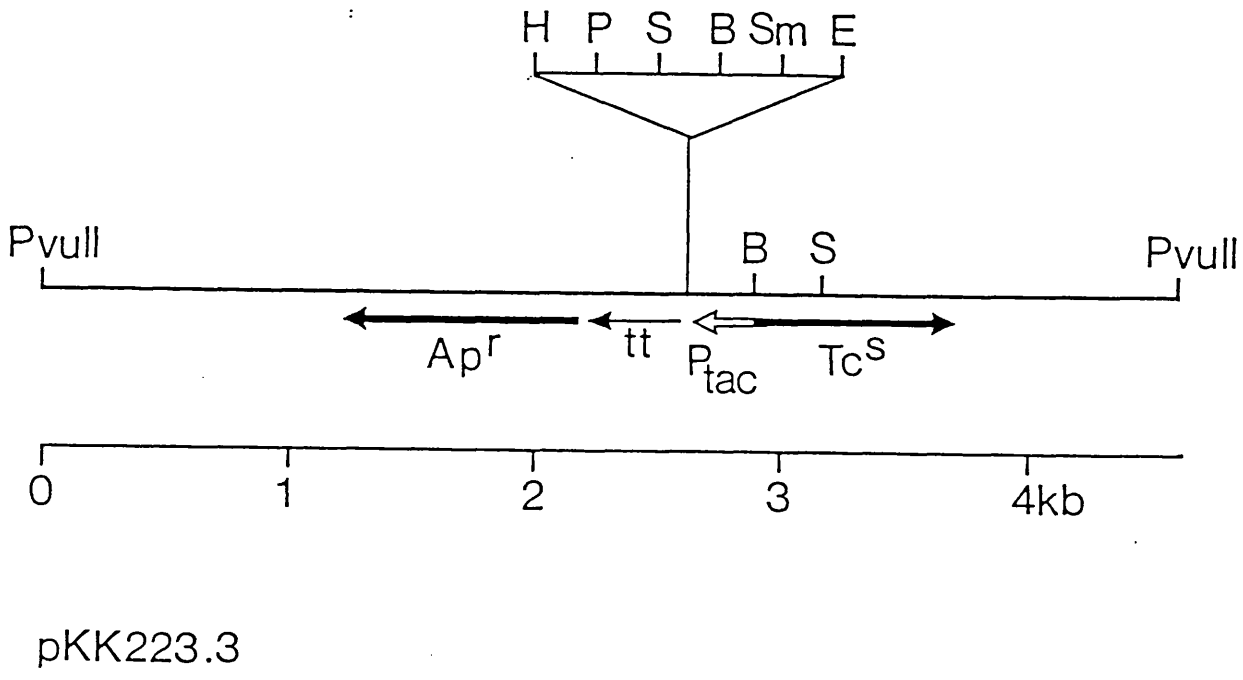
6.2 EXPRESSION OF THE C230 ASSAY RESULTS

The question arises as to how the data should be expressed in view of the 12 fold difference in C230 activity between the two hosts harbouring pRLA5. For example, should it be as actual values of activity of C230 directly comparable between the two hosts, or, as an increase in activity over the vector background level in the respective host? To determine which of these may provide a more realistic representation of the true *in vivo* activities of the clones an *E.coli* promoter, known to function at high levels in both hosts, was cloned into the vector and the C230 activities determined.

6.2.1 Subcloning the *tac* promoter in pRLA5

The *tac*I promoter, known to function at high levels in both *E.coli* and *P.putida* was cloned from pKK223.3 (Brosius and Holy 1984) (Figure 43) into pRLA5. The *tac*I promoter (De Boer *et al* 1983) is a hybrid promoter which combines the -35 region of the *trp* promoter with the -10 region of the *lac* promoter with a spacing of 16bp between. Hence, it exhibits the *E.coli* concensus sequences in the -35 and -10 regions. The vector pKK223.3 is an expression vector based on pBR322 and specifies resistance to Ap; it contains the multiple cloning site from pUC8 adjacent to the *tac* promoter for the cloning of genes to achieve high level expression. The *tac* promoter

FIGURE 43



pKK223.3

FIGURE 43 Restriction map of pKK223.3 linearised at the unique *Pvu*II site.

pKK223.3 (4585bp), an expression vector derived from pBR322 contains the *tac* promoter, denoted by P_{tac}, the polylinker from pUC8 and the *rrn*B ribosomal RNA transcription terminators, denoted by tt. Restriction sites are denoted by B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I; Sm, *Sma*I.

is normally used in a lacI^q background in order to achieve controlled high level expression by induction with IPTG. In this experiment all that was required was an examination of the activity of a typical *E.coli* promoter sequence in the two hosts and so the assays were carried out in *lac*$^-$ backgrounds so there was no repression and IPTG should have no inducible effect. In pKK223.3 there is a *Hin*dIII and a *Sal*I site in the polylinker and another *Sal*I site the other side of the tac promoter (Figure 43). The *tac* promoter could be easily isolated on a *Sal*I/*Sal*I fragment and inserted into pRLA5 linearised with *Xho*I. However the insert could potentially ligate to the vector in either orientation and so in order to achieve correct orientation for promoter activity across *xylE* the promoter was isolated on a *Hin*dIII/*Sal*I fragment and cloned into pRLA5 cut with *Hin*dIII and *Xho*I. This would be expected to be a more efficient process than dephosphorylation of the linearised vector and insertion of a *Sal*I/*Sal*I fragment, coupled with the possibility of 50% of recombinants having the insert in the opposite orientation to that required.

To obtain the *tac* promoter on the *Hin*dIII/*Sal*I fragment, the *Sal*I site in the polylinker must not be cut. If the plasmid was first cut with *Hin*dIII, then partially with *Sal*I, the *Sal*I site close to the *Hin*dIII site would be inefficiently cleaved; this increases the likelihood of the 0.6kb fragment having *Hin*dIII/*Sal*I ends rather than *Sal*I/*Sal*I ends. pKK223.3 was linearised with *Hin*dIII, by digesting the DNA with *Hin*dIII. This was ethanol precipitated then partially digested with *Sal*I. The reactions were analysed by agarose gel electrophoresis and the 0.6kb fragments isolated.

The vector, pRLA5, was digested with *Hin*dIII and *Xho*I and the resultant 13.1kb fragment was isolated from an agarose gel so that the intervening small fragment would be lost and would therefore not interfere with the subsequent ligation. The pRLA5 HindIII/XhoI fragment was ligated to each insert preparation to give approximately equal ratios of vector:insert as judged by agarose gel electrophoresis. pRLA5 *Hin*dIII/*Xho*I fragment was

ligated alone as a control.

Both MC1061 and JM83(pLG221) were transformed with half of each ligation mix, so that the promoter activity of the resultant clones could be assayed in MC1061 (a standard *E.coli* laboratory strain) and the clones mobilised from JM83 (using pLG221 as a helper plasmid) into 2440 and the promoter activity measured therein. A portion of each transformation mix was plated onto L-agar plates containing Ap (for MC1061) and L-agar plates containing Ap and Km (for JM83(pLG221)) and the plates incubated overnight at 37°C. The resultant transformant colonies were sprayed with 100mM catechol, the numbers of pale yellow and bright yellow noted and the bright yellow colonies, which were presumed to be due to the presence of the vector molecule containing the *tac* promoter in the correct orientation to give expression across the *xylE* gene, were inoculated onto fresh Luria agar plates containing the appropriate antibiotic.

The bright yellow colour of MC1061 recombinants exhibited instability since when streaked out they gave rise to small white colonies resulting from the large bright yellow colonies some of which had obvious white sectors. This suggests that the construct in MC1061 was unstable such that the cloned *tac* promoter is rapidly lost from the plasmid. The cloned *tac* promoter appeared totally stable in JM83. Most of the measurements of C230 levels of the promoter clones in the *E.coli* strain were carried out in MC1061 (Appendix) and occasionally this did result in loss of the insert. It would, therefore, have been better to use JM83 for all determinations, since by using MC1061 there may have been selective pressure for lower activity, by deletions or mutations in the insert DNA.

The C230 activities of the *tac* promoter cloned in pRLA5 (pRLA5.*tac*) were measured in JM83 and 2440 on several occasions and termed experiments I, II and III. There is some disparity in the results (Table 15), but in all cases the level of enzyme activity over vector in *E.coli* is greater than that in *P.putida*. The results show that the actual value for activity

TABLE 15

C23O ASSAY RESULTS OF THE *XYlE* GENE UNDER CONTROL OF THE *TAC* PROMOTER IN *E.COLI* AND *P.PUTIDA*

	OD_{375nm} min^{-1} mg^{-1} protein	Level over pRLA5 in respective host
EXPERIMENT I		
2440(pRLA5.*tac*)	255.02	36
JM83(pRLA5.*tac*)	32.89	42
2440(pRLA5)	7.05	
JM83(pRLA5)	0.078	
EXPERIMENT II		
2440(pRLA5.*tac*)	105.26	10
JM83(pRLA5.*tac*)	57.40	68
2440(pRLA5)	10.52	
JM83(pRLA5)	0.844	
EXPERIMENT III		
2440(pRLA5.*tac*)	159.65	43
" "	156.95	42
" "	194.61	52
JM83(pRLA5.*tac*)	12.29	82
" "	14.59	97
" "	28.63	191
2440(pRLA5)	3.75	
JM83(pRLA5)	0.15	

expressed as OD_{375nm} units $minute^{-1}$ mg^{-1} total cell protein is about 8-10 fold lower in JM83 than 2440. Bagdasarian *et al* (1983) measured the expression of *xylE* when under the control of the *tac* promoter in *E.coli* and *P.putida*. They found that the values obtained in specific activity units min^{-1} mg^{-1} protein are approximately equal for the two species whether regulated ($lacI^q$ background + IPTG) or unregulated (lac^- background +/- IPTG). Hence the values obtained here for the C230 activity of the *tac* promoter in pRLA5 in the two species are not therefore directly comparable, since such a large difference in activity is not expected. When the data are expressed as relative levels over vector a result closer to that expected is obtained with the activity of the *tac* promoter more equal in the two hosts but slightly more active in *E.coli* than in *P.putida*. This suggests that the best way to present the data is as a level over the corresponding value of the vector.

6.3 MEASUREMENT OF C230 ACTIVITY

6.3.1 Effect of variation in protein concentration in the C230 assay

The cell-free extracts frequently showed considerable variation in their protein concentration, but it was desirable from a practical point of view and for accuracy not to dilute the extracts and to use a constant protein concentration in each assay, but rather to use a constant volume of cell-free extract and then correct the results for protein concentration. It was necessary, therefore, to determine whether protein concentration showed a direct proportionality to rate of production of 2-hydroxymuconic semialdehyde. An active promoter clone, pRLA5.272, was used and the level of C230 activity measured in *P.putida*. A series of dilutions of the cell-free extract was made, each differing by 20% and 5µl of each dilution was measured for rate of 2-HMS production. The results (Figure 44) show that there is close proportionality of protein concentration to OD_{375nm} units per minute. Hence, it is acceptable to employ a constant volume of

EFFECT OF VARYING PROTEIN CONCENTRATION
ON C23O ACTIVITY

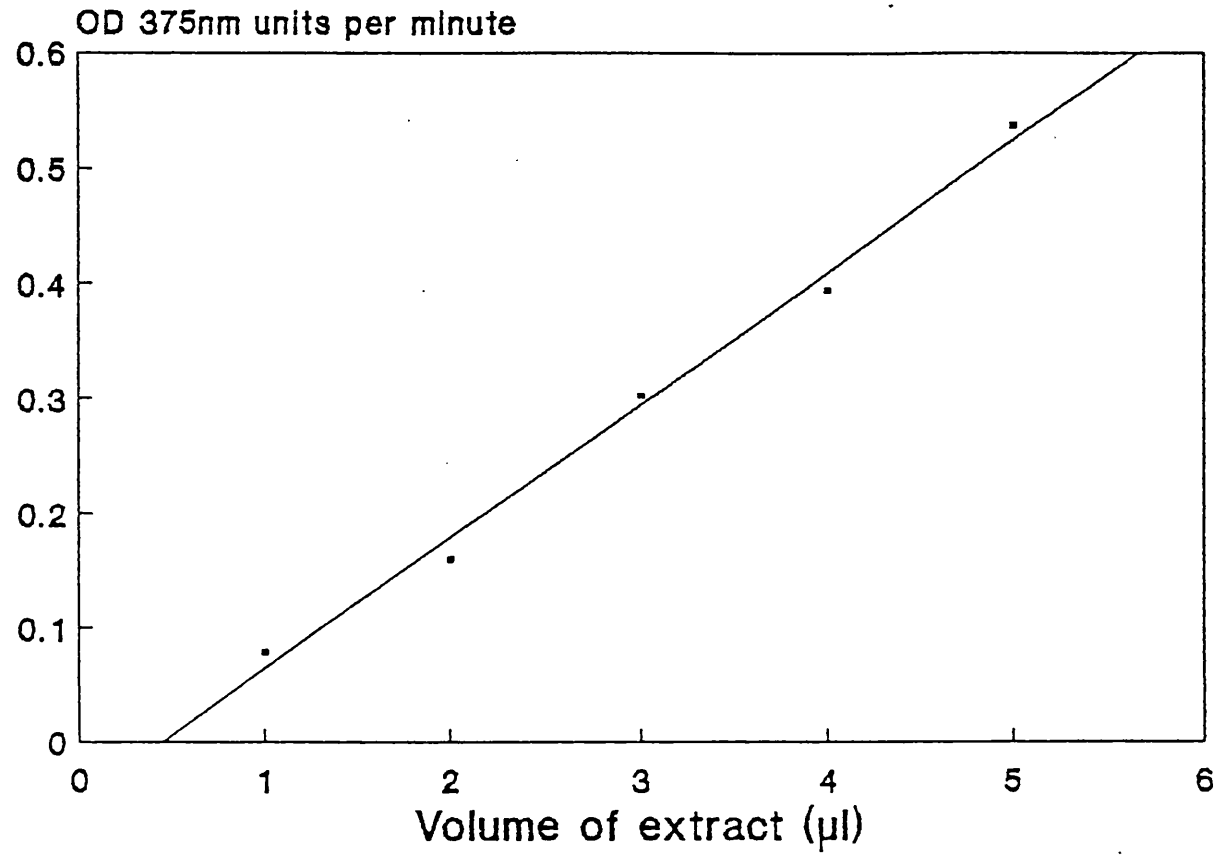


FIGURE 44

cell-free extract in the assay procedure, rather than a constant protein concentration.

6.3.2 Temperature of C230 assay

The temperature optima for growth of the two hosts is different, that of *P.putida* is 27°C and that of *E.coli* is 37°C. This temperature difference may have an effect on the rate of production of 2-HMS in the assay, since it is conceivable that the enzyme, C230, may exhibit maximal activity at 37°C in *E.coli*, but at 27°C in *P.putida*. Therefore, it might be necessary to perform the assays at these two temperature optima. The possibility that the enzyme has maximal activity at different temperatures in the two hosts was investigated by measuring the rate of 2-HMS production by 2440(pRLA5), JM83(pRLA5) and a high activity promoter 2440(pRLA5.90) and JM83(pRLA5.90) at the two temperatures. Each assay was carried out at a range of cell-free extract concentrations (from the above it was know that there was direct proportionality). The results (Figures 45-48) show that there was a maximum increase of overall activity of 10% at the higher temperature and that the reactions in both hosts follow the same kinetics with an increased rate at the higher temperature. There was no apparent species-dependent difference, therefore initial screening of the promoter clones was carried out at room temperature). Those clones which exhibited high promoter activity in *Pseudomonas*, but low activity in *E.coli*, were then assayed several times at a controlled temperature of 37°C (see 6.4 and Table 19).

6.3.3 Use of freshly-prepared extracts

Enzyme activity must be measured on the same day as production of the cell-free extract since storage of the extract at -20°C overnight reduced activity by about 25%. All reactions were therefore carried out with freshly-prepared extracts. The activity of the promoter clones was determined by measuring the maximum rate of production of 2-HMS by

EFFECT OF TEMPERATURE ON C23O ACTIVITY
2440(pRLA5)

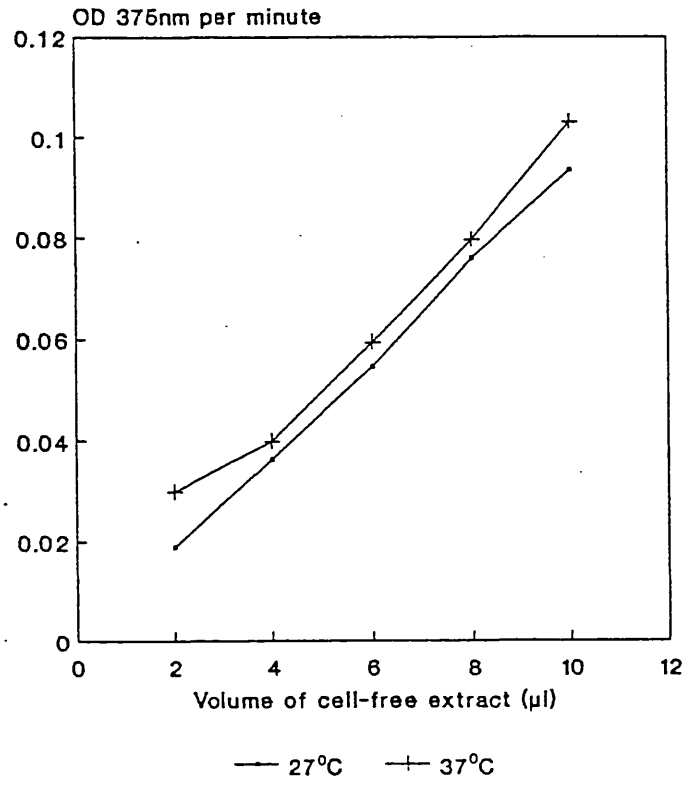


FIGURE 45

EFFECT OF TEMPERATURE ON C23O ACTIVITY
JM83(pRLA5)

Amount of cell-free extract (μl)	OD_{375nm} min^{-1} 27°C	37°C
6	0.00985	0.0138
10	0.0193	0.0230

FIGURE 46

EFFECT OF TEMPERATURE ON C23O ACTIVITY
2440(pRLA5.90)

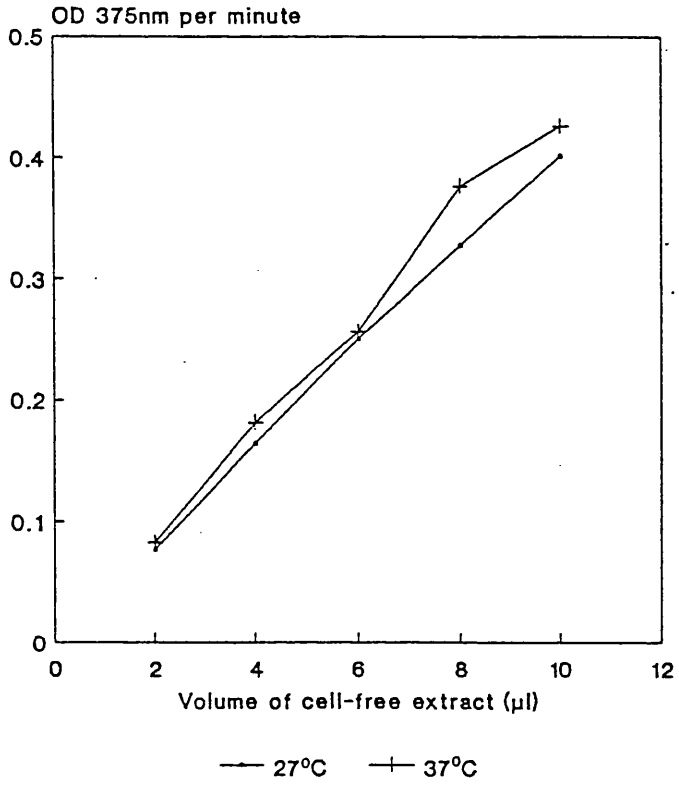


FIGURE 47

EFFECT OF TEMPERATURE ON C23O ACTIVITY
JM83(pRLA5.90)

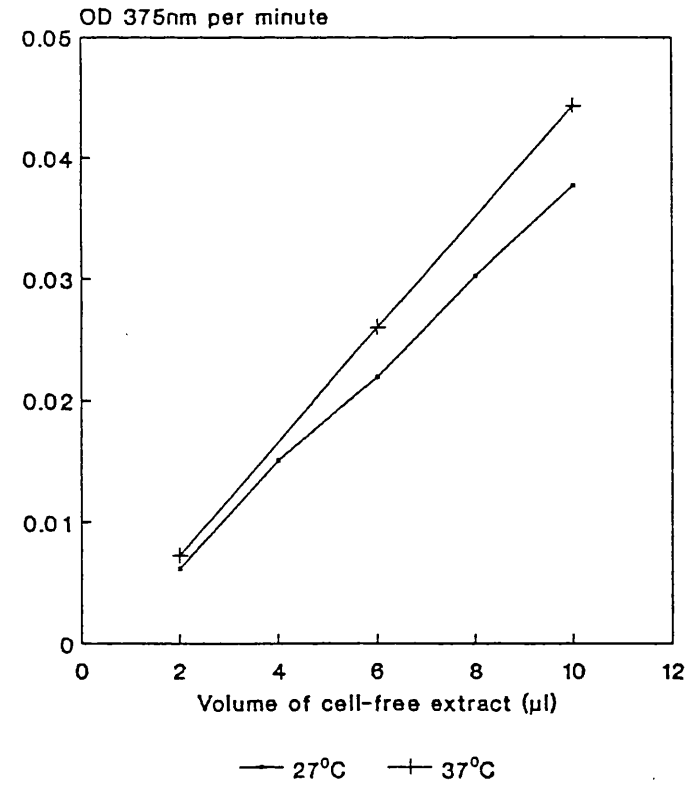


FIGURE 48

freshly-prepared cell-free extract (10μl) at RT over 10 minutes or at 37°C over 2.5 minutes, except for vector controls which were measured over 15 minutes at RT or over 10 minutes at 37°C. The data was expressed as OD_{375nm} units min^{-1} mg^{-1} total cell protein.

6.4 C230 ASSAY RESULTS

Figure 49 shows a typical curve of rate of production of 2-HMS from a highly active clone, pRLA5.272. Table 16 shows the C230 results for the 340 clones isolated which appeared to have an increased C230 activity over the vector background level. Of these 127 (37%) exhibited activity of at least 5 times that of pRLA5 measured on the same occasion. The DNA from these was extracted and *E.coli* JM83 or MC1061 was transformed with this DNA. One resultant transformant grown up in 10ml and assayed for C230 activity in parallel with the same clone in 2440. Table 17 shows the results of the comparative assays of the clones in both hosts. An arbitrary grouping system was used, taking as evidence for enhanced promoter activity a value of 5x the C230 activity exhibited by the vector pRLA5 assayed in parallel. The clones were grouped into three classes according to the value of the ratio of C230 activity of the clone to that of the vector in *P.putida* compared with the ratio of C230 activity of the clone to that of the vector in *E.coli*; greater than 5 (Group I), between 0.2 and 5 (Group II) and 0.2 or less (Group III).

67% of the total number of clones in all three groups belonged to Group II. This is quite interesting, since it suggests that, although it is not known whether all of these represent true active promoters, a sequence which is quite active as a promoter in one species is also active in the

A TYPICAL CURVE OF C23O ACTIVITY

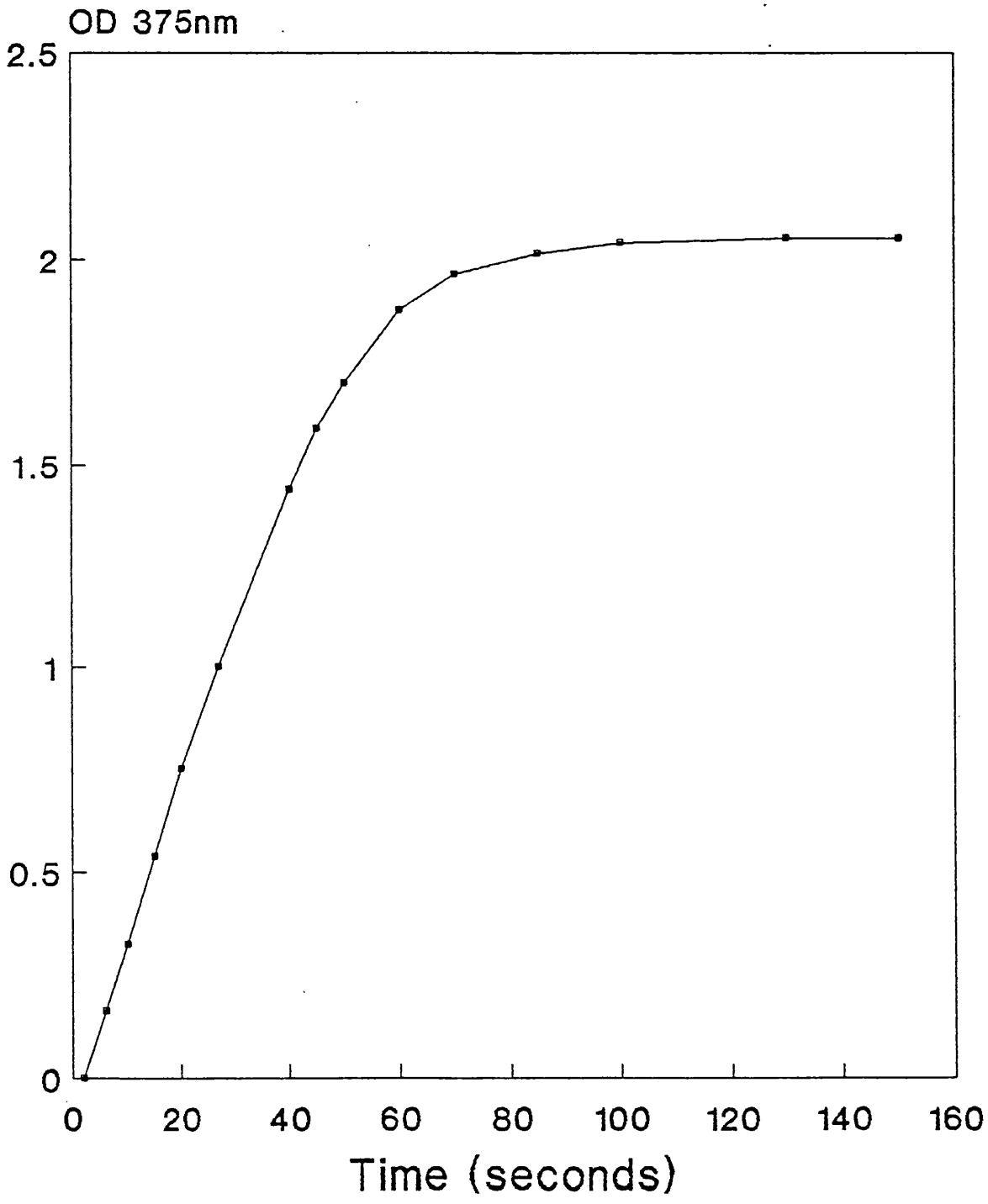


FIGURE 49

TABLE 16

COMPARISON OF C230 ACTIVITIES EXHIBITED BY THE PROMOTER CLONES IN *E.COLI* AND *P.PUTIDA*

	C230 ACTIVITY (OD_{375nm} $min^{-1}mg^{-1}protein$) IN			RATIO OF C230 ACTIVITY TO THAT OF VECTOR pRLA5 IN		
	P.PUTIDA (2440)	*E.COLI* (MC1061 M) (JM83 J)		*P.PUTIDA* (x)	*E.COLI* (y)	x/y
GROUP I						
pRLA5.4	9.12	0.166	M	6.9	1.3	5.3
pRLA5.8	55.72	0.252	M	25	1.5	16.7
pRLA5.64	12.71	0.146	J	13	1.5	8.7
pRLA5.66	44.29	0.550	J	45	5.6	8.0
pRLA5.69	11.23	0.369	J	5.0	0.9	5.6
pRLA5.90	305.08	1.886	J	53	3.9	13.6
pRLA5.110	42.25	0.538	M	8.9	1.1	8.1
pRLA5.121	28.15	0.677	M	7.5	1.0	7.5
pRLA5.185	28.74	1.94	M	11	1.6	6.9
pRLA5.187	73.79	4.103	M	28	3.3	8.5
pRLA5.208	8.31	0.349	M	4.2	0.6	7.0
pRLA5.258	47.58	0.571	M	18	2.9	6.2
pRLA5.279	38.72	0.576	M	14	2.7	5.2
pRLA5.273	124.25	5.061	M	35	3.4	10.3
pRLA5.275	10.23	0.730	M	2.9	0.5	5.8
GROUP II						
pRLA5.1	27.30	4.05	M	21	31	0.7
pRLA5.7	26.89	1.47	M	20	11	1.8
pRLA5.14	15.49	3.52	M	12	27	0.4
pRLA5.17	14.95	3.46	M	11	26	0.4
pRLA5.18	13.20	3.66	M	10	28	0.4
pRLA5.19	25.20	3.19	J	11	32	0.3
pRLA5.23	31.70	2.85	J	13	29	0.4
pRLA5.54	19.28	1.27	J	8.3	13	0.6
pRLA5.55	21.49	0.536	J	9.3	5.7	1.6
pRLA5.56	10.14	0.408	J	4.4	4.3	1.0
pRLA5.57	9.00	0.115	J	3.9	1.2	3.3
pRLA5.58	28.46	0.882	J	12.3	9.3	1.3
pRLA5.59	16.87	0.842	J	17	8.6	2.0
pRLA5.71	30.70	3.95	J	14	9.4	1.5
pRLA5.73	169.25	26.33	J	57	50	1.1
pRLA5.75	138.07	31.23	J	46	60	0.8
pRLA5.82	70.30	7.13	J	24	14	1.7
pRLA5.83	116.41	4.70	M	18	9.3	1.9
pRLA5.85	117.65	10.18	J	20	21	1.0
pRLA5.87	205.5	60.21	J	36	123	0.3
pRLA5.91	5.30	0.480	J	0.92	0.98	0.9
pRLA5.96	56.60	25.71	M	24	56	0.4
pRLA5.97	43.94	13.83	M	19	30	0.6
pRLA5.98	69.04	4.52	M	30	9.8	3.0
pRLA5.102	12.44	6.11	M	5.4	13	0.4
pRLA5.104	18.10	6.61	M	7.8	14	0.6
pRLA5.111	215.18	19.15	M	44	39	1.1

pRLA5.121	28.15	0.677	M	7.5	1.0	7.5
pRLA5.123	26.44	1.21	M	7.0	1.8	3.9
pRLA5.131	18.03	7.94	M	5.5	2.4	2.3
pRLA5.141	7.27	1.84	M	2.2	2.5	0.9
pRLA5.143	66.52	15.81	M	5.3	13	0.4
pRLA5.148	18.70	0.72	M	4.2	1.1	3.8
pRLA5.151	22.63	8.57	M	5.1	14	0.4
pRLA5.154	48.37	6.43	M	11	10	1.1
pRLA5.166	91.64	5.91	M	21	9.4	2.2
pRLA5.168	30.26	1.07	M	6.9	1.7	4.1
pRLA5.170	41.51	5.35	M	10	3.6	2.8
pRLA5.171	4.17	2.23	M	1.0	1.5	0.7
pRLA5.178	33.03	15.28	M	8.1	10	0.8
pRLA5.179	46.61	27.49	M	11	19	0.6
pRLA5.182	77.51	17.38	M	29	14	2.1
pRLA5.183	73.82	18.15	M	28	15	1.9
pRLA5.184	81.58	25.11	M	31	20	1.6
pRLA5.190	62.54	15.18	M	23	12	1.9
pRLA5.193	135.02	8.96	M	89	22	4.0
pRLA5.198	82.72	16.09	M	55	39	1.4
pRLA5.203	14.06	1.08	M	9.3	2.6	3.6
pRLA5.204	19.70	7.29	M	10	12	0.8
pRLA5.205	7.75	5.01	M	3.9	8.9	0.4
pRLA5.206	37.62	19.05	M	19	34	0.6
pRLA5.207	30.59	12.00	M	15	21	0.7
pRLA5.210	11.25	7.74	M	5.7	14	0.4
pRLA5.213	40.69	2.39	M	14	5.2	2.7
pRLA5.220	15.37	5.00	M	5.3	11	0.5
pRLA5.221	13.49	4.58	M	4.6	10	0.5
pRLA5.230	2.74	0.844	M	1.1	2.8	0.4
pRLA5.237	25.96	12.31	M	10	40	0.3
pRLA5.251	10.97	0.630	M	3.5	1.0	3.5
pRLA5.238	14.69	7.08	M	4.7	1.1	4.3
pRLA5.240	140.55	11.71	M	45	18	2.5
pRLA5.245	5.03	0.274	M	1.9	1.4	1.4
pRLA5.246	28.77	1.59	M	11	8.0	1.4
pRLA5.247	21.36	1.66	M	8.0	8.4	1.0
pRLA5.255	19.86	1.95	M	7.4	9.8	0.8
pRLA5.256	13.16	1.38	M	4.7	6.5	
pRLA5.259	22.22	0.349	M	8.3	1.8	4.6
pRLA5.260	18.28	1.86	M	6.8	9.4	0.7
pRLA5.265	22.29	1.10	M	8.0	5.2	1.5
pRLA5.268	141.69	12.17	M	51	58	0.9
pRLA5.271	67.93	19.20	M	19	14	1.4
pRLA5.272	76.87	44.35	M	22	31	0.7
pRLA5.280	21.60	1.30	M	7.8	6.1	1.3
pRLA5.306	70.55	17.56	J	13	32	0.4
pRLA5.308	142.05	16.59	J	26	30	0.9
pRLA5.319	51.84	10.50	J	9.4	19	0.5
pRLA5.320	42.50	2.18	J	7.7	4.0	1.9
pRLA5.322	137.02	21.85	J	18	24	0.8
pRLA5.323	87.47	22.14	J	11	24	0.5
pRLA5.324	156.15	24.44	J	20	27	0.7
pRLA5.325	107.15	22.57	J	14	25	0.6
pRLA5.326	75.07	20.46	J	10	22	0.5
pRLA5.327	54.32	20.92	J	7.0	23	0.3
pRLA5.341	178.51	22.68	J	16	33	0.5
pRLA5.337	214.82	23.70	J	19	34	0.6
pRLA5.335	143.59	8.15	J	13	12	1.1
pRLA5.339	168.93	25.56	J	15	37	0.4
pRLA5.340	238.67	15.21	J	22	22	1.0
pRLA5.356	219.75	26.02	J	44	32	1.4

pRLA5.357	53.27	4.72	J	11	5.8	1.9
pRLA5.358	48.58	12.65	J	10	16	0.6

GROUP III

pRLA5.20	21.99	14.60	J	9.3	147	0.06
pRLA5.21	22.96	4.63	J	9.7	47	0.2
pRLA5.24	22.96	10.49	J	9.7	106	0.09
pRLA5.34	67.82	62.96	J	29	27	0.1
pRLA5.67	10.53	4.78	J	11	49	0.2
pRLA5.78	14.45	12.12	J	4.8	23	0.2
pRLA5.79	15.72	17.96	J	5.3	34	0.2
pRLA5.81	12.87	18.66	J	4.3	36	0.1
pRLA5.89	55.67	17.62	M	8.5	35.0	0.2
pRLA5.92	75.00	36.33	J	13	74	0.2
pRLA5.93	72.87	31.85	M	11	63	0.2
pRLA5.94	18.86	9.05	M	2.9	18	0.2
pRLA5.101	10.69	12.22	M	4.6	27	0.2
pRLA5.122	104.93	84.36	M	28	124	0.2
pRLA5.142	18.97	13.39	M	1.5	11	0.1
pRLA5.144	16.42	14.86	M	1.3	12	0.1
pRLA5.145	17.07	9.80	M	1.4	7.9	0.2
pRLA5.152	20.02	25.08	M	4.5	40	0.1
pRLA5.172	90.64	173.97	M	22	118	0.2
pRLA5.176	51.67	146.00	M	13	99	0.1
pRLA5.189	48.12	139.55	M	18	113	0.2
pRLA5.216	46.94	62.17	M	16	135	0.1
pRLA5.222	8.73	21.06	M	3.0	46	0.07
pRLA5.236	13.06	6.58	M	5.1	22	0.2
pRLA5.239	11.60	14.51	M	4.5	48	0.09
pRLA5.249	21.13	8.57	M	7.9	43	0.2
pRLA5.250	14.08	6.69	M	5.1	32	0.2
pRLA5.262	17.89	5.68	M	6.4	27	0.2
pRLA5.317	47.06	22.65	J	8.6	41	0.2
pRLA5.328	52.78	45.04	J	6.8	49	0.1
pRLA5.347	58.25	22.02	J	5.3	32	0.2

other. Included in this group may be some which have *E.coli*-like promoter elements and *Pseudomonas*-like elements within a single sequence, such that transcription occurs in response to a different signal contained in the same DNA fragment in the two species. 22% of clones belonged to Group III and these are expected to have good *E.coli*-like promoter sequences which are also recognised by the *P.putida* transcription machinery. Only 11% of clones belonged to Group I. This may be because either the conditions for allocation to the groups is too stringent, or it is a reflection of the proportion of promoters which have promoter activity only in *Pseudomonas*. This group presumably contains either positively regulated promoters whose activator is specific to *Pseudomonas* or high constitutive expression from a promoter recognisable by the transcription machinery of *P.putida* but not by that of *E.coli*.

Those clones in Group I were assayed for C230 activity several times at 37°C and in JM83 as the host *E.coli* strain, to give more reproducible and accurate measurements (Table 17). The inserts of promoter clones from Group I were sequenced. Some (pRLA5.8 and pRLA5.110), apart from the fact that they may contain more than one active promoter, were too large to be readily sequenced and so they were not analysed further. pRLA5.69, pRLA5.208 and pRLA5.275 were also not sequenced, since they exhibited relatively low C230 activity and therefore they presumably did not contain a *Pseudomonas* promoter. In addition to those from Group I, one clone from each of Groups II and III, pRLA5.98 and pRLA5.172, was also sequenced for comparison. Two other clones were analysed in addition to the Group I clones, although their promoter activity in *E.coli* was not less than 5 times of the activity of pRLA5 in *E.coli*. These were selected because they had very high activity in *P.putida*, but lower activity in *E.coli*. Thus, they may contain sequences not highly related to *E.coli* promoter sequences, but may be similar to sequences which may constitute an efficient promoter signal in *P.putida*.

TABLE 17

REPEATED C230 ASSAYS ON GROUP I PROMOTER CLONES

CLONE	C230 ACTIVITY (OD_{375nm} $min^{-1}mg^{-1}protein$) IN		LEVEL OVER pRLA5 IN		
	2440	JM83	2440 (x)	JM83 (y)	x/y
pRLA5.90	214	1.4	37	2.3	16.1
	117;237	3.7	18;37	4.8	3.8;7.7
	178	2.5	27	2.7	10.0
	200	1.8	35	3.0	11.7
pRLA5.187	204	13	43	3.8	11.3
	196	6.2	29	18	1.6
	132	4.9	23	8	2.9
	348	14	55	19	2.9
	138	7.8	24	13	1.8
pRLA5.193	360	20	53	57	0.9
	262	25	46	42	1.1
	523	27	82	36	2.3
	345	29	44	34	1.3
pRLA5.240	314	18	63	37	1.7
	112	4.2	38	34	1.1
pRLA5.273	300	34	44	98	0.4
	248	17	44	28	1.6
	408	25	64	32	2.0
	252	23	44	43	1.0

From our current knowledge of promoter signals in *Pseudomonas*, it is possible to speculate as to what the cloned inserts represent in terms of promoter sequences. One would expect the majority of clones to have no detectable promoter activity since they represent cloned DNA fragments from within structural genes or operons. Some clones would be likely to exhibit *E.coli*-like promoters which are active in *Pseudomonas*. These would give rise to a range of promoter activities dependent on the degree of similarity to the *E.coli* concensus sequence and other regions that may influence promoter activity. Others will encode the nitrogen-regulated (ntr or nif) promoters known to be the functional promoter for some *Pseudomonas* genes. Some clones which exhibit high promoter activity in *Pseudomonas*, but low activity in *E.coli*, will be positively regulated where the activator molecule is not synthesised by *E.coli*. It is likely that there will be a set of promoters which are constitutively active and give rise to high level expression in *Pseudomonas*, but have no similarity to the concensus *E.coli* or nif sequences and thus have low activity in *E.coli*. In addition to regions of DNA which encode promoters that are active *in vivo*, it is probable that other regions of DNA will be cloned which are not used as promoters *in vivo*, but will have fortuitous similarity to *E.coli* or nif or *Pseudomonas* promoter sequences and will give rise to detectable activity when cloned in a promoter-probe vector.

6.4.1 Inconsistencies in the data

There was some variation between repeated assays of the Group I clones in terms of level over the vector (Table 17). In order to test whether this reflected a variation between cultures, duplicate cultures of *P.putida* carrying a clone were grown on the same day and assayed individually (Table 18). The variation in C230 activity of duplicate samples measured on the same day was between 6% and 27%. The variation in level over vector for *P.putida* cultures assayed on the same day, since values of C230 activity

TABLE 18

C230 ASSAYS OF *P.PUTIDA* CULTURES GROWN IN DUPLICATE

CLONE	C230 Activity in *P.putida* (OD_{375nm} $min^{-1}mg^{-1}$ protein)	Level over vector in *P.putida*
pRLA5.240	219.08 ; 299.60	76 ; 104
pRLA5.241	5.94 ; 5.24	2 ; 1.8
pRLA5.357	61.54 ; 68.76	13 ; 15
pRLA5.358	64.96 ; 69.21	14 ; 15
pRLA5	1.37 ; 1.55	

for 2440(pRLA5) were the same in each case, was also between 6% and 27% There was much greater variation in C230 activity between cultures assayed on different days, from 39% to 64% in *P.putida* and 31% to 77% in *E.coli*. The variation in enzyme activity level over vector for cultures assayed on different days was based on the value obtained for pRLA5 on that day giving variation from 31% to 58% in *P.putida* and from 8% to 80% in *E.coli*. In general there was greater variation in C230 activity by clones assayed on different days than on the same day, and although the variation in level over vector for *P.putida* cultures assayed on different days was within the same range as for C230 activity, for *E.coli* there was much greater variation. It is possible that this increased variation was due to the value for C230 activity of JM83(pRLA5); this being small it is more subject to inaccuracy in its determination.

6.5 SEQUENCING THE PROMOTER INSERTS

The inserts were to be sequenced on both strands by cloning them into both M13mp18 and M13mp19. They were each sequenced several times on at least one strand. Due to time limitations some inserts (pRLA5.187 and pRLA5.240), were only sequenced on one strand. To be certain of sequencing the insert and not any contaminating DNA, rather than isolating the insert on the small fragment by agarose gel electrophoresis, it was isolated on a fragment of DNA including part of the vector molecule. However, this was only possible for cloning in M13mp19, since there were no suitable restriction sites upstream of the insert. Therefore, for cloning in mp18 the insert was isolated on a *Hin*dIII/*Xho*I fragment. Cloning into mp19 was either by isolating a *Cla*I/*Eco*RI fragment (this includes a 1.5kb region from pRLA5) and inserting this into M13mp19 cut with *Acc*I and *Eco*RI, or by isolating the insert on an *Eco*RI/*Kpn*I fragment and inserting into mp19 cut with the same enzymes, or by isolating a *Hin*dIII/*Xho*I fragment and inserting into M13mp19 cut with *Hin*dIII and *Sal*I.

The cloned sequences were determined by analysis on 6% acrylamide denaturing gels. The nucleotide sequences are shown in Figures 50-56. Only the insert region is shown, but not all of the sequence was obtained for all of the fragments due to limitations in the available time.

6.6 ANALYSIS OF THE CLONED PROMOTER SEQUENCES (Computer analysis was carried out using the Wisconsin Data Handling Package)

To carry out any meaningful analysis of the sequence data, it is necessary to determine whether the cloned DNA was derived from *P.putida*. The most accurate method would be to use the labelled insert as a probe to chromosomal *P.putida* DNA. Unfortunately, due to time limitations, this was not carried out. However, on the basis of the nucleotide sequence some checks could be made by, for example, a comparison of the inserts with known sequences in a database. The sequences were compared against the EMBL and GENBANK databases and no homology of >50% over the whole insert sequence was revealed. Although this does not of course mean that the cloned sequences are not of some other derivation other than the *P.putida* chromosome, it does mean that the DNA is not from M13 a probable source of contaminant DNA. There are very few *Pseudomonas* sequences contained in the nucleotide sequence databases and therefore it is not surprising that there was no homology of the cloned inserts to them. The only *P.putida* sequences held in the databases are from the degradative plasmids.

Having isolated regions of DNA showing promoter activity and ensuring the DNA is from *P.putida*, the next step would be to determine the start of transcription initiation in both *E.coli* and *P.putida* by S1 nuclease or reverse transcriptase mapping experiments. This would provide several important pieces of information, namely, the base or bases at which transcription is initiated in both species, which would indicate whether the same promoter was being utilised in the two species. This information would indicate the probable region (generally up to 40bp upstream of the

transcription start point) where the promoter lies. With insert sequences of up to 500bp in length this experimental data is obviously essential for determining the possible promoter sequence utilised *in vivo*.

Since this information was not available for the cloned fragments described here, any analysis of this sequence data is necessarily highly tentative and awaits confirmation or otherwise by the experiments described above. The sequences were searched using the "FIND" program for the best possible -35, -10 sequences allowing a spacing of 17 ± 1bp between them, and nif sequences allowing only a spacing of 5bp between the -24 and -12 sequences and the putative homologies are indicated in Figures 50-56. Since there was the possibility that during the cloning experiment *in vitro* ligation of *Sau*3A fragments had occurred, the sequences were examined for *Sau*3A recognition sequences. The presence of such sites indicates that the fragments bounded by these sites may not have come from contiguous regions on the *P.putida* chromosome. A database was constructed of all the regions upstream of the ATG translation initiation codon from published *Pseudomonas* genes and those of related species (Table 19). Notation for homology comparisons is as follows, upper case signifies an exact match, lower case a mismatch and "N" represents any base.

6.6.1 GROUP I CLONES

pRLA5.90 (Figure 50)

This was sequenced on both strands but neither terminus was obtained. It was 384bp long and had 2 *Sau*3A sites. During the subcloning of some of the insert sequences into M13 an unusual sequence was sometimes inadvertently cloned. This sequence was also contained within the clone pRLA5.90, ligated at the *Sau*3A site at position 102bp. It was not determined whether the high promoter activity exhibited by this clone was due to this sequence 3' to the *Sau*3A site at 102bp, or the sequence 5' to the *Sau*3A site. It is not known whether this 5' region was derived from

FIGURES 50-56 Nucleotide sequences of promoter clones.

Sequences are shown single-stranded 5' to 3'. Homologies to -35 and -10 *E.coli* promoter sequence elements are indicated by single overlining and to nif promoters by double overlining. Putative ribosome binding sites are indicated by dashed overlining. The arrow in Figure 50 (pRLA5.90) indicates the position of the *Sau*3A site at which it appears ligation has taken place.

FIGURE 50

pRLA5.90

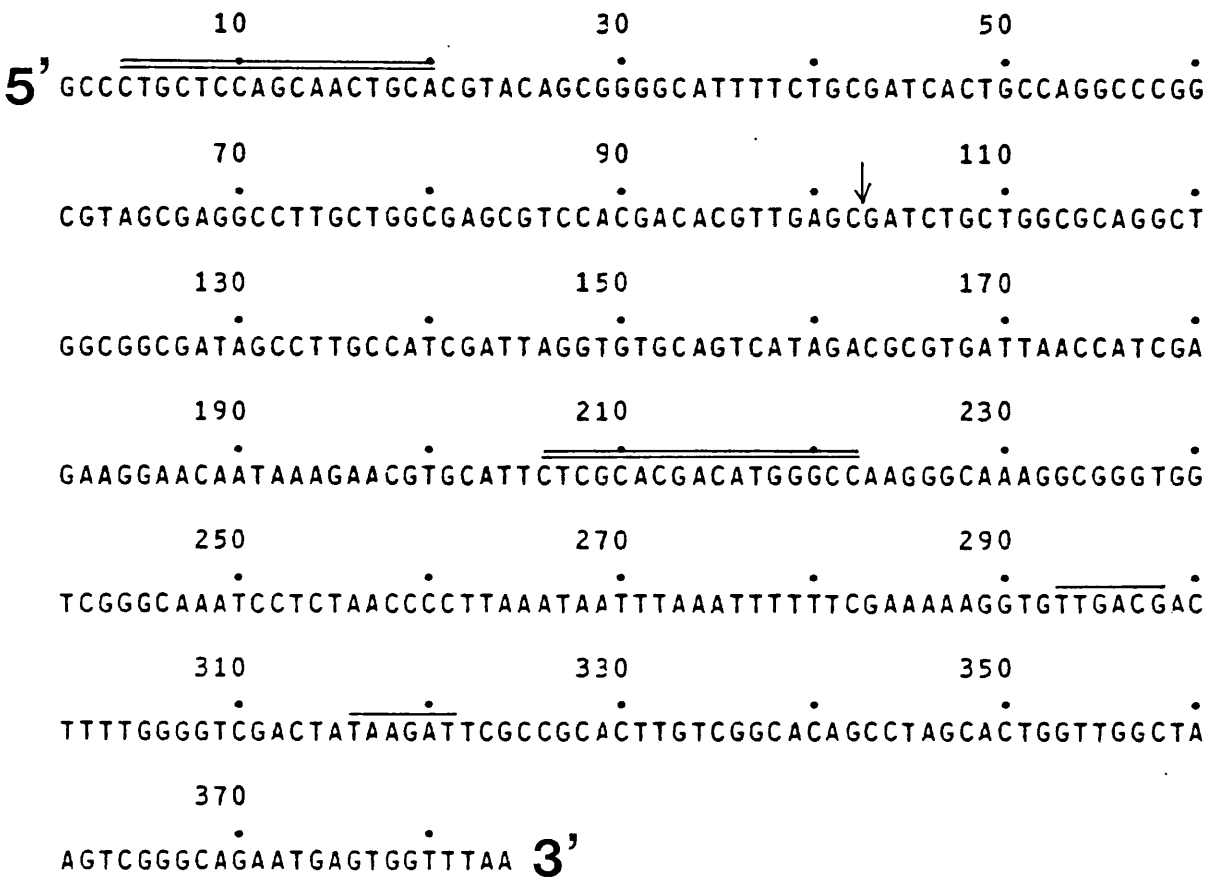


FIGURE 51

pRLA5.187

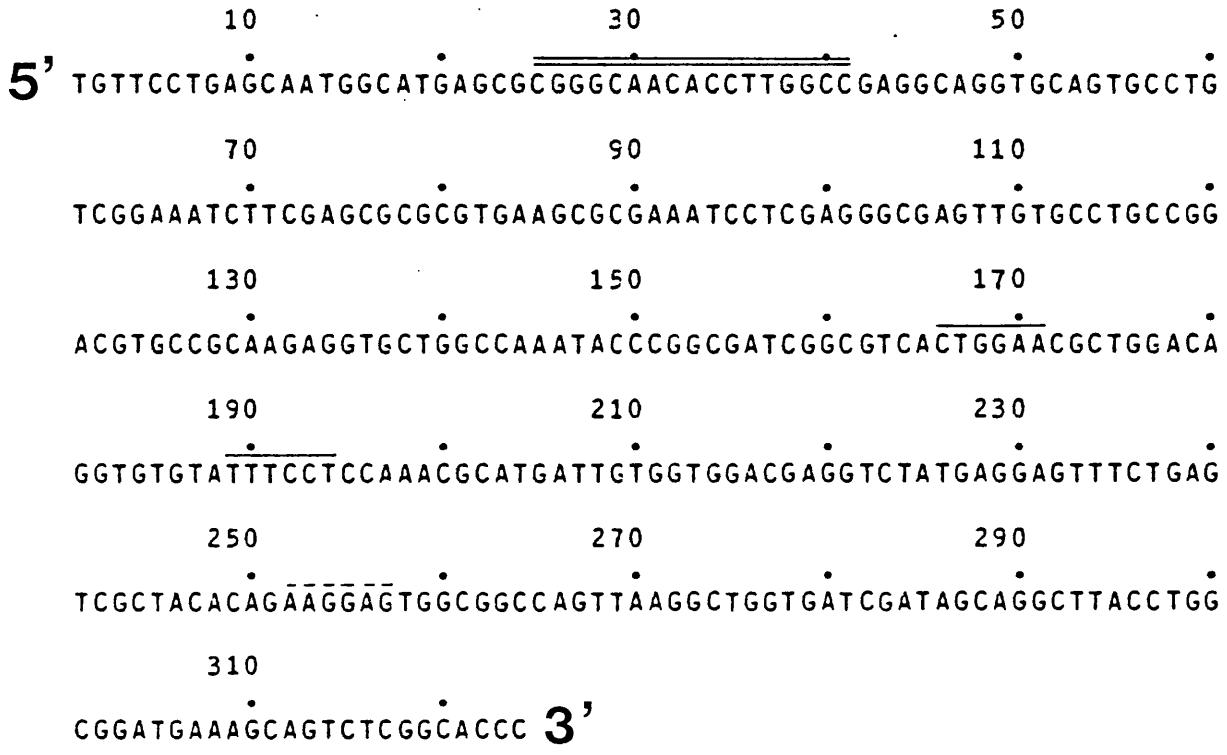


FIGURE 52

pRLA5.193

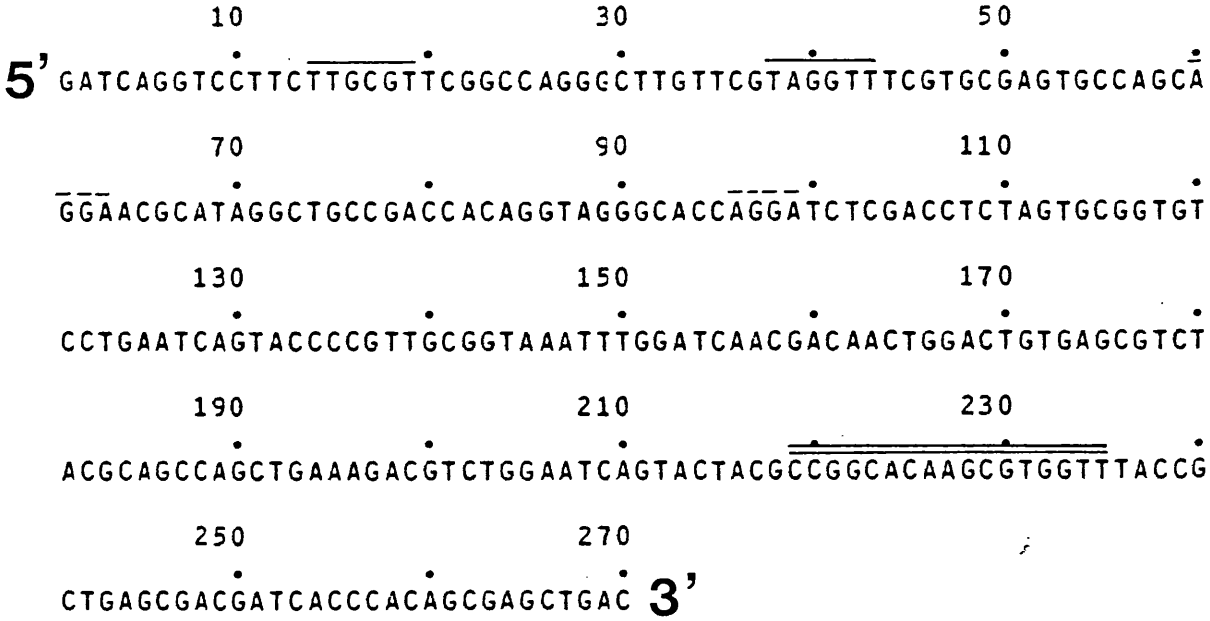


FIGURE 53

pRLA5.240

```
          10                  30                  50
5' ATTAAACAACAAATATTTAATAATGAGGGATAAAAAAGATACCGCCGCAGATATAATTCAG

          70                  90                  110
   TCCAGCCCTGAGTTAAGGCCTTTTGTAGAACGTAATTCGAAGTTGACTGTGCTTGCTCGC

         130                 150
   TTGAGGGGGTCGCCTTCATCTAGCAGAAATATATTACCGGA 3'
```

FIGURE 54

pRLA5.273

```
          10                  30                  50
5' ACCAGTTTCAACGGTTCCTCCAGGCGTCCTTGCTGCTCCACTTGCCGAGCAAGGCCTTGG

          70                  90                  110
   CGGCTGCTGGTTGCTGGGCAAAGGCTCATGGAAGGTCACCAGACAACCGCCCAGCAGCAG

         130                 150                 170
   GAATGACAAAGTCAACAACACCACACGCCAGGCTCTCATCAGGTTCTCCTTGAAAAGGTC

         190                 210                 230
   GTTCTCGGCTCAGGCCATGCCAGCACCAGCTGCAGGTATCGGGTAAGCATAAGCGAGCAT

         250                 270                 290
   CTGCCCATCGGCTTGCGCATTCGCTTGCTTGAGCAGGCCTTGATATTCCAACTGCTCGAT

         310                 330                 350
   AATCGCCGTCAACATCTGTCATCCTCTCCGCTCCCAACCCACGACGTGCAGCAATTGCCG

         370
   GTGCCGTGCAGCAG 3'
```

FIGURE 55

pRLA5.98

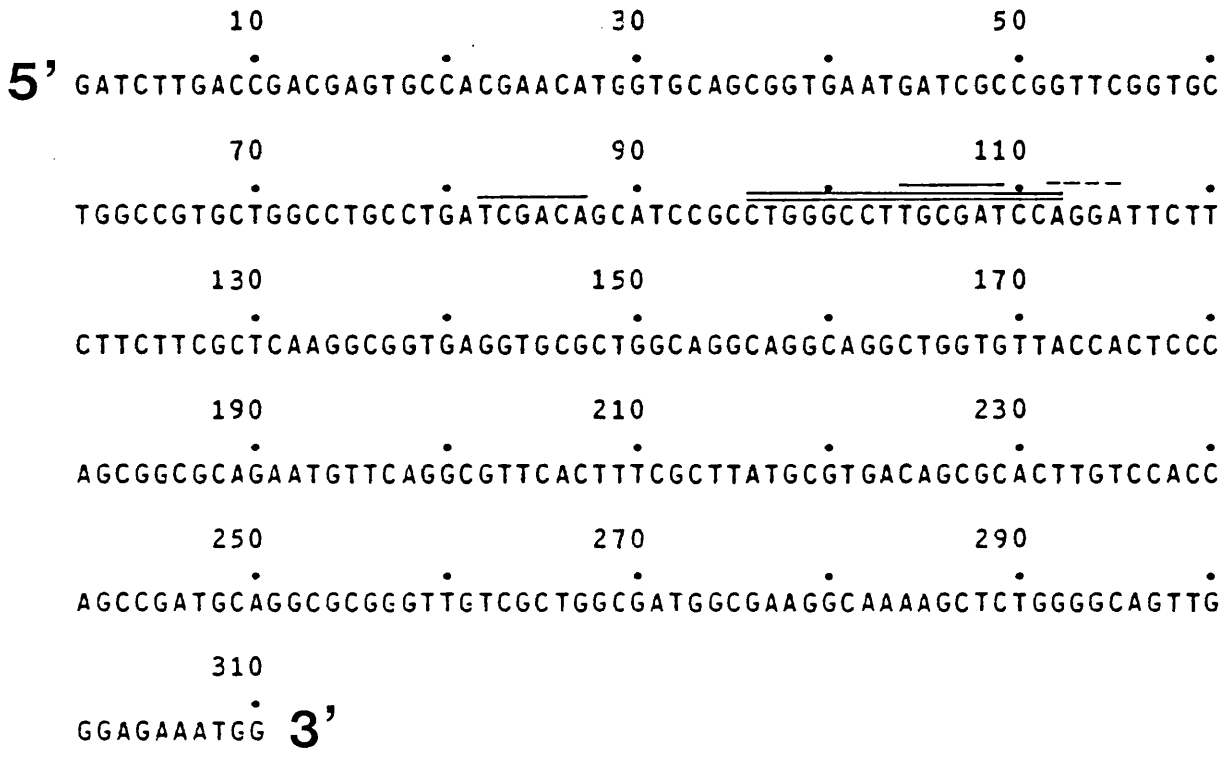


FIGURE 56

pRLA5.172

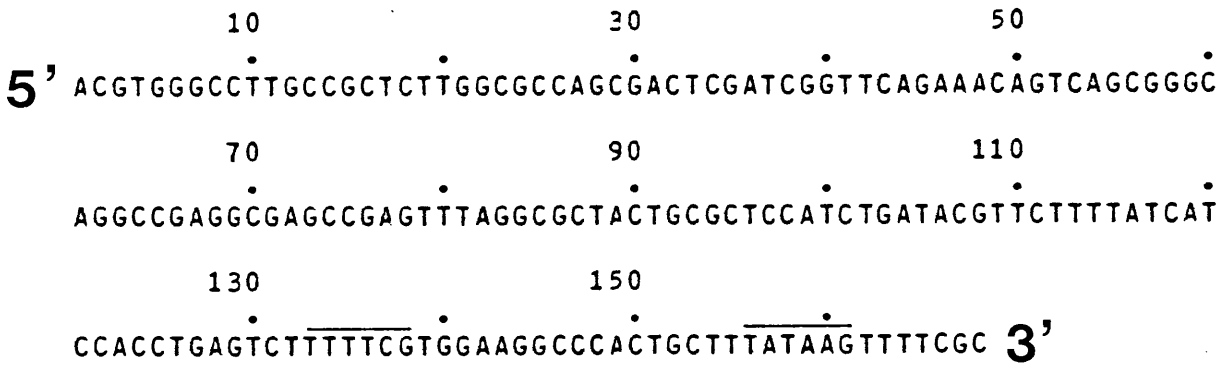


TABLE 19

DATABASE OF UPSTREAM REGIONS OF GENES FROM *PSEUDOMONAS* AND RELATED SPECIES

P.aeruginosa	
argF	Itoh *et al* 1988
algD alginate production	Deretic *et al* 1987c
toxA exotoxin A	Grant and Vasil 1986
toxR exotoxin A regulation	Wozniak *et al* 1987
phosphate regulated heamolysin	Pritchard and Vasil 1986
ompF outer membrane protein	Duchene *et al* 1988
pilin K122-4	Pasloske *et al* 1988
pilin P1	Pasloske *et al* 1988
P.putida	
xylCAB	Inouye *et al* 1984a
xylDEGF	Inouye *et al* 1984b
xylE constitutive mutant	Inouye *et al* 1986
xylR	Spooner *et al* 1986
xylS	Spooner *et al* 1986
ndoA naphthalene dioxygenase	Kurkela *et al* 1988
catBC dissimilation of benzoate	Aldrich and Chakrabarty 1988
clcB 3-chlorocatechol degradation	Frantz and Chakrabarty 1987
benzene oxidation gene 1	Irie *et al* 1987b
P.atlantica	
agrA B-agarase	Belas 1989
P.fluorescens	
inaW ice nucleation	Warren *et al* 1986
P.savastanoi	
ptz cytokinin biosynthesis	Powell and Morris 1986
iaaM indoleacetic acid biosynthesis	Yamada *et al* 1985
P.stutzeri	
nosZ nitrous oxide reductase	Viebrock and Zumft 1988
P.syringae	
copper resistance	Mellano and Cooksey 1988
Pseudomonas sp.	
carboxypeptidase G2	Minton and Clarke 1985
acyI cephalosporin acylase	Matsuda *et al* 1987
acyII cephalosporin acylase	Matsuda *et al* 1987
clcD dienelactone hydrolase	Frantz *et al* 1987
vanAB vanillate demethylase	Brunel and Davison 1988
Caulobacter crescentus	
28kD flagellin	Gill and Agabian 1983
hook P11 promoter	Mullin *et al* 1987
hook P11.1 promoter	Mullin *et al* 1987
Methylomonas clara	
a shotgun cloned promoter	Metzler *et al* 1988

Table 19 ctd.

Aeromonas sobria	
aerA aerC aerolysin	Husslein *et al* 1988
Zymomonas mobilis	
shotgun cloned promoters A,B,C	Conway *et al* 1987a
gap Glyceraldehyde-3-phosphate dehydrogenase	Conway *et al* 1987b
Rhodobacter capsulatus	
pufB light-harvesting I complex	Bauer *et al* 1988

P.putida. It had no homology to M13 nor to sequences in the EMBL database. In both the 5' sequence of pRLA5.90 and the 3' sequence there were 2 sequences which showed homology with 4 mismatches to the nif concensus sequence. The clone also exhibited a potential ribosome binding site but the *Sau*3A site at which it appears religation has taken place lies between the nif sequence and this potential ribosome binding site and there is no translation initiation codon at the distance required to constitute the start of an open reading frame. The overall G+C content of the insert was 54%, the upper sequence having a G+C content of 66% and the lower sequence 50%. The best -35 and -10 *E.coli*-like sequences occurred in the 3' region, they had a spacing of 17bp and 5 out of 6 matches to the concensus in the -35 region including all the most highly conserved bases and 4 out of 6 matches to the concensus in the -10 region including all 3 most highly conserved bases. In the 5' region there was no homology to the *E.coli* concensus promoter sequence.

pRLA5.187 (Figure 51)

This insert was only sequenced on one strand and in one orientation. Neither terminus was obtained. It was 324bp long and had 2 internal *Sau*3A sites. It had a G+C content of 59%, closer to that of the *P.putida* chromosome than the *E.coli* chromosome. The best fit to the nif concensus had 4 mismatches. The best homology to the *E.coli* concensus had 3 out of 6 matches in both regions with a spacing of 17bp and the hexamers had 2 of the most highly conserved bases in both regions. There were two sequences with homology to the predicted ribosome binding site 3' to both the putative nif and *E.coli* promoter sequences, but there were no translation initiation codons at the requisite distance.

pRLA5.193 (Figure 52)

Both strands were sequenced and both the termini of the sequence were

obtained. The insert was 270bp long and had 4 internal *Sau*3A sites. The sequence had a G+C content of 57%, closer to that of the *P.putida* chromosome than the *E.coli* chromosome. The best homology to the nif concensus included 4 mismatches. The best homology to the *E.coli* concensus had a spacing of 18bp between the hexamers and 3 out of 6 matches in both regions, with all 3 most highly conserved bases present in both regions. There were two sequences with homology to the predicted ribosome binding site present downstream of the *E.coli* promoter elements, but as above there were no translation initiation codons at the requisite distance.

pRLA5.240 (Figure 53)

Only one strand was sequenced, and only the 3' terminus was obtained. The insert was 163bp long and had no internal *Sau*3A sites. It had a G+C content of 40%. This sequence had no homology with the nif concensus when up to 4 mismatches were allowed. There was a perfect -10 region and the corresponding -35 sequence at a spacing of 18bp had 3 matches to the *E.coli* concensus including one of the most highly conserved bases. Another potential -35, -10 sequence had a spacing of 17bp and 4 out of 6 matches to the *E.coli* concensus in the -35 region including the 3 most highly conserved bases and 3 out of 6 matches in the -10 region including 1 of the highly conserved bases.

pRLA5.273 (Figure 54)

Both strands were sequenced and both termini were obtained. It was 374bp long and had no internal *Sau*3A sites. The sequence had a G+C content of 58%, which is closer to that of the *P.putida* chromosome than the *E.coli* chromosome. There were two fits to the nif concensus sequence including 4 mismatches and the best fit to the *E.coli* concensus had a spacing of 18bp between the hexamers and 3 out of 6 matches to the concensus in the -35 region and 3 out of 6 matches in the -10 region. 2 of the 3 most conserved

bases were present in both the -35 and -10 regions.

6.6.2 GROUP II CLONE

pRLA5.98 (identical to pRLA5.82 and pRLA5.85) (Figure 55)

This was sequenced on both strands but only the 3' terminus was obtained. The insert was 310bp long and had 3 internal *Sau*3A sites and therefore religation may have occurred. It had a G+C content of 61%, identical to that of the *P.putida* chromosome. The best homology to the nif concensus had 4 mismatches. There was no good homology to the *E.coli* concensus, the best fit had 5 out of 6 matches in the -35 region including 2 of the conserved bases and 3 out of 6 matches in the -10 region including 2 of the conserved bases. The spacing between these -35 and -10 hexamers was 16bp. There was a sequence with homology to the predicted ribosome binding site downstream of the potential *E.coli* promoter, but there was no translation initiation codon at the requisite distance.

6.6.3 GROUP III CLONE

pRLA5.172 (Figure 56)

This was sequenced on both strands and both termini were obtained. The insert was 168bp long and had 1 internal *Sau*3A site. It had a G+C content of 56%, higher than the average for the *E.coli* chromosome. This Group III member was expected to exhibit a good -35/-10 promoter sequence, the best such sequence had 2 out of 3 of the most conserved bases present in the -35 region and 2 other mismatches to the *E.coli* concensus. In the -10 region there was only 1 nonhomologous base to the *E.coli* concensus and this was in one of the most highly conserved bases. The spacing between these hexamers was 17bp. There was no homology to the nif concensus sequence allowing up to 4 mismatches.

As far as it is possible to judge there is no obvious difference in the similarity of the best -35, -10 sequences exhibited by the Group I, II or III clones to the *E.coli* concensus, i.e. the Group III clone did not exhibit a sequence more highly homologous to the *E.coli* concensus than the Group I clones. However there is as yet no evidence that any or all the putative *E.coli*-like promoters function *in vivo* either in *E.coli* or in *P.putida*.

The promoter sequences used to generate a concensus nif sequence and also the *xyl*CAB and CPG2 promoters have no more than 3 mismatches to the nif concensus. None of the sequences analysed here had homology better than 4 mismatches to the nif concensus, hence it is reasonable to assume that these sequences will not function as active nif-like promoters *in vivo*. Also, genes which have active nif promoters are positively activated, therefore, in order for a cloned sequence containing a nif-like promoter to be the active promoter *in vivo*, it is necessary that the binding site for the positive activator has also been cloned. This binding site occurs from 100 to 150 bases upstream of the transcription start (Buck *et al* 1986). The cloned sequences were therefore searched for homology to the concensus NifA binding site ($TGTN_4TN_5ACA$) using the "FIND" program. No homology was found for any of the GroupI, II or III sequences allowing up to 1 mismatch to the concensus, except for pRLA5.90 which had a sequence $TGTN_4TN_5ACg$ at position 149 which is 57 bases upstream of the putative nif promoter sequence. Because of the low level of homology to the nif concensus sequence and the close proximity of this sequence to the putative binding site of the positive activator, it is doubtful that these sequences would function *in vivo*.

The G+C content of the cloned sequences varies from 40% (pRLA5.240) to 66% (5' sequence of pRLA5.90). It has been reported that strong *E.coli* promoter regions (Haughn *et al* 1986) and those of bacteriophage T5 (Gentz and Bujard 1985) are particularly high in A+T content, possibly

facilitating unwinding of the DNA during initiation of transcription. Although most of the cloned sequences reported here have a G+C content similar to that of the *P.putida* genome (61%), pRLA5.240 has a particularly low G+C content of only 40%. Such low G+C contents have also been reported for other *Pseudomonas* promoters, for example the *algD* gene promoter has a G+C content of 46% (Deretic *et al* 1987c). Also Schell (1986) found a similar situation in the *nah* and *sal* operons, which as discussed above have *E.coli*-like promoter sequences. Hence this phenomenon of high A+T in promoter regions may be a feature of both *E.coli* and *Pseudomonas* genes. Determination of the G+C content of the regions upstream of the translation initiation codons of other *Pseudomonas* genes contained in the database also revealed that whereas the phosphate regulated haemolysin gene of *P.aeruginosa* has an overall G+C content of 66% (Pritchard and Vasil 1986), the region of 200bp upstream of the translation initiation codon has a G+C content of only 43%. Also, the *clcB* gene involved in the degradation of 3-chlorocatechol has an overall G+C content of about 60% (Frantz and Chakrabarty 1987), but the region of 200bp upstream of the translation initiation codon has a G+C content of only 47%.

The Group I sequences were examined for any strong homologies by comparing them in pairs using the "BESTFIT" program. This revealed 3 different 6 or 7 basepair sequences which were present in each of 2 of the promoter clones. The sequence GCTTGAG was found in pRLA5.240 and pRLA5.273 and a closely homologous sequence GCaTGAG was found in pRLA5.187. The sequence GACTGTG was present in pRLA5.193 and pRLA5.240 and a closely homologous sequence GAtTGTG was present in pRLA5.187. The sequence GGGCAA was found in the 3' sequence of pRLA5.90 at 2 positions (as a direct repeat) and in pRLA5.187 and pRLA5.273. These motifs were then compared against the *Pseudomonas* database. Some homologies were revealed and for some these may lie in the promoter region, but due to the lack of transcription initiation data it was not possible to confirm this. For the

few genes whose transcription starts have been mapped, the homologies were found not to lie in the predicted promoter regions.

To determine any possible homologous regions between the *Pseudomonas* promoter clones and published *Pseudomonas* gene promoters, the cloned *P.putida*-specific insert sequences were compared against the database of promoter regions from *Pseudomonas* genes using the "FASTN" program. This comparison revealed for each Group I sequence a region that was homologous to more than one of the other sequences in the database (Table 20). These sequences were not identical to those found in the comparison of Group I sequences. For the genes whose transcription start points were known, it was found that the TGGCGGNT motif occurred in the *xylCAB* promoter downstream of the transcription start, but in *toxR* TGGCG occurred at -10bp to the transcription start. The sequence GACGNNNA occurred in the phosphate regulated heamolysin gene of *P.aeruginosa* at -45bp to the transcription start.

6.7 DERIVATION OF A PUTATIVE *PSEUDOMONAS* CONCENSUS PROMOTER SEQUENCE

For the determination of a meaningful concensus sequence for *Pseudomonas* promoter sequences it is necessary to omit from the comparison:
(a) those genes which have recognisable *E.coli* or nif sequences which are active *in vivo*
(b) plasmid-encoded genes since their present location on a *Pseudomonas* plasmid does not preclude their origin from a species other than *Pseudomonas*. Also by virtue of being plasmid-encoded their promoters may have deviated from any *Pseudomonas* concensus if the host range of the plasmid is, or has been in its past evolutionary history, greater than the pseudomonads alone.
(c) genes that are positively activated, since evidence from *E.coli* suggests that such genes in *E.coli* have promoter sequences which are not related to the *E.coli* concensus.

TABLE 20

REGIONS OF HOMOLOGY BETWEEN GROUP I SEQUENCES AND *PSEUDOMONAS* DATABASE

1. pRLA5.90	GCCAGGCCCG		
acyI	cCCcGGCCaG		
acyII	GCaAGGCCgG		
2. pRLA.187	CCTGCCGGA		
acyI	CCTGCtGcA		
acyI	CCTGCgGcA		
pilin P1	gCTGCCaaA		
clcD	CCTGCtccA		
3. pRLA5.193	GATCTCG	pRLA5.193	GACGATCA
inaW	GAcCTCG	*pufB*	GACGggCA
acyII	GATCTCG	heamolysin gene	GACGATtA
4. pRLA5.240	TATAATTC		
xylR	TggAATgC		
pilin P11.1	TtTAATTC		
5. pRLA5.273	TGGCGGCT	pRLA5.273	CAGGCCTT
xylCAB	TGGCGGCT	*xylR*	CAGGCCTT
catBC	cGGCGGCT	*vanAB*	CAGGCgTT
pRLA5.273	GTCAACAA	pRLA5.273	TCATCCT
benzene oxidation	GTCAACgA	*pufB*	TgATCCT
M.clara promoter	GTaAACAA	*toxA*	TCATCCT

In addition to the above limitations it is necessary, for the accurate determination of a promoter region, that comparisons are made for the regions upstream of known transcription initiation points. Unfortunately, however, if all these limitations are made for all the published data for pseudomonad gene promoter regions, it leaves too few sequences for comparison. It is indeed a pity that for the many *Pseudomonas* genes and promoter regions that have now been sequenced (and the number is rapidly growing) that the investigators have not carried out the step of mapping the transcript starts. Had this been done it would have allowed for a more informative comparison of *Pseudomonas* promoter regions. In the following comparison of promoters, points (a) and (b) above were adhered to, but regulated genes were included. Also genes from some closely-related species were included.

Nucleotide sequences from both chromosomally-encoded *Pseudomonas* genes and genes from related species, whose transcription initiation start points had been mapped and which had no homology to the *E.coli* or nif concensus sequences were compared in the region from +1 to -40 relative to the transcription start. These were regions from the *algD* promoter of *P.aeruginosa*, the aerolysin gene, *aer*A, of *Aeromonas sobria*, the two promoter regions of *tox*A from *P.aeruginosa*, the two promoter regions of the *arg*F gene of *P.aeruginosa* and the heamolysin gene promoter of *P.aeruginosa* (Table 21).

A putative concensus sequence was derived:

C T N C N_6 G C C N_2 A N_4 A N C N C G C C N_3 T

4 4 5 4 5 4 4 4 4 7 7 6 4 4

The numbers below the bases indicate the number of sequences out of the total of 7 which had homology to each base in the concensus. The invariant "C" occurs at -23 to the transcription start (+1) in *algD*, -14 in *aer*A, -14 in *arg*F mRNA-2, -10 in *tox*A mRNA-1, -11 in *tox*A mRNA-2, -15 in *tox*R and -17 in the haemolysin gene. Table 21 shows the percentage homology of the

TABLE 21

SEQUENCES USED IN DERIVATION OF A PROMOTER CONCENSUS FOR *PSEUDOMONAS*

	PERCENTAGE HOMOLOGY TO CONCENSUS
aerolysin A gene from *Aeromonas sobria*	
attTCCTGCTTGagtCTAAATAACCTCGgCTGATatggtgg	64
exotoxin A gene from *P.aeruginosa*	
mRNA-1 ttccgCTCCCCGCCAGCCTCcCCGCATCCCGCaCCCTaga	86
mRNA-2 cccgCaCCCTAGACGCCCCgCCGCtCtCCGCCGGCTcgcc	71
*tox*R gene from *P.aeruginosa*	
CTGCGTGCGGGCtCCATGCCcGaGCGCCTTGgcgagattt	71
arginine F gene from *P.aeruginosa*	
tCcGCGACATTtCCTTATAAGATCGCGCCTTCccctattt	79
phoshate regulated haemolysin gene from *P.aeruginosa*	
tTAaTCATCTcgaAAcAAGAAGtACGCAGATTgatggaaatc	50
*alg*D gene from *P.aeruginosa*	
cgagcggGaCAAACGGCCGGAACTTcCCTCGCaGAGaaaacatcctatca	64
CONCENSUS CTNCNNNNNGCCNNANNNNANCNCGCCNNNT	

Upper case indicates homology to concensus, lower case no homology

sequences used to determine the concensus to the concensus. This varies from 50 to 86%. The Group I,II and III sequences were examined for homology to this concensus using the "GAP" program. The Group I sequences all exhibited a region with either 57 or 64% homology, the Group II clone had 64% homology and the Group III only 50%. Whether these homologies are significant in terms of their possible location in promoter regions can only be determined by mapping the transcription start points. It is interesting that the Group III clone had the lowest homology with the concensus, only 50%. This group of clones might be expected to have low homology if the concensus sequence formed the promoter of genes specifically expressed in *Pseudomonas*. The promoter regions up to -40 from the transcription start for the TOL mutant promoters which give rise to constitutive expression were compared against this concensus, but in each case there was poor homology, less than 50%. The *Pseudomonas* database was also searched for homology to this putative *Pseudomonas* concensus. Several of the sequences showed a degree of homology greater than 50%, these are shown in Table 22. For the TOL promoter regions where the transcription initiation points are known the homologous region for *xylDEGF* occurs downstream of the transcript start, for *xylCAB* it occurs at -142 and for the constitutive mutant promoter of *xylE* it occurs at -63, -83 and -96 relative to the three transcript starts. For the other genes whose transcript starts are known it occurs at -79 in the P11.1 promoter of *Caulobacter crescentus*, at +60 in the P11 promoter of *C.crescentus*, and at +67 in the aerolysin gene *aerC* of *A.sobria*.

Deretic *et al* (1987c) has identified homologous regions in the promoters of *algD* and *toxA*, two positively regulated genes of *P.aeruginosa*. There was a sequence of CCGGAACTTCCCTCGCAG in *algD* and a highly homologous sequence of CCGCTCCCCGCCAGC in *toxA* of *P.aeruginosa*. Both of these sequences occur around -30. In addition there was a sequence conserved between these two genes of CATCC around -10. The sequence ATCC has also

TABLE 22

HOMOLOGIES OF DATABASE SEQUENCES TO THE PUTATIVE *PSEUDOMONAS* CONCENSUS

	PERCENTAGE HOMOLOGY TO CONCENSUS
*acy*I of a *Pseudomonas* sp. CgCCGCCATCGCCCGgCAGGAGCGCGgCGAGc	71
aerC of *A.sobria* CTTCATCCCGGgtCAgCACCAGCTCGCgTCCg	64
PII of *C.crescentus* CgCCGCACCCaCCGCcGCTCtTCCCGCCGCGc	64
xylCAB of *P.putida* tTGaTGATTTGCtCAAATACAGCCaGCgTGCT	64
xylDLEGF of *P.putida* CTTtGAGGGCaaCTGgATTTATCTCGCCCACg	64
constitutive mutant *xylE* of *P.putida* CcACAGGATTtgCTCATGATACgACtCCACTT	64
28kD flagellin gene promoter of *C.crescentus* aTCtCGGCGTGCgATtTGATcGCACGCCGAAT	64
benzene oxidase gene promoter 1 of *P.putida* aTCgGCAATTGCCTGcCAAGtACCCGCCATCc	64
copper resistance gene promoter of *P.syringae* tTCaAGCTTAcagAAATGTAATCGCGCCGCTT	64
agrA of *P.atlantica* aaTCATACCTGCCATcCCTTACCCCGCaTAGc	64
vanAB of a *Pseudomonas* sp. tTCaCGAAGGcCCGTACTCTAGCCaGCCGCCc	64
PII.I promoter of *C.crescentus* CTGCTAAGGAGCCATcTTGTtTaGgcCtTTCT	57
xylS of *P.putida* tcGaCTTGGCGCCTTtCTACATCACaCCAAGc	57
an *M.clara* gene promoter CaACAAACGCaaCAGgCGAAgGCCgGCCGGAT	57
pilin PAK122-4 of *P.aeruginosa* CTGCCAAATCGagGAAATCCAGCTgtCaAAAa	57
inaW of *P.fluorescens* CTGCATAGACatCGTATGTGAAaAtGtaTTAT	57
catBC of *P.putida* CTGgCAGCCCGaaACAACGGACCTgGCaACAa	57
iaaM of *P.savastanoi* CTAtACGCAAGCCAGtGCGTgGtCgGCCAGCg	57

been found to be present in the promoter region of another positively regulated gene, *cat*BC (Aldrich and Chakrabarty 1988). Although there is no evidence that they have functional significance, these homologies are particularly striking, therefore the Group I, II and III sequences were examined for the presence of similar structures. Although some of the sequences show considerable homology to the sequences centered around -30, in particular pRLA5.273 which has 73% homology to the 15bp sequence in *tox*A, a sequence of ATCC is not present further downstream. It is possible that the conserved elements in the *tox*A and *alg*D promoter regions are not important for *Pseudomonas*-specific gene expression *per se*, but are elements important in positive regulation of these genes.

6.8 SUMMARY OF RESULTS

Fragments of the *P.putida* chromosome were shotgun cloned into a promoter-probe vector and clones which exhibited *Pseudomonas*-specific promoter activity were isolated and sequenced. None of these sequences exhibit functional nif-like promoters. All have sequences with homologies to *E.coli* promoters, but determination of the activity of these *in vivo* requires further investigation. Various comparisons between Group I sequences, and also between Group I sequences and a constructed *Pseudomonas* promoter database revealed some homologies, but due to the lack of information on transcript start points in *Pseudomonas* genes, it was not possible to determine whether these homologies were located in the promoter regions. By comparison of 7 chromosomally-located sequences from genes of *Pseudomonas* and related species which had no homology with *E.coli* or nif-like promoters upstream of the transcription start, a putative concensus sequence was constructed. The significance of this concensus could not be measured by comparing it with the *Pseudomonas* database since homologies could not be located relative to promoter regions due to the lack of transcription initiation data. However the *xyl*CAB operon (which has

a functional nif promoter) exhibited homology at position -149 relative to the transcription start, suggesting that although it may have some significance in regulation the concensus is not involved in promoter recognition, at least for this gene. None of the Group I sequences contained regions with homology to all conserved elements found in the positively regulated gene promoters of *algD* and *toxA*.

In summary, no strong homologies were observed for the putative *Pseudomonas*-specific promoter clones isolated, either by comparison between themselves or against a database of upstream regions of other *Pseudomonas* genes. Further information is required on the transcription initiation points in *E.coli* and *P.putida* for these isolated sequences and also whether they are true constitutive promoters or whether they are positively activated.

CHAPTER 7

SUMMARY

7.1 THE STUDY OF PSEUDOMONAD PROPERTIES

7.2 THE COUPLED TRANSCRIPTION/TRANSLATION SYSTEM FOR *P.PUTIDA*

7.3 THE BARRIER TO EXPRESSION OF *PSEUDOMONAS* GENES IN *E.COLI*

7.4 FUTURE DEVELOPMENTS

7.1 THE STUDY OF PSEUDOMONAD PROPERTIES

The genus *Pseudomonas* exhibits a plethora of unique and potentially exploitable properties and one species, *P.aeruginosa*, is an important human pathogen. For these reasons several pseudomonads are now being intensively studied at the molecular level. However, a major hindrance to the study of the properties elaborated by these species is that of gene expression. Although some *Pseudomonas* genes can be expressed in the well-defined host *E.coli*, there remain many for which expression can only be achieved in this host when the expression is driven by *E.coli* expression signals. Although this can be of use, particularly for the purification of a gene product, it provides no information on the mechanism of regulation of such genes. It is desirable that the increasing knowledge about *Pseudomonas* genes and their products can be put in the context of an understanding of the fundamentals of gene expression in *Pseudomonas*, rather than, as has by necessity been done, relating this information to the better understood mechanism of gene expression in *E.coli*.

7.2 THE COUPLED TRANCRIPTION/TRANSLATION SYSTEM FOR *P.PUTIDA*

Despite the array of sophisticated techniques now available for the genetic manipulation of *E.coli*, these techniques are generally not applicable to the genetic analysis of *Pseudomonas*. This is particularly pertinent to many of the interesting and unusual properties exhibited by the pseudomonads which are often not expressed in *E.coli*. The lack of amenability of *Pseudomonas* genes to manipulation by *E.coli*-based techniques has brought about the derivation of broad host range cloning vectors and cosmids, transposon-donor plasmids, promoter-probe vectors, expression vectors and *Pseudomonas* host strains that can be easily transformed. Although once cloned, *Pseudomonas* genes can often be expressed in an *E.coli* gene expression system when under control of a strong *E.coli* promoter

signal, it would be more desirable to determine the protein products and thus coding capacity of such genes when expressed by their native expression signals and in their native host background. For this purpose the published method for the generation of an *in vitro* coupled transcription/translation system for *E.coli* (Zubay 1973, Pratt *et al* 1981) was adapted for the production of an active coupled transcription/translation system based on *P.putida*. This, like the comparable *E.coli* system, can be used for the analysis of cloned gene products. This *P.putida* system may have applications for the analysis of cloned *Pseudomonas* genes in general, particularly when information is required concerning such genes which are poorly expressed in gene expression systems based on *E.coli*, for example the *xylS* gene from the *P.putida* TOL pathway (Spooner *et al* 1987). In addition to the expression of genes encoded on supercoiled plasmids molecules, linear fragments can be used to direct protein synthesis in the *P.putida* coupled transcription/translation system as as been demonstrated for the comparable *E.coli* system (Pratt *et al* 1981). This method has been shown to be useful for the localisation of gene products to coding sequences. This *Pseudomonas* gene expression system is thus a valuable addition to the tools for genetic manipulation of *Pseudomonas* properties.

There is an interesting group of plasmids, those of incompatibility groups P and Q, which have an exceptionally wide host range including both *Pseudomonas* and *E.coli*. The nature of broad host range is not yet understood. It is possible that it may require the expression of different genes in different hosts. Alternatively the expression of identical genes may be driven by the use of different promoter signals or by the use of generalised promoter sequences active in many hosts. Evidence from RK2 (Guiney 1982) suggests that host range is associated with the replication proteins expressed by such broad host range plasmids.

The coupled transcription/translation system derived from *P.putida* was

used to examine the polypeptide products directed by the broad host range plasmid RSF1010/R300B and to compare them with the products synthesised in an *E.coli* coupled transcription/translation system. The results of the comparison revealed that although most of the products directed by R300B were identical in both host systems, some products were uniquely synthesised in one system or the other. It is possible that some of these differences could be due to proteolytic activity rather than the use of different expression signals in the two host systems. However of most significance was the synthesis of a large (73kD) polypeptide in the *P.putida* coupled transcription/translation system alone which was identified as the product of the *repB** gene involved in mobilisation. The promoter of this gene also transcribes an overlapping open reading frame for a gene product, RepB, that is synthesised in both the *E.coli* and *P.putida* coupled transcription/translation systems. Hence these data suggest that the differential expression of the 73kD polypeptide cannot be due to ineffectual recognition of the promoter by the *E.coli* RNA polymerase but it may be due to poor recognition of the ribosome binding site in *E.coli* or the polypeptide may be unstable in *E.coli* compared with *P.putida*. The proposition of differential recognition of ribosome binding sites does not hold up well against experimental data. The ribosome binding sites of *E.coli* and *P.aeruginosa* as determined by Shine and Dalgarno (1975) are highly similar and the ribosome binding site of *repB** has homology to both. Also there is little evidence for translational specificity by ribosomes of different Gram negative organisms, although this does occur between Gram negative and Gram positive organisms (Lodish 1970, Stallcup and Rabinowitz 1973) and for an *E.coli* plasmid-encoded gene (Goldfarb *et al* 1982).

E.coli promoter elements have been identified for all of the ORF's of RSF1010 which give rise to functional protein products (E.Scherzinger, Personal Communication). However this does not preclude that there may be other regions specifically used as promoter recognition sequences in other

host environments. Although the results from the *in vitro* coupled transcription/translation system experiments obviously cannot enlighten us greatly on the nature of expression of broad host range genes *in vivo*, it appears that there are differences in the expression of genes directed by the broad host range plasmid, RSF1010, in different host backgrounds *in vitro*. The data also suggest that alternative regions of gene products may function in different hosts. Similar situations have been observed with the *trf*A gene of RK2 and the *sog* gene of the ColI plasmid.

7.3 THE BARRIER TO EXPRESSION OF *PSEUDOMONAS* GENES IN *E.COLI*

The nature of the barrier to expression exhibited by some *Pseudomonas* genes in *E.coli* has not been directly investigated. Although a multitude of possibilities exist in terms of differing transcriptional/translational machinery between the two species, the strongest evidence which pinpoints where this difference lies comes from the analysis of promoter signals in *Pseudomonas* in terms of nucleotide sequence and determinations of mRNA start points and levels of mRNA production. These results have suggested that there may be a barrier to expression at the level of recognition of the *Pseudomonas* promoter signal in *E.coli*, for some *Pseudomonas* genes at least.

Promoters in *E.coli* were initially easily recognised as hexameric nucleotide sequences, but more recent work has shown that other modified sequences can have activity higher than the *E.coli* concensus indicating that promoter recognition by RNA polymerase is not the total determinant of promoter activity. In *Pseudomonas*, studies of promoter sequence revealed that some genes have *E.coli*-like or nif-like promoters, but other genes had no homology to these. Since this latter type of gene is generally poorly expressed in *E.coli* this created the possibility that such genes have *Pseudomonas*-specific promoter sequences presumably recognised by an alternative RNA polymerase holoenzyme. However comparisons of such promoter

sequences (Mermod *et al* 1984, Inouye *et al* 1986) have so far failed to reveal possible recognition sequences. Such a class of promoters are, therefore, considered to be specific to *Pseudomonas* and hence it could be anticipated that the constitutive "house-keeping" genes expressed at significantly high levels throughout the growth cycle of an organism, may have promoters of this type.

The attempts at derivation of a concensus sequence for these promoters from the published data on cloned *Pseudomonas* genes has been hampered in several ways:

(a) the relatively small number of genes for which transcription initiation data has been obtained

(b) many of the data are from plasmid-encoded genes and therefore may possibly not have similarity to *Pseudomonas* chromosomal genes

(c) most of the gene systems investigated are highly regulated and therefore the promoters of these may also differ compared with unregulated gene promoters. Previously published concensus sequences for *Pseudomonas* promoters have proved to be of little significance in the light of data which show that *Pseudomonas* promoters are heterogenous, some having *E.coli* promoters active in *Pseudomonas* and others having nif-like promoters active in *Pseudomonas.*

Therefore in order to derive a set of promoter sequences which were more directly comparable in terms of being chromosomally-encoded, constitutively-expressed and with activity specific to *Pseudomonas* this work was concerned with the construction of a broad host range promoter-probe vector which allowed for both the shotgun cloning of promoters from *P.putida* and the assay of their activity in both *E.coli* and *P.putida.* In this way a subset of promoters was isolated which exhibited activity more specific to *Pseudomonas.* All of these promoter clones exhibited some activity in *E.coli.* The proportion of this type of promoter obtained, compared with all clones having any promoter activity, was only

about 11%. This class of clones, designated as Group I, in addition to representing those promoters which give rise to constitutive expression in *P.putida*, are also expected to include promoters which require a positive activator not synthesised by *E.coli*. Hence at the level of stringency set in this work for the determination of *Pseudomonas*-specific promoters, it appears that most *P.putida* genes have promoters which can be recognised to some degree by both *P.putida* and *E.coli*. Thus the data imply that the hypothesis that most or all *Pseudomonas* "house-keeping" genes have a promoter specificity optimised for expression in *Pseudomonas* rather than *E.coli* may be incorrect.

Since mRNA initiation data were not obtained for the *Pseudomonas*-specific promoters isolated in this work, a comparison was made of published chromosomally-encoded promoter sequences from +1 to -40 relative to the trancription start from *Pseudomonas* genes and an *Aeromonas* gene. There were no strong homologies, but a tentative concensus sequence was obtained. it is possible that one reason for the lack of obvious homologies is that positively regulated genes were included in the analysis and they might be expected to have an alternative structure to that of unregulated promoters, as has been found for positively regulated *E.coli* genes. It remains to be seen whether the putative concensus is found in similar regions of other *Pseudomonas* genes or indeed those promoters isolated in this work. If a true *Pseudomonas*-specific promoter really exists then it is possible that elements involved in the promoter recognition sequence may be more subtle than those for an *E.coli* promoter, in view of the difficulties in identification of conserved sequences.

7.4 FUTURE DEVELOPMENTS

In order to derive any useful meaning in terms of the sequences active in *E.coli* and *P.putida* as the promoters on the *Pseudomonas*-specific clones isolated in the course of this work, it is necessary that transcript

mapping analysis is carried out in both species. Having obtained this data the promoter regions could be aligned relative to their transcription starts and homologies determined between the sequences more accurately than can be achieved without this additional data. These sequences could also be compared with the other published *Pseudomonas* promoter sequences for which transcript mapping data is available and a more accurate determination of a possible *Pseudomonas* concensus could be carried out. The nature of the genes whose promoters these sequences represent is of interest, particularly in terms of whether they are constitutively expressed "house-keeping" genes as predicted, or genes for catabolic pathways. One piece of information that is required is whether the isolated *Pseudomonas*-specific promoter clones are from constitutively expressed genes, or whether they are from genes under positive control. If the latter were correct then it is possible that if the positive activator was provided then the promoter may be as active in *E.coli* as in *P.putida*.

It would be interesting to determine what sigma factor of RNA polymerase is involved in the recognition of these sequences, i.e. does *Pseudomonas* possess a novel sigma factor that specifically recognises a set of promoters not recognised by *E.coli* RNA polymerase. A comparison of σ factors from both *E.coli* and *Pseudomonas* could reveal the potential of these two hosts for the recognition of different promoter signals. Since the discovery of σ^{70}, three other σ subunits have been identified in *E.coli* and therefore a similar potential for the synthesis of alternate σ factors exists for *Pseudomonas*. It is even conceivable that different species of *Pseudomonas* possess a different range of such subunits, particularly in view of their diverse habitats and ubiquitous existence in a wide range of nutrient impoverished environments. It is known that the alternate σ subunits of *B.subtilis* are involved in the response to environmental stress and in the regulation of developmental and nonessential genes such as extracellular enzyme genes and sporulation genes. It is conceivable that

the expression of genes when growth has ceased is controlled by alternative σ factors, for example the synthesis of secondary metabolites.

The genus *Pseudomonas* constitutes a wide range of species which in turn inhabit a large array of ecological niches. Mainly from a study of *P.aeruginosa* and *P.putida*, which are closely related relative to the genus as a whole, it is apparent that gene expression in both is broadly similar to that of *E.coli* but incorporates some fine tuning which possibly reflects the varied and unusual genetic properties of the genus. Hopefully in future years we will gain some insight into the nature of these differences in terms of the existence of promoter sequences and transcription termination sequences specific to *Pseudomonas*. A survey of the more diverse members of the genus may reveal other features important in the expression and regulation of *Pseudomonas* genes.

REFERENCES

Aldrich TL, Chakrabarty AM (1988) J Bacteriol 170;1297-1304

Allan B, Kropinski AM (1987) Biochem Cell Biol 65;776-782

Allan B, Linseman M, MacDonald LA, Lam JS, Kropinski AM (1988) J Bacteriol 170;3668-3674

Bagdasarian M, Lurz R, Rueckert B, Franklin FCH, Bagdasarian MM, Frey J, Timmis KN (1981) Gene 16;237-247

Bagdasarian M, Bagdasarian MM, Lurz R, Nordheim A, Frey J, Timmis KN (1982) In:Bacterial Drug Resistance (Mitsuhashi S, Ed) Japan Scientific Society Press, Tokyo p.183-197

Bagdasarian M, Timmis KN (1982) Curr Top Microbiol Immunol 96;47-67

Bagdasdarian MM, Amann E, Lurz R, Ruckert B, Bagdasarian M (1983) Gene 26;273-282

Bagdasarian MM, Scholz P, Frey J, Bagdasarian M (1986) In:Banbury Report 24: Antibiotic Resistance Genes: Ecology, Transfer and Expression. Cold Spring Harbor, New York p.209-233

Barth PT, Grinter NJ (1974) J.Bacteriol 120;618-630

Barth PT, Tobin L, Sharpe GS (1981) In:Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids (Levy SB, Clowes RC, Koenig EL, Eds) Plenum Press New York p.439-448

Bauer CE, Young DA, Marrs BL (1988) J Biol Chem 263;4820-4827

Belas R (1989) J Bacteriol 171;602-605

Betz JL, Clarke PH (1972) J Gen Microbiol 73;161-174

Beynon J, Cannon M, Buchanon-Wollaston V, Cannon F (1983) Cell 34;665-671

Birkmann A, Sawers RG, Bock A, (1987) Mol Gen Genet 210;535-542

Birnboim HC, Doly J (1979) Nucl Acids Res 7;1513-1523

De Boer HA, Comstock LJ, Vasser M (1983) Proc Natl Acad Sci 80;21-25

Boulnois GJ, Wilkins BM, Lanka E (1982) Nucl Acids Res 10;855-869

Boulnois GJ, Varley JM, Sharpe GS, Franklin FCH (1985) Mol Gen Genet 200;65-67

Brammar WJ, Clarke PH, Skinner AJ (1967) J Gen Microbiol 47;87-102

Brammar WJ, Charles IG, Matfield M, Cheng-Pin L, Drew RE, Clarke PH (1987) FEBS Letts 215;291-294

Brosius J, Holy A (1984) Proc Natl Acad Sci 81;6929-6933

Brown JE, Brown PR, Clarke PH (1969) J Gen Microbiol 57;273-285

Brown NL, Ford SJ, Pridmore RD, Fritzinger DC (1983) Biochemistry 22;4089-4095

Brown PR, Clarke PH (1972) J Gen Microbiol 70;287-298

Brunel F, Davison J (1988) J Bacteriol 170;4924-4930

Brunner M, Bujard H (1987) EMBO J 6;3139-3144

Buchanan-Wollaston V, Passiatore JE, Cannon F (1987) Nature 328;172-175

Buck M (1986) J Bacteriol 166;545-551

Buck M, Miller S, Drummond M, Dixon R (1986) Nature 320;374-378

Buckel P, Zehelein E (1981) Gene 16;149-159

Casadaban MJ, Cohen SN (1980) J Mol Biol 138;179-207

Chak K-F, James R (1985) Nucl Acids Res 13;2519-2531

Chakrabarty AM, Gunsalus CF, Gunsalus IC (1968) Proc Natl Acad Sci 60;168-185

Chatterjee DK, Chatterjee P (1987) J Bacteriol 169;2962-2966

Chen H-Z, Zubay G (1983) J Bacteriol 154;650-655

Clarke PH, Ornston LN (1975) In:Genetics and Biochemistry of *Pseudomonas* (Clarke PH, Richmond MR, Eds) J.Wiley and Sons, Bristol, UK. p.191-340

Collins J (1979) Gene 6;29-42

Consevage MW, Porter RD, Phillips AT (1985) J Bacteriol 162;138-146

Conway T, Osman YA, Ingram LO (1987a) J Bacteriol 169;2327-2335

Conway T, Sewell GW, Ingram LO (1987b) J Bacteriol 169;5653-5662

Cowing DW, Bardwell JCA, Craig EA, Woolford C, Hendrix RW, Gross CA (1985) Proc Natl Acad Sci 82;2679-2683

Davison J, Heusterspreute M, Chevalier M, Thi VH, Brunel F (1986) Gene 51;275-280

Derbyshire KM, Hatfull G, Willets N (1987) Mol Gen Genet 206;161-168

Deretic V, Gill JF, Chakrabarty AM (1987a) Biotechnology 5;469-477

Deretic V, Gill JF, Chakrabarty AM (1987b) J Bacteriol 169;351-358

Deretic V, Gill JF, Chakrabarty AM (1987c) Nucl Acids Res 15;4567-4581

Deretic V, Chandrasekharappa S, Gill JF, Chatterjee DK, Chakrabarty AM (1987d) Gene 57;61-72

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

Ditta G, Stanfield S, Corbin D, Helinski DR (1980) Proc Natl Acad Sci 77;7347-7351

Dixon R (1986) Mol Gen Genet 203;129-136

Doi RH, Wang L-F (1986) Microbiol Revs 50;227-243

Dougan G, Sherratt D (1977) Mol Gen Genet 151;151-160

Duchene M, Schweizer A, Lottspeich F, Krauss G, Marget M, Vogel K, von Specht B-U, Domdey H (1988) J Bacteriol 170;155-162

Dunn NW, Gunsalus IC (1973) J Bacteriol 114;974-979

Durland RH, Helinski DR (1987) Plasmid 18;164-169

Dusha I, Shroder J, Putnoky P, Banfalvi Z, Kondoros A (1986) Eur J Bioch 160;69-75

Eggink G, Lageveen RL, Altenburg B, Witholt B (1987) J Biol Chem 262;17712-17718

Ensley BD, Ratz BJ, Osslund TD, Simon MJ, Wackett LP, Gibson DT (1983) Science 222;167-169

Farin F, Clarke PH (1978) J Bacteriol 135;379-392

Figurski DH, Meyer RJ, Helinski DR (1979) J Mol Biol 133;295-318

Figurski DH, Pohlman RF, Bechhofer DH, Prince AS, Kelton CA (1982) Proc Natl Acad Sci 79;1935-1939

Franklin FCH, Bagdasarian M, Bagdasarian MM, Timmis KN (1981) Proc Natl Acad Sci 78;7458-7462

Franklin FCH, Lehrbach PR, Lurz R, Ruueckert B, Bagdasarian M, Timmis KN (1983) J Bacteriol 154;676-685

Franklin FCH (1985) In:DNA Cloning Vol I, A Practical Approach (DM Glover, Ed) RL Press Ltd. Oxford, UK p.165-184

Frantz B, Chakrabarty AM (1987) Proc Natl Acad Sci 84;4460-4464

Frantz B, Ngai K-L, Chatterjee DK, Ornston LN, Chakrabarty AM (1987) J Bacteriol 169;704-709

Frey J, Bagdasarian M, Feiss D, Franklin FCH, Deshusses J (1983) Gene 24;299-308

Fujita N, Ishihama A, Nagasawa Y, Ueda S (1987) Mol Gen Genet 210;5-9

Furukawa K, Arimura N, Miyakazi T (1987) J Bacteriol 169;427-429

Gentz R, Bujard H (1985) J Bacteriol 164;70-77

Gilbert W (1976) In:RNA Polymerase (Losick R and Chamberlin MJ, Eds) Cold Spring Harbour, New York p.193-205

Gill PR, Agabian N (1983) J Biol Chem 258;7395-7401

Gillman MZ, Wiggs JL, Chamberlin MJ (1981) Nucl Acids Res 9;5991-6000

Goldfarb DS, Rodriguez RL, Doi RH (1982) Proc Natl Acad Sci 79;5886-5890

Govan JRW, Harris GS (1986) Microbiol Sci 3;302-308

Gragerov AI, Chenchik AA, Aivasashvilli VA, Beabealashvilli RSH, Nikiforov VG (1984) Mol Gen Genet 195;511-515

Grant CCR, Vasil ML (1986) J Bacteriol 168;1112-1119

Gray GL, Smith DH, Baldridge JS, Harkins RN, Vasil ML, Chen EY, Heyneker HL (1984) Proc Natl Acad Sci 81;2645-2649

Greener A, Helinski DR (1987) Poster at EMBO Workshop on "Promiscuous plasmids of Gram negative bacteria" Birmingham University, UK.

Grosjean H, Fiers W (1982) Gene 18;199-209

Gross CA, Cowing DW, Erikson JW, Grossman AD, Straus D, Walter W, Zhou Y-N (1987) In:RNA Polymerase and the regulation of transcription. (WS Reznikoff *et al*, Eds) Elsevier Science Publishing Co. Inc. p.151-157

Guerry P, Van Embden J, Falkow S (1974) J Bacteriol 117;619-630

Guiney DG (1982) J Mol Biol 162;699-703

Gunsalus IC, Gunsalus CF, Chakrabarty AM, Sikes S, Crawford IP (1968) Genetics 60;419-435

Haas D, Holloway BW, Schambock A, Leisinger T (1977) Mol Gen Genet 154;7-22

Haas D, Holloway BW (1978) Mol Gen Genet 158;229-237

Hahn V (1986) FEMS Letts 36;293-297

Haring V, Scholz P, Scherzinger E, Frey J, Derbyshire K, Hatfull G, Willetts NS, Bagdasarian M (1985) Proc Natl Acad Sci 82;6090-6094

Harley CB, Reynolds RP (1987) Nucl Acids Res 15;2343-2361

Haughn GW, Wessler SR, Gemmill RM, Calvo JM (1986) J Bacteriol 166;1113-1117

Helmann JD, Chamberlin MJ (1987) Proc Natl Acad Sci 84;6422-6424

Higgins CF, Ames GF-L, Barnes WM, Clement JM, Hofnung M (1982) Nature 298;760-762

Hirschman J, Wong PK, Sei K, Keener J, Kustu S (1985) Proc Natl Acad Sci 82;7525-7529

Holm L (1986) Nucl Acids Res 14;3075-3087

Holloway BW (1969) Bacteriol Revs 33;419-443

Holloway BW, Krishnapillai V, Morgan AF (1979) Microbiol Revs 43;73-102

Holloway BW, Morgan AF (1986) Ann Rev Microbiol 40;79-105

Hunt TP, Magasanik B (1985) Proc Natl Acad Sci 82;8453-8457

Husslein V, Huhle B, Jarchau T, Lurz R, Goebel W, Chakraborty T (1988) Molec Microbiol 2;507-517

Inouye S, Nakazawa A, Nakazawa T (1981) J Bacteriol 148;413-418

Inouye S, Nakazawa A, Nakazawa T (1983) J Bacteriol 155;1192-1199

Inouye S, Yousuke E, Nakazawa A, Nakazawa T (1984a) Proc Natl Acad Sci 81;1688-1691

Inouye S, Nakazawa A, Nakazawa T (1984b) Gene 29;323-330

Inouye S, Nakazawa A, Nakazawa T (1985) J Bacteriol 163;863-869

Inouye S, Asai Y, Nakazawa A, Nakazawa T (1986) J Bacteriol 166;739-745

Inouye S, Nakazawa A, Nakazawa T (1987a) Proc Natl Acad Sci 84;5182-5186

Inouye S, Nakazawa A, Nakazawa T (1987b) J Bacteriol 169;3587-3592

Irie S, Shirai K, Doi S, Yorifuji T (1987a) Agric Biol Chem 51;1489-1493

Irie S, Doi S, Yorifuji T, Takagi M, Yano K (1987b) J Bacteriol 169;5174-5179

Ishihama A, Kajitani M, Enami M, Nagasawa H, Fukuda R (1983) In:Microbiology 1983 (Schlessinger D, Ed) American Society for Microbiology, Washington DC p.4-6

Itoh Y, Haas D (1985) Gene 36;27-36

Itoh Y, Soldati L, Stalon V, Falmagne P, Terawaki Y, Leisinger T, Haas D (1988) J Bacteriol 170;2725-2734

Jackson M, Pratt JM, Holland IB (1983) FEBS 163;221-224

Jeenes DJ, Soldati L, Baur H, Watson JM, Mercenier A, Reimmann C, Leisinger T, Haas D (1986) Mol Gen Genet 203;421-429

Johnson JC, DeBacker M, Boezi JA (1971) J Biol Chem 246;1222-1232

Johnson K, Parker ML, Lory S (1986) J Biol Chem 261;15703-15708

Johnson WC, Moran CP Jnr, Losick R (1983) Nature 302;800-804

Kahn M, Kolter R, Thomas CM, Figurski DH, Meyer R, Remaut E, Helinski DR (1979) Methods Enzymol 68;268-280

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

Kartasova T, Huisman H, Schilperoort R (1981) Nucl Acids Res 9;6763-6772

Konyecsni WM, Deretic V (1988) Gene 74;375-386

Kurkela S, Lehvaslaiho H, Palva T, Teeri TH (1988) Gene 73;355-362

Laemmli UK (1970) Nature 227;680-685

Leemans J, Langenakens J, De Greve H, Deblaere R, Van Montagu M, Schell J (1982) Gene 19;361-364

Leidigh BJ, Wheelis ML (1973) J Molec Evolution 2;235-242

Leventhal JM, Chambliss GH (1979) Biochim Biophys Acta 564;162-171

Lodish HF (1970) Nature 226;705-707

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

McClure WR (1985) Ann Rev Biochem 54;171-204

McLaughlin JR, Murray CL, Rabinowitz JC (1981) J Biol Chem 256;11273-11282

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular Cloning: A Laboratory Manual Cold Spring Harbor, New York

Martin C, Cami B, Borne F, Jeenes DJ, Haas D, Patte J-C (1986) Mol Gen Genet 203;430-434

Marugg JD, Nielander HB, Horrevoets AJG, Van Megen I, Van Genderen I, Weisbeek PJ (1988) J Bacteriol 170;1812-1819

Matsuda A, Toma K, Komatsuk K-I (1987) J Bacteriol 169;5821-5826

Matthaei JH, Nirenberg MW (1961) Proc Natl Acad Sci 47;1580-1588

Meagher RB, Tait RC, Betlach M, Bayer HW (1977) Cell 10;521-536

Mellano MA, Cooksey DA (1988) J Bacteriol 170;2879-2883

Mergeay M, Boynen A, Legrain C, Glansdorff N (1978) J Bacteriol 136;1187-1188

Mermod N, Lehrbach PR, Reineke W, Timmis KN (1984) EMBO J 3;2461-2466

Mermod N, Ramos JL, Lehrbach PR, Timmis KN (1986) J Bacteriol 167;447-454

Metzler T, Marquardt R, Prave P, Winnacker E-L (1988) Mol Gen Genet 211;210-214

Meyer R, Boch G, Shapiro J (1979) Mol Gen Genet 171;7-13

Minton NP, Atkinson T, Sherwood RF (1983) J Bacteriol 156;1222-1227

Minton NP, Clarke LE (1985) J Mol Appl Genet 3;26-35

Mullin D, Minnich S, Chen L-S, Newton A (1987) J Mol Biol 195;939-943

Muralikrishna P, Suryanarayana T (1987) Eur J Biochem 167;299-305

Murray K, Duggleby CJ, Sala-Trepat JM, Williams PA (1972) Eur J Biochem 28;301-310

Nakai C, Kagamiyama H, Nozaki M, Nakazawa T, Inouye S, Ebina Y, Nakazawa A (1983) J Biol Chem 258;2923-2928

Neidhart FC, Van Bogelen RA, Vaughn V (1984) Ann Rev Genet 18;295-329

Newbury SF, Smith NH, Robinson EC, Hiles ID, Higgins CF (1987a) Cell 48;297-310

Newbury SF, Smith NH, Higgins CF (1987b) Cell 51;1131-1143

Nirenberg MW, Matthaei JH (1961) Proc Natl Acad Sci 47;1588-1602

Olsen RH, Shipley P (1973) J Bacteriol 113;772-780

Osbourn AE, Barber CE, Daniels MJ (1987) EMBO J 6;23-28

Palleroni NJ (1975) In:Genetics and Biochemistry of *Pseudomonas* (Clarke PH, Richmond MR, Eds) J.Wiley and Sons Bristol, UK. p.1-36

Pasloske BL, Sastry PA, Finlay BB, Paranchych W (1988) 170;3738-3741

Persson C, Nordstrom K (1986) Mol Gen Genet 203;189-192

Pinkney M, Theophilus BDM, Warne SR, Tacon WCA, Thomas CM (1987) Plasmid 17;222-232

Poindexter JS (1981) Microbiol Revs 45;123-179

Postgate JR, Kent HM (1987) J Gen Microbiol 133;2563-2566

Powell GK, Morris RO (1986) Nucl Acids Res 14;2555-2565

Pratt JM, Boulnois GJ, Darby V, Orr E, Wahle E, Holland IB (1981) Nucl Acids Res 9;4459-4474

Pritchard AE, Vasil ML (1986) J Bacteriol 167;291-298

Ramos JL, Gonzalez-Carrero M, Timmis KN (1988) FEBS Letts 226;241-246

Reeve JN, Mendelson NH, Coyne SI, Hallock LL, Cole RM (1973) J Bacteriol 114;860-873

Riordan ML (1987) Biotechnology 5;442-448

Rubens C, Heffron F, Falkow S (1976) J Bacteriol 128;425-434

Russell DR, Bennett GN (1982) Gene 20;231-243

Sancar A, Hack AM, Rupp WD (1979) J Bacteriol 137;692-693

Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci 74;5463-5467

Schell MA (1983) J Bacteriol 153;822-829

Schell MA (1986) Proc Natl Acad Sci 83;369-373

Scherzinger E, Bagdasarian MM, Scholz P, Lurz R, Ruckert B, Bagdasarian M (1984) Proc Natl Acad Sci 81;654-658

Schneider K, Beck CF (1986) Gene 42;37-48

Scholz P, Haring V, Scherzinger E, Lurz R, Bagdasarian MM, Schuster H, Bagdasarian M (1985) In:Plasmids in Bacteria (Helinski DR *et al*, Eds) Plenum Press New York p.243-259

Shanley MS, Neidle EL, Parales RE, Ornston LN (1986) J Bacteriol 165;557-563

Sharpe GS (1984) Gene 29;93-102

Shen B-F, Tai PC, Pritchard AE, Vasil ML (1987) J Bacteriol 169;4602-4607

Shimotsu H, Kawamura F, Kobayashi Y, Saito H (1983) Proc Natl Acad Sci 80;658-662

Shine J, Dalgarno L (1974) Proc Natl Acad Sci 71;1342-1346

Shine J, Dalgarno L, (1975) Nature 254;34-38

Shingler V, Thomas CM (1984) J Mol Biol 175;229-249

Shinomiya T, Shiga S, Kageyama M (1983) Mol Gen Genet 189;375-381

Skelly S, Coleman T, Fu C-F, Brot N, Weissbach H (1987) Proc Natl Acad Sci 84;8365-8369

Smith CA, Thomas CM (1983) Mol Gen Genet 190;245-254

Smith CA, Shingler V, Thomas CM (1984) Nucl Acids Res 12;3619-3630

Soldati L, Jeenes DJ, Haas D (1987) FEMS Letts 42;163-167

Southern (1975) J Mol Biol 98;503-517

Spooner RA, Lindsay K, Franklin FCH (1986) J Gen Microbiol 132;1347-1358

Spooner RA, Bagdasarian M, Franklin FCH (1987) J Bacteriol 169;3581-3586

Stallcup MR, Rabinowitz JC (1973) J Biol Chem 248;3209-3215

Stanier RY, Palleroni NJ, Doudoroff M (1966) J Gen Microbiol 43;159-271

Stern MJ, Ames GF-L, Smith NH, Robinson EC, Higgins CF (1984) Cell 37;1015-1026

Straus DB, Walter WA, Gross CA (1987) Nature 329;348-351

Tait RC, Close TJ, Lundquist RC, Hagiya M, Rodriguez RL, Kado CI (1983a) Biotechnology 1;269-275

Tait RC, Kado CI, Rodriguez RL (1983b) Mol Gen Genet 192;32-38

Takanami M, Sugimoto K, Sugisaki H, Okamoto T (1976) Nature 260;297-302

Thomas CM, Meyer R, Helinski DR (1980) J Bacteriol 141;213-222

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

Twigg AJ, Sherratt DS (1980) Nature 283;216-218

Vann WF, Silver RP, Abeijon C, Chang K, Aaronson W, Sutton A, Finn CW, Lindner W, Kotsatos M (1987) J Biol Chem 262;17556-17562

Viebrock A, Zumft WG (1988) J Bacteriol 170;4658-4668

Vieira J, Messing J (1982) Gene 19;259-268

Walker JA, Allen RL, Falmagne P, Johnson MK, Boulnois GJ (1987) Infect Immun 55;1184-1189

Warren GJ, Corotto L, Wolber P (1986) Nucl Acids Res 14;8047-8060

Warren GJ, Lindemann J, Suslow TV, Green RL (1987) ACS Symposium 334;215-227

West RW Jnr, Neve RL, Rodriguez RL (1979) Gene 7;271-288

Wetekam W, Staack K, Ehring R (1971) Mol Gen Genet 112;14-27

Willetts N, Crowther C (1981) Genet Res 37;311-316

Williams PA, Murray K (1974) J Bacteriol 120;416-423

Williams PA, Worsey MJ (1976) J Bacteriol 125;818-828

Windass JD, Worsey MJ, Pioli EM, Pioli D, Barth PT, Atherton KT, Dart EC, Byron D, Powell K, Senior PJ (1980) Nature 287;396-401

Worsey MJ, Franklin FCH, Williams PA (1978) J Bacteriol 134;757-764

Wozniak DJ, Cram DC, Daniels CJ, Galloway DR (1987) Nucl Acids Res 15;2123-2135

Yamada T, Palm CJ, Brooks B, Kosuge T (1985) Proc Natl Acad Sci 82;6522-6526

Yang H-L, Heller K, Gellert M, Zubay G (1979) Proc Natl Acad Sci 76;3304-3308

Yang H-L, Ivashkiv L, Chen H-Z, Zubay G, Cashel M (1980) Proc Natl Acad Sci 77;7029-7033

Yanisch-Perron C, Vieira J, Messing J (1985) Gene 33;103-119

Yen K-M, Gunsalus IC (1985) J Bacteriol 162;1008-1013

Zimmer SG, Millette RL (1975) Biochemistry 14;290-299

Zubay G (1973) Ann Rev Gens 7;267-287

ACKNOWLEDGEMENTS

I would like to thank Dr GJ Boulnois for his help and interest during the course of this work and for his meticulous reading of this thesis. Also Dr GS Sharpe for his enthusiasm, provision of strains and the hospitality at Runcorn.

I am indebted to Dr IS Roberts for advice, on scientific matters both practical and otherwise and to Dr JA Walker for communication of his practical expertise.

Thanks also go to Professor M Bagdasarian and Dr FCH Franklin for the generous provision of strains and to Dr E Scherzinger for communication of unpublished data. Also to J. Keyte for synthesis of oligonucleotides and the staff of CPU for photographic work.

I would like to thank the SERC and ICI plc for funding this work.

And finally I acknowledge the many lives sacrificed by those generally amenable bugs, *Pseudomonas*.

APPENDIX

C230 ACTIVITIES OF *P.PUTIDA* HARBORING SHOTGUN CLONES

CLONE	C230 ACTIVITY (OD_{375nm} $min^{-1}mg^{-1}$protein)	LEVEL OVER VECTOR pRLA5
pRLA5.19	49.94	14
pRLA5.20	40.31	11
pRLA5.21	34.93	10
pRLA5.22	11.80	3.3
pRLA5.23	47.72	14
pRLA5.24	35.62	10
pRLA5.25	12.05	3.4
pRLA5.26	10.37	2.9
pRLA5.27	8.85	2.5
pRLA5.28	3.93	1.6
pRLA5.29	5.25	2.2
pRLA5.30	5.41	2.2
pRLA5.31	4.25	1.8
pRLA5.32	3.88	1.6
pRLA5.33	6.09	2.5
pRLA5.34	110.08	46
pRLA5.35	3.86	2.8
pRLA5.37	5.03	2.1
pRLA5.38	5.75	2.4
pRLA5.39	9.26	2.6
pRLA5.40	4.32	1.2
pRLA5.41	3.02	0.9
pRLA5.42	4.29	1.2
pRLA5.43	3.47	1.0
pRLA5.44	2.03	0.6
pRLA5.45	2.88	0.8
pRLA5.46	2.26	0.6
pRLA4.47	3.43	1.4
pRLA5.48	1.87	0.5
pRLA5.49	2.82	1.1
pRLA5.50	4.54	1.8
pRLA5.51	3.92	1.5
pRLA5.52	3.79	1.5
pRLA5.53	3.34	1.3
pRLA5.54	23.64	9.3
pRLA5.55	37.97	15
pRLA5.56	8.99	3.5
pRLA5.57	9.98	4.0
pRLA5.58	23.67	9.3
pRLA5.59	40.80	12
pRLA5.60	9.98	2.9
pRLA5.61	3.68	1.1
pRLA5.62	3.41	1.0
pRLA5.63	5.34	1.5
pRLA5.64	50.8	14
pRLA5.65	135.2	38
pRLA5.66	165.2	46
pRLA5.67	25.4	7.0
pRLA5.68	9.31	4.2
pRLA5.69	12.92	6.0
pRLA5.70	8.36	3.8
pRLA5.71	51.19	23
pRLA5.72	137.65	20
pRLA5.73	137.32	20

pRLA5.74	5.46	0.8
pRLA5.75	107.88	16
pRLA5.76	5.34	0.8
pRLA5.77	5.05	0.7
pRLA5.78	13.19	1.9
pRLA5.79	20.24	2.9
pRLA5.80	19.36	2.8
pRLA5.81	17.35	2.5
pRLA5.82	69.37	10
pRLA5.83	69.84	10
pRLA5.84	14.40	4.9
pRLA5.85	57.12	19
pRLA5.86	6.99	2.4
pRLA5.87	124.85	42
pRLA5.88	14.14	4.8
pRLA5.89	28.34	9.6
pRLA5.90	239.23	81
pRLA5.91	156.30	53
pRLA5.92	56.42	19
pRLA5.93	40.84	14
pRLA5.94	14.88	5.1
pRLA5.95	27.00	4.6
pRLA5.96	73.17	13
pRLA5.97	71.06	12
pRLA5.98	83.64	14
pRLA5.99	10.01	1.7
pRLA5.100	20.74	3.6
pRLA5.101	15.99	5.4
pRLA5.102	15.00	5.0
pRLA5.103	5.64	1.9
pRLA5.104	15.42	5.2
pRLA5.105	9.39	3.2
pRLA5.106	8.84	1.5
pRLA5.107	10.19	1.8
pRLA5.108	11.17	2.0
pRLA5.109	22.46	4.0
pRLA5.110	36.07	8.1
pRLA5.111	88.14	20
pRLA5.112	5.25	1.2
pRLA5.113	7.69	1.7
pRLA5.114	4.79	1.1
pRLA5.115	2.80	0.6
pRLA5.116	8.53	1.1
pRLA5.117	11.01	1.4
pRLA5.118	12.95	1.7
pRLA5.119	9.73	1.3
pRLA5.120	28.93	3.8
pRLA5.121	53.99	7.0
pRLA5.122	295.79	38
pRLA5.123	51.68	6.7
pRLA5.124	8.48	1.1
pRLA5.125	5.76	0.7
pRLA5.126	18.04	2.3
pRLA5.127	28.73	3.7
pRLA5.128	10.88	1.4
pRLA5.129	14.04	1.8
pRLA5.130	3.38	0.9
pRLA5.131	21.02	5.6
pRLA5.132	7.76	2.1
pRLA5.133	12.02	3.2
pRLA5.134	9.90	2.6
pRLA5.135	4.12	1.1

pRLA5.136	2.76	0.7
pRLA5.137	4.57	1.2
pRLA5.138	6.80	1.8
pRLA5.139	6.10	1.6
pRLA5.140	4.60	1.2
pRLA5.141	27.41	7.3
pRLA5.142	18.64	5.7
pRLA5.143	55.85	17
pRLA5.144	21.67	6.6
pRLA5.145	20.78	6.3
pRLA5.146	2.96	0.9
pRLA5.147	4.51	1.4
pRLA5.148	11.92	4.8
pRLA5.149	5.32	2.1
pRLA5.150	3.88	1.6
pRLA5.151	19.91	8.0
pRLA5.152	20.89	8.4
pRLA5.153	1.11	0.4
pRLA5.154	35.84	14
pRLA5.155	33.18	13
pRLA5.156	1.96	0.9
pRLA5.157	4.96	2.0
pRLA5.158	6.86	2.8
pRLA5.159	2.70	0.9
pRLA5.160	3.14	1.1
pRLA5.161	2.36	0.8
pRLA5.162	2.79	1.0
pRLA5.163	6.41	2.2
pRLA5.164	2.94	1.0
pRLA5.165	7.61	2.7
pRLA5.166	77.21	27
pRLA5.167	4.87	1.7
pRLA5.168	12.02	4.2
pRLA5.169	2.47	0.9
pRLA5.170	41.38	10
pRLA5.171	32.35	7.6
pRLA5.172	55.06	13
pRLA5.173	3.54	0.8
pRLA5.174	2.86	0.7
pRLA5.175	5.95	1.4
pRLA5.176	42.43	10
pRLA5.177	4.84	1.1
pRLA5.178	31.86	7.5
pRLA5.179	32.52	7.6
pRLA5.180	14.17	3.3
pRLA5.181	14.47	4.0
pRLA5.182	106.36	29
pRLA5.183	111.43	30
pRLA5.184	144.26	39
pRLA5.185	39.46	11
pRLA5.186	15.20	4.2
pRLA5.187	98.72	27
pRLA5.188	17.12	4.7
pRLA5.189	116.92	32
pRLA5.190	97.40	27
pRLA5.191	13.53	3.7
pRLA5.192	3.21	0.6
pRLA5.193	145.40	29
pRLA5.194	5.09	1.0
pRLA5.195	3.18	0.6
pRLA5.196	8.38	1.7
pRLA5.197	14.29	2.8

pRLA5.198	116.28	23
pRLA5.199	13.53	2.7
pRLA5.200	9.36	1.8
pRLA5.201	4.78	0.9
pRLA5.202	5.17	1.0
pRLA5.203	37.31	7.4
pRLA5.204	29.54	12
pRLA5.205	16.93	7.0
pRLA5.206	32.40	13
pRLA5.207	35.60	15
pRLA5.208	11.83	4.9
pRLA5.209	5.07	2.1
pRLA5.210	17.85	7.5
pRLA5.211	10.79	4.5
pRLA5.212	8.99	3.7
pRLA5.213	63.92	27
pRLA5.214	7.84	3.3
pRLA5.215	3.40	1.4
pRLA5.216	74.14	31
pRLA5.217	3.05	1.7
pRLA5.218	2.88	1.6
pRLA5.219	3.65	2.0
pRLA5.220	9.88	5.5
pRLA5.221	17.17	10
pRLA5.222	11.55	6.4
pRLA5.223	4.73	2.6
pRLA5.224	2.85	1.6
pRLA5.225	2.43	1.3
pRLA5.226	7.96	4.4
pRLA5.227	8.51	4.7
pRLA5.228	2.46	1.4
pRLA5.229	4.03	2.2
pRLA5.230	13.14	7.1
pRLA5.231	1.84	1.0
pRLA5.232	2.31	1.2
pRLA5.233	2.55	1.4
pRLA5.234	2.93	1.6
pRLA5.235	6.56	4.5
pRLA5.236	11.35	7.8
pRLA5.237	15.20	10
pRLA5.238	11.84	8.1
pRLA5.239	14.06	10
pRLA5.240	219.08;299.60	76;104
pRLA5.241	5.94;5.24	2.0;1.8
pRLA5.242	7.20	2.5
pRLA5.243	3.39	1.5
pRLA5.244	4.35	1.9
pRLA5.245	15.97	7.2
pRLA5.246	19.71	8.5
pRLA5.247	23.93	10
pRLA5.248	4.87	2.1
pRLA5.249	20.37	8.7
pRLA5.250	20.63	8.8
pRLA5.251	10.38	4.5
pRLA5.252	4.32	1.9
pRLA5.253	3.99	1.7
pRLA5.254	5.67	2.4
pRLA5.255	20.47	8.8
pRLA5.256	28.00	12
pRLA5.257	4.00	1.7
pRLA5.258	13.66	5.6
pRLA5.259	30.53	13

pRLA5.260	16.77	6.9
pRLA5.261	2.35	1.0
pRLA5.262	24.95	10
pRLA5.263	5.66	2.3
pRLA5.264	2.46	1.0
pRLA5.265	11.28	4.7
pRLA5.266	5.08	2.1
pRLA5.267	4.22	1.4
pRLA5.268	95.12	32
pRLA5.269	4.90	1.7
pRLA5.270	5.86	2.0
pRLA5.271	67.01	23
pRLA5.272	63.06	22
pRLA5.273	109.65	37
pRLA5.274	2.77	0.94
pRLA5.275	24.20	8.3
pRLA5.276	38.83	13
pRLA5.277	1.72	0.6
pRLA5.278	2.15	0.7
pRLA5.279	35.03	12
pRLA5.280	31.28	11
pRLA5.281	5.63	1.6
pRLA5.282	2.52	0.7
pRLA5.283	2.17	0.6
pRLA5.284	2.09	0.6
pRLA5.285	4.81	1.4
pRLA5.286	2.62	1.0
pRLA5.287	3.24	1.2
pRLA5.288	2.62	1.4
pRLA5.289	2.48	1.3
pRLA5.290	1.89	1.0
pRLA5.291	2.70	1.4
pRLA5.292	2.81	1.5
pRLA5.293	2.12	1.4
pRLA5.294	3.80	2.4
pRLA5.295	2.66	1.7
pRLA5.296	2.25	1.4
pRLA5.297	7.07	4.5
pRLA5.298	7.02	4.5
pRLA5.299	6.94	4.5
pRLA5.300	2.34	1.5
pRLA5.301	1.55	1.0
pRLA5.302	3.84	2.5
pRLA5.303	1.74	1.1
pRLA5.304	13.89	3.1
pRLA5.305	7.61	1.7
pRLA5.306	66.80	15
pRLA5.307	9.66	2.2
pRLA5.308	166.57	38
pRLA5.309	10.46	2.4
pRLA5.310	15.51	3.5
pRLA5.311	13.45	3.0
pRLA5.312	14.73	3.3
pRLA5.313	6.01	1.4
pRLA5.314	13.31	3.0
pRLA5.315	4.73	1.0
pRLA5.316	12.84	2.8
pRLA5.317	35.40	7.8
pRLA5.318	19.78	4.4
pRLA5.319	69.81	16
pRLA5.320	39.53	8.8
pRLA5.321	16.86	3.7

pRLA5.322	212.17	39
pRLA5.323	217.40	39
pRLA5.324	171.38	31
pRLA5.325	169.11	31
pRLA5.326	127.32	23
pRLA5.327	55.71	10
pRLA5.328	59.68	11
pRLA5.329	12.29	2.7
pRLA5.330	5.54	1.2
pRLA5.331	4.47	0.9
pRLA5.332	21.43	3.4
pRLA5.333	24.71	3.9
pRLA5.334	10.91	1.9
pRLA5.335	113.99	20
pRLA5.336	40.18	7.1
pRLA5.337	162.45	29
pRLA5.338	27.82	4.9
pRLA5.339	184.95	33
pRLA5.340	176.03	31
pRLA5.341	165.02	29
pRLA5.342	41.06	5.2
pRLA5.343	29.81	3.8
pRLA5.344	11.33	2.3
pRLA5.345	6.11	1.3
pRLA5.346	9.27	1.9
pRLA5.347	10.08	2.3
pRLA5.348	11.15	2.6
pRLA5.349	7.51	1.7
pRLA5.350	4.33	1.0
pRLA5.351	152.15	17
pRLA5.352	152.73	17
pRLA5.353	198.98	22
pRLA5.354	210.96	23
pRLA5.355	310.01	34
pRLA5.356	327.26	69
pRLA5.357	61.54;68.76	13;15
pRLA5.358	64.96;69.21	14;15
pRLA5.359	9.86	2.1

ABSTRACT FOR "A STUDY OF GENE EXPRESSION IN *PSEUDOMONAS*" BY R.L.ALLEN

An inherent problem in the study of the genetics of the interesting and potentially commercially useful properties of the pseudomonads is that of gene expression, since many of the genes encoding these properties are not well expressed in an *E.coli* background. The evidence available at the present time indicates that some *Pseudomonas* genes may possess different promoter sequences not recognised by *E.coli* RNA polymerase.

An *in vitro* coupled transcription/translation system based on *P.putida* has been developed. A comparison of *E.coli* and broad host range plasmid DNA in this and the equivalent *E.coli* system showed that although cloned *E.coli* and vector polypeptides were synthesised in both systems, there was a difference in the polypeptide products directed by broad host range plasmid DNA in the two systems. In particular RSF1010 directed the synthesis of a 73kD polypeptide uniquely in the *P.putida* system. This was shown to be a polypeptide involved in mobilisation of the plasmid.

A broad host range promoter-probe vector based on RSF1010 was constructed and used for the shotgun cloning of *P.putida* promoters. A small subset of fragments which were active as promoters in *P.putida* but exhibited much lower activity in *E.coli* were isolated, sequenced and analysed with respect to concensus *E.coli* and nitrogen-regulated promoter sequences. These isolated DNA fragments may represent promoters which have sequences specifically recognised by *Pseudomonas* RNA polymerase. An analysis of published *Pseudomonas* chromosomally-encoded promoters revealed putative *Pseudomonas*-specific concensus regions.