DOCTOR OF PHILOSOPHY

The catabolism of aromatic esters by Acinetobacter sp.ADP1

Jones, Rheinallt M.

Award date:
2000

Awarding institution:
Bangor University

The catabolism of aromatic esters by *Acinetobacter* sp. ADP1

Thesis a gyflwynwyd i ddyfarniad Prifysgol Cymru Bangor

gan

Rheinallt M. Jones

mewn ymgeisyddiaeth am radd *Philosophiae Doctor*

Goruchwyliwr: Yr Athro Peter A. Williams

Mai 2000

Summary

A 10 kbp region upstream of the *ben-cat* supraoperonic cluster of catabolic genes was cloned from *Acinetobacter* sp. strain ADP1. This region contains genes that encode the catabolism of aromatic esters, converting them to the corresponding aromatic carboxylic acids by the sequential action of inducible enzymes.

A set of genes for three inducible enzymes: *areA*-encoding an esterase, *areB*-encoding a benzyl alcohol dehydrogenase, and *areC*-encoding a benzaldehyde dehydrogenase, were found to be located 3.5kbp upstream of the *ben* genes. Individual disruptions of the *areA*, *areB* and *areC* genes on the chromosome by insertion of a Km-resistance cassette, resulted in strains which were severely reduced in the ability to utilise aryl esters as sole carbon growth substrates as compared to the wild-type strain.

Each *are* gene was individually expressed to high specific activity in *Escherichia coli.* The relative activities against different substrates of the cloned enzymes were, within experimental error, identical to that of wild-type *Acinetobacter* sp. ADP1 grown on benzyl acetate, benzyl alcohol or 4-hydroxybenzyl alcohol as the carbon source. The substrate preferences of all three enzymes were broad, encompassing a range of substituted aromatic compounds and in the case of the AreA esterase, different carboxylic acids.

Mutants blocked in genes of the β-ketoadipate pathway with or without an additional *areA* mutation confirmed that benzyl acetate, and 2- and 4-hydroxybenzyl acetates were channelled into the β-ketoadipate pathway at the level of benzoate, catechol and protocatechuate respectively.

Studies of the expression of *areCBA* by RT-PCR, by biochemical assays, and by *are*::*lacZ* chromosomal gene fusions found that the genes are co-transcribed on the same mRNA and are induced as an operon by benzyl acetate, benzyl alcohol or benzaldehyde.

Adjacent to *areC* is the gene *areR* that encodes for a protein that has amino acid sequence homology to members of the σ^{54}-dependent NtrC family of transcriptional regulators. Sequence analysis of the promoter region upstream of *areC* showed that it contained all the elements typical of σ^{54}-dependent gene expression. A disruption of

areR on the chromosome by insertion of a Km-resistance cassette resulted in a mutant that was unable to induce AreA, B, or C activity.

Sequence analysis of the region adjacent to *areA* identified four open reading frames: *salA, salR, salE* and *salD*, which are organised in two convergent transcription units, *salAR* and *salDE*, as was confirmed by RT-PCR.

salD encodes a protein with undetermined function but which shows homologies with membrane proteins thought to be involved in hydrocarbon uptake. A Km^r insertion in *salD* deleteriously affected cell growth and viability. *salE* was cloned into expression vector pET5a and its product was shown to have esterase activity against short-chain alkyl esters of 4-nitrophenol. SalE was also shown to have hydrolytic activity against ethyl salicylate, producing ethanol and salicylic acid. The insertional inactivation of *salE* with a Km^r cassette resulted in a mutant that lost the ability to utilise only ethyl and methyl salicylates of the esters tested as sole carbon sources and no esterase activity against ethyl salicylate could be detected in cell extracts. SalE was induced during growth on ethyl salicylate but not on salicylate itself.

The SalA and SalR gene products are homologous to two *Pseudomonas* proteins NahG and NahR, respectively encoding salicylate hydroxylase and the positive LysR-family regulator of both salicylate and naphthalene catabolism. *salA* was cloned into pUC18 together with *salR* and *salE* and showed salicylate-inducible hydroxylase activity against a range of substituted salicylates, with the same relative specific activity as found in wild type ADP1 grown on salicylate. Knockout mutations with Km^r cassettes of *salA* and *salR* respectively, resulted in mutants that were unable to grow on salicylate or ethyl salicylate, or to express salicylate hydroxylase activity.

Mutants blocked in genes of the β-ketoadipate pathway with or without an additional *salE* mutation confirmed that ethyl salicylate and salicylate were channelled into the β-ketoadipate pathway at the level of catechol and thence dissimulated by the *cat* genes products.

The function of the *are* and *sal* genes is to catabolise a range of aryl esters and alkyl salicylate into carboxylic acids, which are further catabolised to central metabolites by the adjacent *ben* and *cat* genes. This contributes to the metabolic diversity of ADP1 and enables it to grow on substrates that are likely to be found as natural breakdown products of plant compounds.

Contents

List of Figures

Tables

Acknowledgements

I would firstly like to extend my gratitude to my supervisor, Professor Peter Williams for his steady and consistent guidance and friendship throughout the period of my study. I would also like to thank the Biotechnology and Biological Sciences Research Council (BBSRC) for their financial support during the project.

I would like to thank my colleagues in the laboratory, namely Dr. Linda Shaw, Dr. Keith James, Dr. Richard Ogden, Dr. Mike Barnes, Dr. Sergio Fuenmayor, Dr. Kevin Hallberg, Dr. Michelle Hughes, Dr. Ian Ross, Miss Bethan Stallwood, and Dr. Ning-Yi Zhou for their time and help with advice on various techniques, and with maintaining a happy and pleasant working atmosphere.

I wish to thank Dr. Ellen Neidle and Miss Lauren Collier of The University of Georgia, for the gift of pBAC98 which really got the project off to a running start. I also wish to thank Dr. Helen Wing of The University of Birmingham for her guidance with the Primer Extension Assays.

Carwn ddiolch i fy Mam a 'Nhad am bob cefnogaeth a gefais dros y blynyddoedd. Tybiaf fod D.O. yn rhywle gyda gwên lydan ar ei wyneb o weld 'Y Llyfr' wedi ei gwblhau o'r diwedd! Carwn hefyd ddiolch i weddill y teulu am eu cefnogaeth hwythau.

Yn olaf, carwn ddiolch i Esyllt am fod yn gefn imi yn ystod y cyfnod hwn. Diolch am dy gwmni a'th amynedd, ac am ddallt gydol bob Sul a hwyrnos pan fyddai gwaith yn galw.

Abbreviations

- Molarity quantities and solutions are stated as moles (mmol and μmol) or M (mM and μM).
- DNA bases are referred to as basepairs (bp) or kilobasepairs (kbp).
- Standard SI units are used throughout this thesis.

A_{600}:	Absorbance at 600nm.
Amp:	Ampicillin.
BADH:	Benzyl alcohol dehydrogenase.
BLAST:	Basic Local Alignment Search Tool.
BZDH:	Benzaldehyde dehydrogenase.
C23O:	Catechol 2,3-dioxygenase.
CAPS:	3-(Cyclohexylamino)-1-propanesulfonic acid.
dNTP:	Deoxy nucleotides.
DTT:	Dithiothreitol.
ε_m:	Molar extinction coefficient.
EDTA:	Ethylenediamine tetra-acetate.
Km:	Kanamycin.
LB:	Luria Broth.
MM:	Minimal media.
NAD:	Nicotinamide Adenine Dinucleotide.
NADH:	Nicotinamide Adenine Dinucleotide reduced form.
ORF:	Open reading frame.
PCR:	Polymerase Chain Reaction.
RT-PCR:	Reverse Transcriptase Polymerase Chain Reaction.
SDS:	Sodium dodecyl sulphate.
Tris:	2-amino-2hydroxymethylpropan-1,3-diol.
UV:	Ultra-violet light.
V:	Volts.
v/v:	Percentage solution made volume per volume.
w/v:	Percentage solution made weight per volume.

CHAPTER 1

INTRODUCTION

1.1.1. Microorganisms and the Carbon Cycle.

The biosphere is comprised of a complex and diverse mixture of carbon compounds. This mixture is in a constant state of flux as compounds are synthesised, transformed, or mineralised. The flux of carbon through the biosphere is known as the Carbon Cycle. The mineralisation and recycling of assimilated carbon compounds in the carbon cycle is dependent upon microorganisms. Microorganisms utilise a vast array of organic compounds as a sole carbon source that supports their growth. These compounds can vary from simple sugars, alcohols, acids, hydrocarbons, complex carbohydrates, or lipid polymers. The microorganisms use highly specialised pathways to catabolise these compounds to products that enter the common pathways of metabolism, such as the TCA cycle and β-oxidation.

1.1.2. Decaying plants are a food-source to microorganisms.

The decay of plant material is a diverse source of organic compounds for microorganisms to utilise as a carbon source for growth. These include lignin, cellulose, alkaloids, terpenes, and tannins to name but a few. Of these, the most abundant is lignin, which comprises about 25% of the land-based biomass on earth, and the recycling of lignin is a vital link in the carbon cycle.

Lignin is a complex three-dimensional polymer produced by an enzyme-initiated dehydrogenative polymerisation of three phenylpropane monomers, *p*-hydroxycinnamyl, coniferyl, and sinapyl alcohols. The monomeric phenylpropane units in lignin are linked to each other not by a single intermonomeric linkage but by several different carbon-carbon and ether linkages, most of which are not readily hydrolysable. Thus lignin is resistant to degradation by most microorganisms. Only certain fungi and probably certain strains of bacteria are capable of decomposing lignin. The best-known lignin-degrading microorganisms are the white-rot fungi. These are wood-decaying fungi that decompose lignin macromolecules. Lignin degradation involves the oxidative cleavage of side chains between α- and β- carbons

leading to the formation of aromatic acids. Many benzoic acids, benzyl alcohol, and benzaldehyde derivatives have been identified in among the low molecular weight products of fungus-degraded lignins. Such monomeric compounds are potential carbon-growth substrates for microorganisms. However, the relative stability of such aromatic compounds poses a formidable energy challenge to soil microorganisms due to the stable nature of the benzene ring.

1.1.3. The stability of the benzene ring.

The benzene ring, first described by Kekulé in 1865 as a cyclic hexatriene structure comprises six carbon atoms, each of which bonds to a single hydrogen atom and to two other carbons forming a symmetrical sigma-bonded framework. The ring is a regular hexagonal structure in which all carbon-bonds are equal (140 pm) with bond angles of 120°. Each sp^2 – hybridised trigonal carbon atom has a single p orbital containing one electron. The individual p orbitals overlap creating a continuous π system above and below the ring within the p electron circulate. The experimentally derived value for the stability of the benzene ring shows that it is more stable than the derived value of the theoretical non-delocalised cyclohexatrene unit. Thus the delocalization provided by the formation of the delocalised π system in benzene has a stabilising effect. Hence the chemistry of benzene predominantly involves substitution reactions in which the products retain the π system and only the hydrogen atoms are replaced by other groups. Nevertheless, to maintain the flux through the carbon cycle, the benzene rings assimilated by plant secondary metabolism must be cleaved and mineralised.

1.1.4. Isolation of microorganisms that are capable of aromatic catabolism.

The first microorganisms isolated and identified as having a diverse metabolic capability against aromatic compounds were all classified as *Pseudomonas* (Stanier *et al.* 1966). Subsequently, some of these microorganisms have been further classified and assigned to separate genera. One of these groups of microorganisms, which were non-motile diplococci, was classified into the genus *Moraxella*, now known as *Acinetobacter*. The *Acinetobacter* genus, as well as the *Pseudomonas* genus, proved to have a diverse metabolic capability against a range of aromatic substrates

(Baumann *et al.* 1968). However, *Pseudomonas* and *Acinetobacter* are quite different in morphology, motility, and the G+C content of their DNA. Subsequent 16S RNA nucleotide sequence studies have placed *Acinetobacter* together with the *Pseudomonas* group in the γ subdivision of Proteobacteria (Rainey *et al.* 1994).

One member of the *Acinetobacteriaceae*, strain ADP1 (BD413) thus far has not been sub-classified within the genus although for a long time it had been incorrectly designated as *A. calcoaceticus*. Originally isolated by Elliot Juni, it was shown to be highly competent for natural transformation (Juni, 1972). Transformation of a *trpE* mutant of this strain yielded prototrophs when any other member of the *Acinetobacter* genus served as the donor DNA strain (Juni, 1972). The natural transformation genetic system has been of great assistance and has enabled a genetic approach to biochemical and physiological questions that could not have been readily addressed in other ways. The ease by which mutants can be created by replacement of chromosomal genes with modified DNA is a highly advantageous aspect of *Acinetobacter* sp. ADP1 genetics. Also, the physiology of the organism allows for the selection of mutations causing loss of function, or any other specified change (Parke *et al.* 2000). These factors make *Acinetobacter* sp. ADP1 a highly desirable organism for studying the microbial catabolism of aromatic compounds. This is why *Acinetobacter* sp. ADP1 is the choice organism for this study, and in the remainder of this chapter various aspects of the biology of *Acinetobacter* sp. ADP1 will be introduced.

1.2. Natural Transformation.

1.2.1. Natural transformation has been observed across several genera of bacteria.

Several species from various genera of bacteria have been found to be naturally competent. Natural transformation enables the bacteria to take up DNA from the environment without the need of induction by physical or chemical pre-treatment (Stewart & Carlson, 1986). Transformation is a process where DNA is transferred from the extracellular environment into the recipient cell's interior. Providing the transformed DNA is complementary to the recipient's DNA, it can recombine and alter the recipient's DNA, which may lead to an altered genotype and phenotype. The

bacterial cells take up DNA when they are in a competent state and competence may be induced in a number of ways. These include growth limitations in the case of *Bacillus* (Dabnau, 1991), or the concentration of calcium and a competence factor in the case of *Streptococcus pneumoniae* (Trombe *et al.* 1992). A common form of competence induction occurs when the bacterial cells are in exponential growth. Some of the bacteria that have this property include *Neisseria gonorrhoeae* (Biswas *et al.* 1989), *Deinococcus radiodurans* (Tirgari & Moseley, 1980), and the choice organism for study in this thesis, *Acinetobacter* sp. ADP1 (Juni, 1972). Natural transformation in *Acinetobacter* sp. ADP1 has been extensively studied. The mechanism of DNA uptake has been studied and reviewed Palmen & Hellingwerf (1997), the genes involved with competence have been cloned (Busch *et al.* 1999), and a study of the transformation of ADP1 in soil microcosms has been carried out (Nielsen *et al.* 2000).

1.2.2. Natural Transformation in *Acinetobacter* sp. ADP1.

Natural transformation is induced in *Acinetobacter* sp. ADP1 as a result of increased nutrient availability. Maximum competence is induced when a nutrient up-shift is imposed on stationary cells. The ability to take up DNA is gradually lost during the late stages of exponential growth and after entrance into the stationary state. It was determined by Palmen *et al.* (1994) that competence in ADP1 is regulated. However, it was also found that the 'amount of DNA taken up by ADP1 is not sufficient to support significant growth', which led to the conclusion that the biological function of natural transformation is not to provide the cell with nutrients.

ADP1 has a DNA translocation system, which does not differentiate between chromosomal or plasmid DNA, nor does it display sequence specificity. The DNA enters the cell in single stranded form due to the presence of an exonuclease enzyme in the DNA transport system. In addition, transformation is also an energy dependent process as shown by being inhibited by the addition of uncouplers (Palmen *et al.* 1992). Transformation was inhibited in studies using DNA that carried a covalently bound protein. The translocation system was not able to translocate a DNA/protein complex, due to it blocking the DNA uptake site. It was also found that transformation efficiency depends on the length of the transforming DNA, where the longer the DNA strand, the higher the efficiency. It has been shown that up to 25 % of

the cells in a culture can be transformed by saturating amounts of DNA, over a 3 hour incubation (Palmen *et al.* 1993).

1.2.3. Competence factors are involved in Natural Transformation of *Acinetobacter* sp. ADP1.

Four genes, *comC, comP, comE,* and *comF* have been shown to encode proteins that are necessary for natural transformation in *Acinetobacter* sp. ADP1. The first of these genes to be identified was *comP*. ComP is similar to prepilins of type IV pili and to pilin-like components of protein translocation machinery. Insertional mutagenesis of *comP* resulted in a reduction in the transformation efficiency of ADP1 (Porstendorfer *et al.*1997). A second gene, *comC* was also found to be essential for natural transformation. ComC has amino acid sequence similarities to various type IV pilus assembly factors. A *comC* deficient mutant was not able to bind or take-up DNA (Link *et al.* 1998). Downstream of *comC*, there are two genes designated *comE* and *comF*. Both genes encode proteins that are similar to pilins and pilin-like components. Both *comE* and *comF* genes have been mutated, which resulted in the natural transformation in *comF* mutants being reduced by 1,000-fold, whereas comE mutants exhibited 10-fold-reduced transformation frequencies. It was also determined that ComE and ComF are specific for DNA translocation, since *comE* and *comF* defects did not affect other secretion mechanisms (Busch *et al.* 1999).

1.2.4. The availability of DNA to *Acinetobacter* sp. ADP1 in soil Microcosms.

Factors affecting natural transformation of *Acinetobacter* sp.ADP1 in soil microcosms have been investigated. It was found that ADP1 cells in soil can be easily induced by nutrients to undergo natural transformation with chromosomal DNA. It was also shown that access to nutrients is critical for the development and maintenance of competence in soil, which is also affected by abiotic factors like moisture level and phosphate salt concentration (Nielson *et al.* 1997a). In a further study, a small microcosm, based on optimised *in vitro* transformation conditions, was used to study the ecological factors affecting the transformation. When chromosomal DNA was released into soil, it rapidly became unavailable for transformation of ADP1. In

addition, strain ADP1 quickly lost the ability to receive, stabilise, and express exogenous DNA after its introduction into soil (Nielson *et al.* 1997b).

Acinetobacter sp. BD413 (pFG4) was the choice model organism to study the biological significance of dead bacterial cells in soil to the intra- and interspecies transfer of gene fragments by natural transformation. BD413(pFG4) has a partial kanamycin resistant gene. The kanamycin sensitive BD413(pFG4) was exposed to lysates of kanamycin resistant *Acinetobacter* spp., *Pseudomonas*, and *Burkholdaria.* The investigation revealed the ability of all three DNA sources to transform BD413(pFG4) after 4 days incubation in soil, which demonstrated ADP1's ability to take up DNA regardless of the genera of bacteria from which the DNA originated (Nielson *et al.* 2000).

1.2.5. Uses of the Natural Transformation System in *Acinetobacter* sp. ADP1.

Studies have found that *Acinetobacter* sp. ADP1 is extremely competent for natural transformation and therefore replacement of chromosomal genes with modified DNA is a highly efficient process. Various studies have used natural transformation to modify the ADP1 chromosome.

The transformation system was used to recover mutant chromosomal DNA within recombinant plasmids by a technique known as 'Gap Repair'. Here, a linearised plasmid containing the DNA that encompasses the mutation, was transformed into the mutant ADP1 stain. The gap in the plasmid was repaired by using the chromosomal fragment as template DNA in a recombinational event. The resultant plasmid was thus circularised, and in addition, was able to replicate. The plasmid contained an antibiotic resistance gene, and was selected for, and recovered. The technique was successful in recovering fragments of up to 7.0 kbps in length (Gregg-Jolly & Ornston, 1990).

Cloned ADP1 genes have been identified using the natural transformation system. Wild type ADP1 DNA cloned into *E. coli* transforms mutant ADP1 cells with high efficiency. Therefore, *E. coli* colonies containing cloned genes can be identified by replica plating onto a lawn of ADP1 mutant cells. This procedure was used to demonstrate the clustering of ADP1 genes required for utilization of 4-hydroxybenzoate (Averhoff *et al.* 1992).

The natural transformation system was exploited to produce a high–resolution genetic and physical map of ADP1. In this technique, previously isolated ADP1 genes were modified to incorporate the recognition site for the restriction endonuclease *Not*I. Transformation of the modified DNA into ADP1 allowed for the precise placement of the *Not*I restriction site within 40 known genes. Analysis of the resulting restriction fragments with gel electrophoresis and Southern Blotting allowed a high resolution map of the ADP1 chromosome to be constructed (Gralton *et al.* 1997).

Large numbers of ADP1 mutants were obtained in the study of the *pobR* gene encoding the transcriptional activator of pobA, by using natural transformation with fragments of DNA from error-prone polymerase chain reaction (PCR). Here, amplified fragments of *pobR* from error-prone PCR were randomly transformed into ADP1. The mutations were transferred directly to the chromosome via homologous recombination. Several independent mutations of wild type *pobR* were recovered. Sequence analysis of the mutant alleles showed that the mutations were predominantly single-nucleotide substitutions broadly distributed within *pobR* (Kok *et al.* 1997).

Natural transformation has been widely used for the insertion of an antibiotic resistance cassette into the ADP1 genome. The insertion of such a heterologous gene is used to disrupt the function of a given gene. This technique requires a recombinant fragment of ADP1 DNA to have been sequenced and mapped. Then an antibiotic resistance marker cassette is inserted into the open reading frame of the gene of interest, thus disrupting the function of the gene by premature termination. The plasmid is then linearised and transformed into ADP1. Recombination of the heterologous gene into the genome will occur at the point where the flanking sequence has homology to the wild type DNA. The recombination is selected for by antibiotic pressure. The result is the creation of a mutant ADP1 stain with a disruption in the gene of interest. The phenotype of the mutant strain can be studied with respect to the hypothesised function of the gene.

The molecular techniques listed above have facilitated the study of the β-ketoadipate pathway in ADP1. This pathway catalyses the enzymatic cleavage of the aromatic rings in catechol (1,2-dihydroxybenzene) or protocatechuate (3,4-dihydroxybenzoate), and channels the cleavage products into the tricarboxylic acid cycle.

1.3. Catabolism of the aromatic substrates.

1.3.1. Aerobic microbial cleavage of the benzene ring.

Certain microorganisms have evolved enzymes that are able to mineralise naturally occurring aromatic compounds. They are able to utilise these compounds to produce energy and to serve as their sole carbon growth substrate. Mineralisation effectively involves the complete dissociation of the aromatic ring, producing carbon dioxide as the end-product of respiration, which subsequently is recycled by photosynthesis. This recycling of organic matter is largely predicated on the enormous catalytic diversity of microbial enzymes. The enzymes are organised into well-regulated metabolic pathways, which provides for the release of energy that can be captured as ATP. The microbial enzymes responsible for the mineralisation have been well-studied and their mechanisms eludicated. Aromatic catabolic pathways consist of three distinct stages: (1) preparation of the aromatic ring for cleavage; (2) ring cleavage; (3) dissimilation into central metabolites (Harwood & Parales, 1996).

1.3.2. Preparation of the aromatic ring for cleavage.

The first step in the aerobic mineralisation of aromatic compounds is the formation of the dihydroxylated benzene ring compounds catechol (1,2-dihydroxybenzene) and protocatechuate (3,4-dihydroxybenzoate). These two compounds are central players in aromatic catabolism and virtually all the studied pathways for the catabolism of the naturally occurring aromatic compounds feed through catechol or protocatechuate (Harwood & Parales, 1996).

1.3.3. Benzene rings are cleaved by specialised microbial enzymes.

The second phase of degradation involves ring cleavage. There are three main mechanisms which microorganisms have developed for ring cleavage. These are known as *ortho*-, *meta*-, and gentisate cleavage (Fig. 1.3.1.) (Harwood & Parales, 1996).

The *ortho* cleavage of catechol and protocatechuate occurs between the two hydroxyl side groups (also known as intradiol cleavage) producing *cis*,*cis*-muconate, and β-carboxymuconate respectively. The *ortho* cleavage pathway is also known as

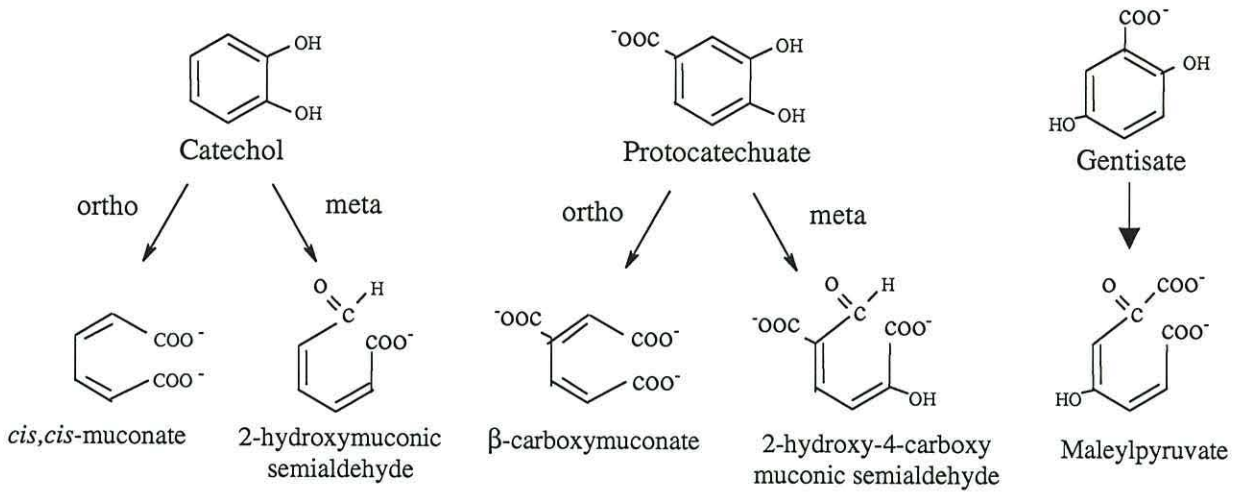


Fig. 1.3.1. Major aerobic routes of aromatic ring cleavage. The three major routes for ring cleavage are *ortho-*, *meta-*, and gentisate cleavage. *Ortho* cleavage occurs between the two hydroxyl side groups and is also known as intradiol cleavage, while *meta* cleavage occurs adjacent to one of the hydroxyls and is also known as extradiol cleavage (Harwood & Parales, 1996). The products of ring cleavage are also shown.

the β-ketoadipate pathway, because the products of catechol and protocatechuate cleavage both converge into β-ketoadipate (Fig.1.4.1.) (Ornston & Neidle, 1991; Harwood & Parales, 1996).

The *meta* cleavage (also known as extradiol cleavage) of catechol occurs adjacent to one of the hydroxyls and produces 2-hydroxymuconic semialdehyde as a product. There are two different *meta* cleavages of protocatechuate, catalysed by protocatechuate 2,3 dioxygenase (P2,3O) or by protocatechuate 4,5 dioxygenase (P4,5O) respectively. The P2,3O cleavage of protocatechuate produces 2-hydroxy-5-carboxymuconic semialdehyde, and the P4,5O cleavage of protocatechuate produces 2-hydroxy-4-carboxymuconic semialdehyde (Fig. 1.3.1.). The third ring cleavage pathway, the gentisate pathway, is followed when ring cleavage occurs when two hydroxyl groups are *para* to each other, and cleavage occurs between the carboxyl-substituted carbon and an adjacent hydroxylated carbon (Fig. 1.3.1.). Compared to the catechol and protocatechuate fission, gentisate cleavage has been far less intensively studied.

1.4. The β-ketoadipate pathway.

1.4.1. Biological distribution of the β-ketoadipate pathway.

The β-ketoadipate pathway has been found in diverse prokaryotic organisms. These include *Arthrobacter*, *Acinetobacter*, *Bacillus*, *Rhodococcus*, *Alcaligenes*, *Burkholdaria*, *Comomonas*, *Pseudomonas*, *Azotobacter*, *Rhizobium* and *Braydyrhizobium* (Harwood & Parales, 1996). Also, the β-ketoadipate pathway has been found in the plant pathogen *Agrobacterium tumefaciens* (Parke, 1995), in the nitrogen-fixing plant symbiotic species and the actinomyces *Pseudonocardia* and *Micromonospora* (Grund & Kutzner, 1998) as well as in the aniline-assimilating bacterium *Frateuria* (Murakami *et al.* 1999).

1.4.2. Enzyme function in the β-ketoadipate pathway.

The chemistry and enzymology of the β-ketoadipate pathway has been extensively studied in *Acinetobacter* sp. ADP1 (Fig. 1.4.1.). These studies have been facilitated by ADP1's natural transformation system. In ADP1, the enzymes that convert

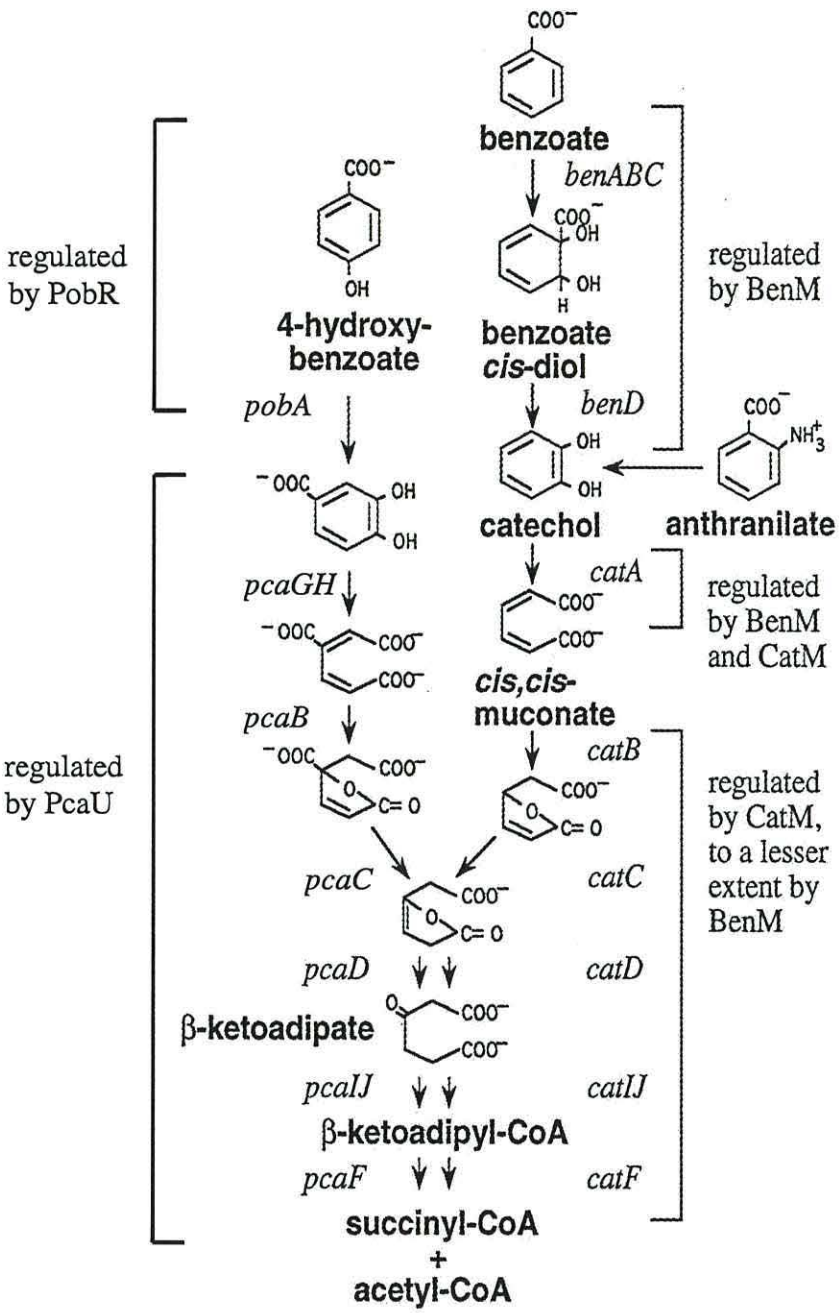


Fig. 1.4.1. The β-ketoadipate pathway of Acinetobacter sp. strain ADP1. The β-ketoadipate pathway is involved in the conversion of protocatechuate and catechol to amphibolic intermediates (Ornston & Neidle, 1991; Harwood & Parales, 1996). The organisation of the genes is presented in Fig. 1.4.2. and the regulators of the catabolic operons are also shown (Gralton *et al.* 1997).

catechol and protocatechuate to tricarboxylic acid intermediates are well characterised. Protocatechuate 3,4 dioxygenase (P3,4O) (*pcaGH*) (Doten *et al.* 1987) and catechol 1,2 dioxygenase (C1,2O) (*catA*) (Neidle & Ornston, 1987) are parallel steps in the two branches of the β-ketoadipate pathway, and both enzymes catalyse analogous reactions.

P3,4O catalyses the intradiol cleavage of protocatechuate by incorporating two atoms of molecular oxygen to form β-carboxy- *cis,cis*-muconate. Studies of P3,4O in *P.putida* have revealed that for activity, the enzyme requires a ferric ion to be ligated by the histidyl and tyrosyl residues of its β subunit (Ohlendorf *et al.*1994). The crystal structure of P3,4O has also been studied in *P. putida*. It has an $\alpha\beta Fe^{3+}$ protomeric structure, of which the two subunits share 30 % amino acid identity (Ohlendorf *et al.*1994).

On the other hand, C1,2O catalyses the incorporation of oxygen to form *cis,cis*-muconate. The structure of C1,2O has been studied in *P. arvilla*. It also has a dimeric form which requires a ferric ion for activity and has conserved tyrosyl and histidyl residues at the positions corresponding to those of P3,4O. The C1,2O structures studied so far are almost all homodimers (Nakai *et al.* 1990).

The next two parallel enzymes in the β-ketoadipate pathway in ADP1 are β-carboxy- *cis,cis*-muconate lactonizing enzyme (CMLE) (*pcaB*), and *cis,cis*-muconate lactonizing enzyme (MLE) (*catB*). The reactions of these two enzymes differ from each other in that CMLE mediates an *anti*- cycloisomerization whereas MLE catalyses a *syn* addition to the double bond of *cis,cis*-muconate (Fig. 1.4.1.) (Chari *et al.*1987a). The amino acid sequences of these two enzymes show no similarity.

The next two parallel enzymes are 3-carboxymuconolactone decarboxylase (CMD)(*pcaC*) which converts carboxymuconolactone to β-ketoadipate enol-lactone, and muconolactone isomerase (MI)(*catC*) which converts muconolactone to β-ketoadipate enol-lactone. These two enzymes are not analogous. Their amino acid sequences have little similarity, and the enzymes have opposite stereofacial specificities for the substrate (Fig. 1.4.1.) (Chari *et al.*1987b).

The last two steps of the β-ketoadipate pathway in ADP1 are catalysed by enzymes that are tightly conserved. The steps involve a coenzyme A transfer (*pcaIJ* and *catIJ*), and a thiolytic cleavage (*pcaF* and *catF*), which convert β-ketoadipate to

succinyl CoA and acetyl CoA (Fig. 1.4.1.) (Shanley *et al.* 1994). The *pcaIJF* and the *catIJF* genes share 98% nucleotide sequence identities for each pair of aligned genes.

1.4.3. Genes coding for β-Ketoadipate pathway enzymes are arranged in supra-operonic clusters.

The β-ketoadipate pathway genes have been cloned and sequenced in *Acinetobacter* sp. ADP1. These studies revealed a supraoperonic clustering of genes (Fig.1.4.2.). The *cat* genes for catechol catabolism are located adjacent to the *ben* genes for the catabolism of benzoate to catechol. Also, the *pca* genes for protocatechuate catabolism are located adjacent to the *qui* and the *pob* genes for the catabolism respectively of quinate and *p*-hydroxybenzoate to protocatechuate. The *cat* and the *pca* genes are located 270Kb apart on the *Acinetobacter* sp. ADP1 chromosome (Gralton *et al.* 1997). The relative positions of the *cat* and *pca* genes have also been determined in *P.aeruginosa.* They are located on opposite sides of the chromosome, with 2000kb separating both sets of genes (Holloway *et al.* 1994).

Recently, the genes adjacent to the *pca-qui-pob* genes have been cloned. On one side of the *pca-qui-pob* genes, adjacent to the *pca* genes, are the *adu* genes (adipate utilisation) for the catabolism of dicarboxylic acids and on the other side, adjacent to *pob* genes are the *ppa* genes (phenylpropenoid utilisation) (L.N Ornston, Yale University, CT, USA, personal communication). The *ppa* genes are part of the pathways that funnel the lignin monomers ferulate, coumarate, caffeate, chlorogenate and p-hydroxyphenyl propionate into the β-ketoadipate pathway (Fig. 1.5.1.) (Parke *et al.* 2000).

1.4.4. β-ketoadipate pathway organisation in ADP1.

An interesting feature of the organisation of the β-ketoadipate pathway genes in ADP1 is the similarities in the genes coding the final two analogous steps of the pathway. The genes, *pcaIJF* of the protocatechuate branch and *catIJF* of the catechol branch share over 98% nucleotide sequence identities for each pair of aligned genes. The *pcaIJF* and the *catIJF* genes are separated by 270 kbp, which raises the question

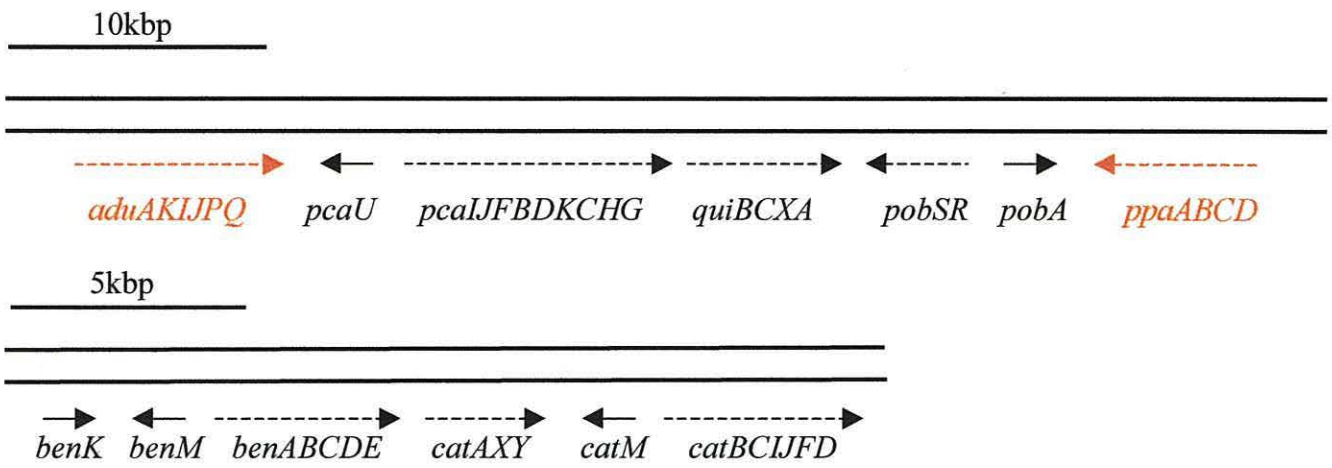


Fig. 1.4.2. Two islands of catabolic diversity in the *Acinetobacter* sp. ADP1 chromosome. The figure presents the locations of the genes in genetic islands of catabolic diversity in *Acinetobacter* sp. ADP1. Arrows below the gene designation indicate directions of transcription. Gene designations highlighted in red type are unpublished (L.N. Ornston, Yale University. CT. USA). The genes designated as *adu* are recently characterized genes that are involved in the catabolism of adipate and other straight chain dicarboxylic acids. The *pca* genes are involved in the catabolism of protocatechuate to amphibolic intermediates. The *qui* genes are involved in the catabolism of quinate to protocatechuate. The *pob* gene products are responsible for the catabolism of 4-hydroxybenzoate to protocatechuate, while the recently characterised *ppa* gene products are involved in the catabolism of the phenylpropenoids: chlorogenate, 4-hydroxylphenyl propionate, coumerate, ferulate, and caffeate. The *ben* genes are involved in the catabolism of benzoate to catechol, while the *cat* genes are involved in the conversion of catechol to amphibolic intermediates (Harwood & Parales, 1996). The two islands of catabolic genes are separated on the ADP1 genome by 280kbp (Gralton *et al.* 1997).

of why the DNA between these two regions is stable and not subject to deletion *via* recombination. Moreover, it is known that exchange of nucleotide sequences between *pcaIJF* and *catIJF* appears to contribute to repair of mutations. Mutations in *pcaIJF* undergo high frequency repair by recombination that depends upon both *catIJF* and *recA* (Kowalchuk *et al.* 1994; Gregg-Jolly & Ornston, 1994). This is in contrast to *pcaD and catD* of *Acinetobacter* sp. ADP1 which share only 52% amino acid identity. It is proposed that in the case of *catIJF/pcaIJF*, genetic recombination and accompanying repair of mutations may help cells to maintain an accurate protein product. In the case of *catD/pcaD* gene sequences, divergence occurred beyond the threshold where genetic exchange would be detrimental.

1.4.5. Regulators of the β-ketoadipate pathway in ADP1.

All of the genes for the degradation of catechol in *Acinetobacter* sp. ADP1 are activated by the LysR type regulator CatM (Neidle *et al.* 1989). It acts with the inducer *cis,cis* muconate to positively regulate *catA* and *catBCIJFD* gene expression (Fig. 1.4.1.). The CatM DNA binding sites upstream of *catA* and *catB* have been identified by footprinting experiments (Romero-Arroyo *et al.* 1995).

There is a second LysR-type regulator in the *ben-cat* operonic cluster designated *benM*. The amino acid sequence of BenM and CatM are 57% identical. An unusual regulatory feature of both BenM and CatM is their ability to recognise the same inducer, *cis,cis*-muconate, to activate the same genes i.e. catA and catB. However, unlike CatM, BenM activates *catA* and *catB* in response to benzoate, or benzoate together with *cis,cis*-muconate. Moreover, BenM also induces the *ben* genes for the catabolism of benzoate to catechol (Fig. 1.4.1). BenM activates the *ben* genes in response to benzoate, or benzoate together with *cis,cis*-muconate (Collier *et al.* 1998).

There are two identified regulator genes present in the *pca-qui-pob* gene cluster in ADP1. These have been designated *pcaU* and *pobR*, and have been extensively studied. The PcaU and PobR primary structures share 54% amino acid sequence identity. Also, PcaU and PobR have similar DNA-binding sites, metabolic modulators, and physiological function (Kok *et al.*1998).

The *pcaU* gene is divergently transcribed from the *pcaIJFBDKCHG* gene cluster. Protocatechuate binds to PcaU and thereby specifically activates transcription of the *pcaIJFBDKCHG* genes for protocatechuate catabolism. Also, PcaU has been shown to bind to DNA in the *pcaU-pcaI* intergenic region (Gerischer *et al.* 1998).

Divergently transcribed from *pobA* (4-hydroxybenzoate hydrolase) is *pobR*. It has been shown that PobR specifically regulates *pobA* expression in response to 4-hydroxybenzoate. Once it is expressed, PobA catalyses the conversion of 4-hydroxybenzoate to protocatechuate, which thence binds to PcaU, which is the activator of the *pca* genes. Unlike CatM and BenM of the *ben-cat* cluster, transcription regulation exercised by the two proteins PcaU and PobR is highly specific. The specificity of the two regulators is highlighted in that 4-hydroxybenzoate only differs from protocatechuate in that it contains one less oxygen atom (Kok *et al.* 1998).

1.5. The catabolism of naturally occurring aromatic monomers by *Acinetobacter sp. ADP1*.

1.5.1. Pathways for the funnelling of aromatic monomers into the β-ketoadipate pathway in *Acinetobacter* sp. ADP1.

Acinetobacter sp. ADP1 can utilise benzoate as a sole carbon growth source. Adjacent to the *cat* genes in *Acinetobacter* sp. ADP1 are the *benABCD* genes encoding a multicomponent oxygenase for the conversion of benzoate to catechol. The enzyme, benzoate 1,2-dioxygenase, is composed of a hydroxylase component, encoded by *benAB*, and an electron transfer component, encoded by *benC* (Neidle *et al.*1991). The BenD protein is a member of the short-chain alcohol dehydrogenase superfamily and is required for the conversion of benzoate *cis*-dihydrodiol to catechol (Fig. 1.4.1.) (Neidle *et al.* 1992).

ADP1 can also use anthranilate (2-amino benzoate) as a sole carbon growth source. Anthranilate is converted to catechol, allowing further degradation to proceed via the β-ketoadipate pathway (Fig. 1.4.1.). This reaction is catalysed by the gene products of *antABC* genes encoding a multi-component aromatic ring-dihydroxylating enzyme, with homologies to *benABC*, for the conversion of benzoate to catechol. Unlike benzoate dioxygenase (BenABC), anthranilate dioxygenase (AntABC)

catalysed catechol formation without requiring a dehydrogenase (BenD). The *antABC* genes were mapped on the ADP1 chromosome and were found to be unlinked to the two known supraoperonic gene clusters involved in aromatic compound degradation (Bundy *et al.*1998). No other compounds other than benzoate or anthranilate have yet been identified which are channelled through the catechol branch of the β-ketoadipate pathway in ADP1.

Adjacent to the *pca* genes for protocatechuate catabolism, are the *qui* genes for the catabolism of quinate to protocatechuate (Fig. 1.4.1; Fig. 1.5.1.). The *qui* genes are clustered in the order *quiBCXA* directly downstream from the *pca* operon. Catabolism of quinate to protocatechuate requires the consecutive action of quinate dehydrogenase (QuiA), dehydroquinate dehydratase (QuiB), and dehydroshikimate dehydratase (QuiC). The *quiX* gene is thought to encode a porin, but the specific function of this protein has not been established (Elsemore & Ornston, 1995).

Directly downstream of the *qui* genes are *pob* genes. The *pobA* gene encodes a 4-hydroxybenzoate 3-hydroxylase which catalyses the catabolism of *p*-hydroxybenzoate to protocatechuate (Fig. 1.4.1; Fig. 1.5.1.) (Fernandez *et al.*1995).

Downstream of the *pob* genes and transcribed convergently is a set of recently characterised genes designated the *ppa* genes (L.N. Ornston, personal communication). The *ppaABCD* genes encode the enzymes responsible for the growth of *Acinetobacter* sp. ADP1 on ferulate, coumarate, caffeate, chlorogenate and p-hydroxyphenylpropionate (Fig. 1.5.1.). The *ppaA* gene encodes an esterase, *ppaB* encodes an acyl CoA dehydrogenase, and *ppaD* encodes a CoA ligase. Sequence analysis of *ppaC* suggests that it encodes a porin-like protein. Genetic defects in the ligase gene (*ppaD*) blocks catabolism of ferulate, coumarate, caffeate, and p-hydroxyphenylpropionate (Fig. 1.5.1.). Growth with p-hydroxyphenylpropionate but not with the other compounds is prevented in a strain in which the acyl CoA dehydrogenase (*ppaC*) has been disrupted. The metabolism of saturated compounds proceeds by ligation to CoA followed by oxidation of the propionyl CoA side chain.

The esterase (*ppaA*) gene is located between the *qui* genes and the other *ppa* genes. Chlorogenate, an ester formed between quinate and caffeate is a common plant product produced in response to stress and injury. Disruption of the esterase gene renders *Acinetobacter* unable to grow with chlorogenate, but does not alter the ability

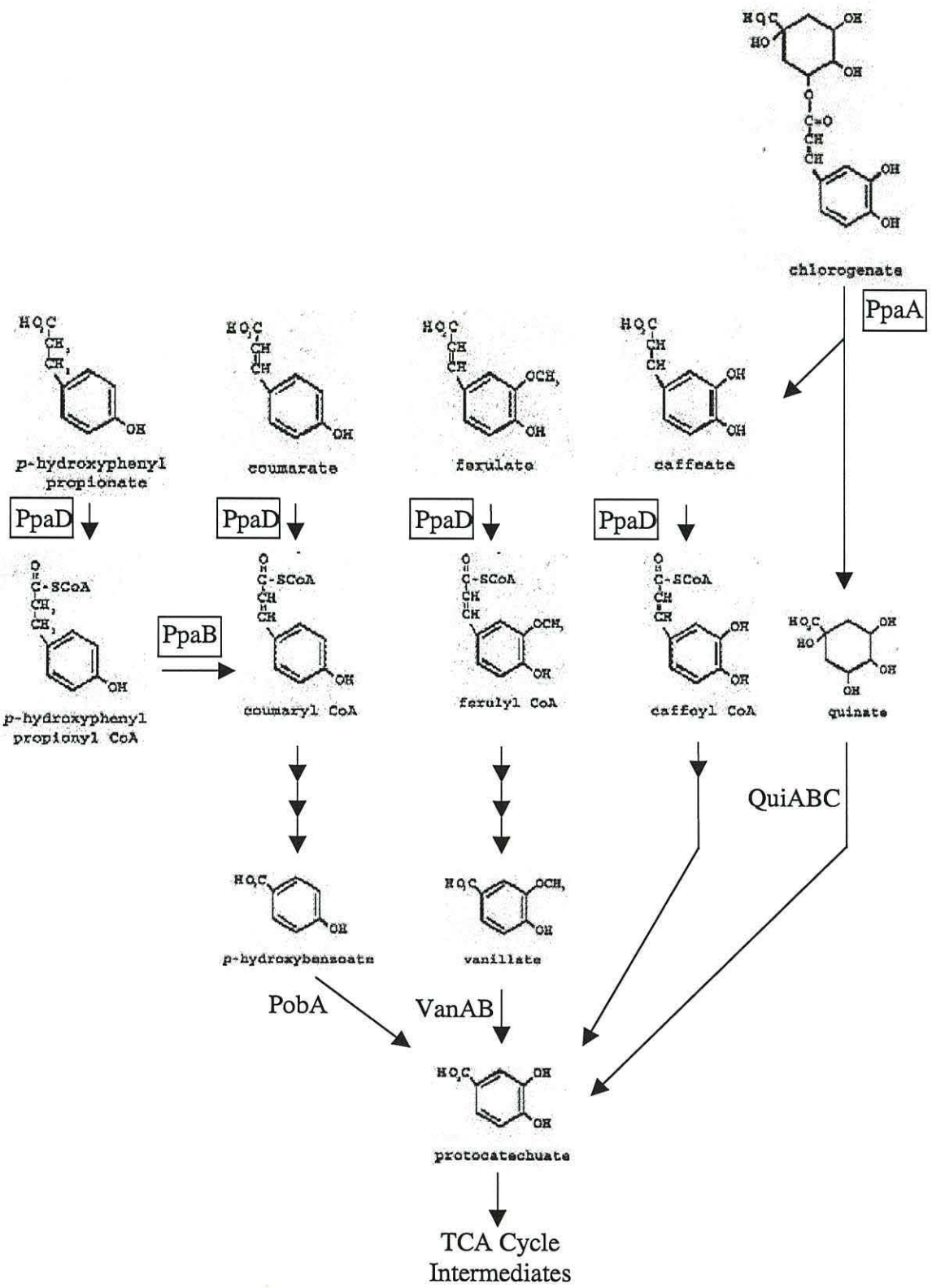


Fig. 1.5.1. The putative functions of the *ppa* genes of *Acinetobacter* sp. ADP1. The *ppa* genes have been shown to be involved in the catabolism of phenylpropenoid catabolism. All the phenylpropenoids shown have been shown to feed into the protocatechuate branch of the β-ketoadipate pathway (L.N. Ornston, Personal communication).

of the bacteria to grown with quinate and caffeate (L.N. Ornston, personal communication) (Fig. 1.5.1.). Further enzymes are required to convert the phenylpropenoid-CoAs to protocatechuate. The characterisation of the missing steps in this pathway is currently being carried out at Ornston's laboratory.

Genes involved in channeling aromatic metabolites into protocatechuate that have an exceptional location have been designated the *van* genes. The *vanAB* genes are positioned in a distant unstable region of the chromosome (Segura *et al.* 1999). VanA and VanB form an oxygenative demethylase that converts vanillate to protocatechuate. The *vanAB* genes are also involved in ferulate catabolism. Ferulate has been shown to be metabolized through a vanillate intermediate in *Acinetobacter* sp. ADP1 (Segura *et al.*1999) (Fig. 1.5.1.).

1.5.2. Aromatic Compound Transporters in *Acinetobacter* sp. ADP1.

Genes located in the *ben-cat* and the *adu-pca-qui-pob-ppa* clusters have been identified as having a role in aromatic compound transport. These genes are members of the large group of transport proteins termed the major facilitator super family of proteins (Marger & Saier, 1993).

The *benK* gene was identified within the *ben-cat* gene cluster. The *benK* gene was found to be expressed in response to the benzoate metabolite, *cis,cis*-muconate. The disruption of *benK* reduced benzoate uptake and impaired the use of benzoate as the carbon source for ADP1 growth (Collier *et al.* 1997).

Another gene, designated *mucK* is involved in muconate transport in ADP1. Insertional inactivation of the *mucK* resulted in the loss of the ability to utilise exogenous muconate. The *mucK* gene, although associated with the benzoate branch of the β-ketoadipate pathway, is close to the *pca-qui-pob* gene cluster (Williams & Shaw, 1997).

Sequence comparisons of PcaK in ADP1 indicate that it is another member of the facilitator super family of proteins. The deduced amino acid sequence of PcaK is 30% identical to BenK from ADP1, and 57% identical to PcaK from *P.putida*. The PcaK from ADP1 has been shown to be involved in the active transport system for both 4-hydroxybenzoate and protocatechuate uptake. Moreover, PcaK has been found

to have overlapping specificity for p-hydroxybenzoate and for protocatechuate with a fourth member of the major facilitator superfamily of transport proteins found in ADP1. This protein is designated VanK, and is coded by *vanK* which is located in the unstable vanillate catabolic genes, which are far removed on the chromosome from the *pca* genes. Inactivation of both PcaK and VanK proteins severely impairs growth with protocatechuate, and the activity of either protein alone can mask the phenotype associated with inactivation of its homologue (D'Argenio *et al.* 1999).

1.5.3. Regulation between branches of the β-ketoadipate pathway.

Using nuclear magnetic resonance spectroscopy techniques, the metabolism of *Acinetobacter* sp. ADP1 has been studied when provided with several carbon sources simultaneously, all degraded via the β-ketoadipate pathway. ADP1 preferentially degraded benzoate, which is degraded via the catechol branch, in preference to *p*-hydroxybenzoate, which is degraded via the protocatechuate branch, when both compounds were present. It was also found that the product of catechol ring cleavage, cis,cis-muconate, inhibited the utilization of 4-hydroxybenzoate in the presence of benzoate (Gaines *et al.* 1996).

Studies of *pcaF* and *pcaK* in *P.putida* revealed that the expression of *pcaK* (gene for 4-hydroxybenzoate transport) is repressed by growth on benzoate when cells were grown on a mixture of 4-hydroxybenzoate and benzoate. Here, regulatory control in the β-ketoadipate pathway appears to extend to the first two steps of 4-hydroxybenzoate degradation, since levels of 4- hydroxybenzoate hydroxylase and protocatechuate 3,4-dioxygenase activities were also depressed by benzoate. These findings indicate that transport plays a role in establishing the preferential degradation of benzoate over 4-hydroxybenzoate in *P.putida* (Nichols & Harwood, 1995). These two sets of results from ADP1 and from *P.putida* demonstrate that there is communication between the two branches of the β-ketoadipate pathway.

1.6. A tight regulation system controls the expression of catabolic genes.

1.6.1. Bacteria Respond to Environmental Stimuli.

Bacteria respond to diverse environmental stimuli such as temperature, pH, and the presence or absence of energy sources to name but a few. Bacteria are able to determine the state of the environment, and then are able to transmit the information intercellularly and thence adapt the physiology of the cell accordingly. This response to the environment is achieved by signal detection and signal relay by proteins that prevent the inefficient production of messenger Ribonucleic acid (mRNA). The control of mRNA production can occur by: 1) suppressing mRNA synthesis from the DNA template in the RNA polymerase-mediated reaction known as transcription; 2) the stability of the mRNA; 3) by post-translational modifications such as the requirement for a co-factor or correct folding into its functional state.

In transcriptional regulation, expression of genes involved in a given process are controlled until their gene products are required. This control can be at a single hit level if the genes are encoded on an operon, or at a multiple gene level by controlling multiple operons or genes in a regulon. By these mechanisms bacteria tightly control modification of cellular processes and adapt their physiology rapidly.

1.6.2. The start of transcription.

Regulation of transcription occurs through the control of the cell's ability to synthesise RNA from a DNA template, in a process called transcription which is catalysed by RNA polymerase. RNA polymerase is made up of a number of subunits of the configuration, $\alpha_2\beta\beta'\sigma$ ($E\sigma$). The core of the polymerase is invariant and has the subunit structure $\alpha_2\beta\beta'$(E). The core polymerase forms a complex with the sigma factor (σ), and it is the σ factor domain of the complex that recognises the DNA promoter (P) that is located upstream of the transcriptional start site. Once bound to the promoter, the $E\sigma$ forms an initial closed (c) complex with the DNA duplex $E\sigma Pc$. Following this, the RNA polymerase undergoes conformational changes, which also results in conformational changes in the DNA to which it is bound. This results in the DNA duplex melting to two single strands, and an open (o) complex is formed ($E\sigma Po$) (Record *et al.* 1996). Nucleoside triphosphate (NTP) then binds to the functional RNA polymerase-promoter complex, followed by the synthesis of the

initial phosphodiester bonds. The RNA polymerase then moves from the promoter as it enters the elongation phase and RNA synthesis.

The σ subunit is bound relatively weakly to the rest of the enzyme (i.e. the core polymerase, E), and following open complex formation the σ protein is released during the initiation process. The core polymerase then walks along the DNA, catalysing phosphodiester bond formation between the growing RNA chain and the next NTP.

Chain termination occurs when the ternary complex of the core RNA polymerase, RNA, and DNA, encounters a secondary structure such as a stem loop containing an inverted repeat. After chain termination and dissociation of the ternary complex the released core can again bind a σ subunit and start a new cycle of RNA synthesis (Burgess *et al.* 1969; Travers & Burgess, 1969).

1.6.3. Sigma factors recognise the promoter sequence.

Control mechanisms may be exerted at any stage of the initiation of transcription. Many bacteria have exploited the start of transcription as a focal point for the control process, and a number of regulatory mechanisms have been uncovered. One mechanism of control involves the different transcriptional factors in the cell. Through these different factors, RNA polymerase can be targeted to promoters with different sequences (Helmann, 1994; Lonetto *et al.* 1992; Record *et al.* 1996). The sigma factor which has been most extensively studied in *E.coli* and other bacteria is known as the sigma 70 (σ^{70} or σ^{A}) and is encoded by the *rpoD* gene in *E.coli*. The σ^{70} recognises and binds to the consensus sequence known as the -10,-35 promoter box, and commonly features in the transcription of housekeeping genes.

Another sigma factor that binds to a similar sequence as the -10,-35 box is σ^{s}, which is involved in the regulation of general stress response genes and stationary phase specific genes in *E.coli* (Nguyen & Burgess, 1997). A set of minor σ factors has also been identified in *Bacillus subtilis* that are involved in the control of its sporulation process (Kroos & Cutting, 1994). Another type of sigma factor is the sigma 54 (σ^{54} or σ^{N}) which is involved in the transcription of specialised metabolic genes. It is structurally different to the σ^{70} and binds to an alternative promoter

sequence known as the -12,-24 box (Merrick, 1993). In this chapter, we will discuss the interactions of both sigma 70 and sigma 54 transcription factors.

1.6.4. Sigma70 recognises the -10, -35 promoter.

The *Escherichia coli* genome encodes several different sigma factors, of which σ^{70} (encoded in *E.coli* by *rpoD* and named for its molecular weight of 70,000 KDa) is by far the most abundant. The subunit is an important determinant of the sequence-specific recognition of promoter DNA. It has been shown that σ^{70} must be in the context of the holoenzyme before it can bind to the DNA promoter region (Dombroski *et al.* 1992). The $E\sigma^{70}$ binds to a consensus -10,-35 box at the promoter. Specifically, four important promoter elements can be identified as playing a role in transcription initiation (Fig. 1.6.1.a). These include two hexamers at positions -10 and -35 upstream from the transcription start site, a 'spacer' stretch of DNA separating the -10 and –35 elements, and a region between -40 and -60 known as the UP element. Different compilations of promoter sequences recognised by $E\sigma$ 70 have established the importance of the promoter -10 and -35 regions (Harley & Reynolds, 1987; Hawley & McClure, 1983). The consensus sequences of the -35 and -10 regions are TTGACA and TATAAT respectively (Fig. 1.6.1.a). The greater the similarity of the -10 and -35 regions to their consensus sequence, the higher the efficiency of the promoter (Record *et al.* 1996). The consensus sequence of the $E\sigma^{70}$ promoter is a near perfect binding site and variations can result in low affinity binding. This is how the rate of transcription of unregulated genes such as transcriptional regulator proteins is expressed at different levels compared to housekeeping genes.

An A+T-rich region (the UP element) between positions -40 and -60 has been recognised in some promoters. This is also an additional determinant of promoter activity. This region is contacted by the α subunit (Rao *et al.* 1994; Ross *et al.* 1993).

A subclass of *E. coli* promoters function without a recognisable -35 region or the involvement of any regulating proteins. Such promoters have an "extended -10 region" with the sequence TGNTATAAT (Dombroski, 1997, Sabelnikov *et al.* 1995; deHaseth *et al.* 1998) (Fig. 1.6.1.a).

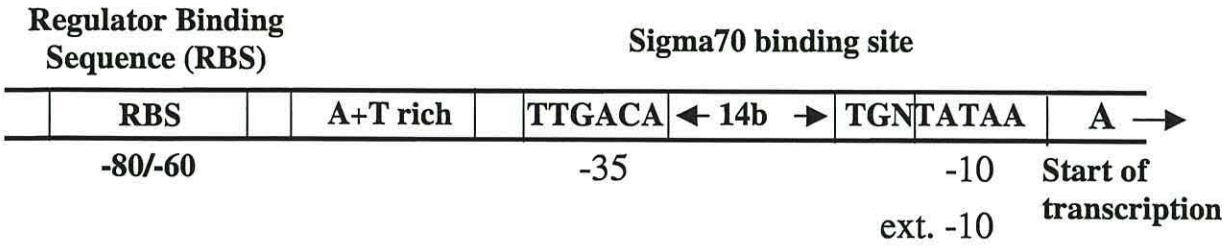


Fig.1.6.1.a. The sigma 70-type promoter region. The start of transcription (+1) is indicated and normally occurs at an adenine residue. Upstream of the start is the -10,-35 conserved region which is the site at which the sigma70 binds. The extended -10 region normally occurs in situations where there is an unconserved -35 region. The Regulator Binding Sequence (RBS) normally occurs between -80 and -60 and is the site at which the transcriptional regulator binds to the DNA (Record *et al.* 1996; deHaseth *et al.* 1998).

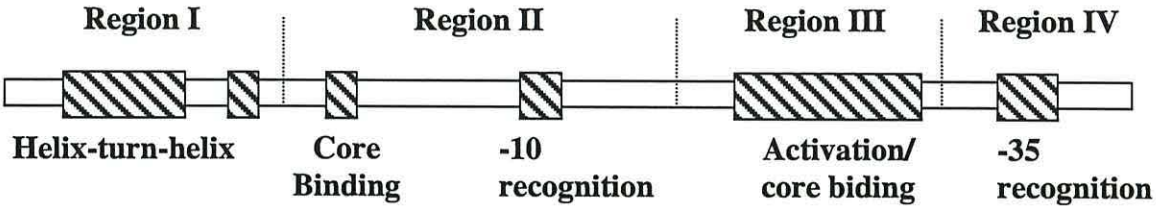


Fig 1.6.1.b. The structure of a typical sigma70 transcription factor. The typical sigma70 transcription factor can be separated into four regions. Region I contains a helix-turn-helix that has been implicated in DNA binding. Region II contains regions that are involved in binding to the core polymerase, and in recognition of the -10 promoter. Region III has been implicated in binding to the core polymerase, and region IV contains a region which has been implicated in the -35 promoter region (Schell, 1993).

The amino acid sequence of the sigma 70 has been determined and has been shown to contain defined regions with specific functions (Fig. 1.6.1.b).

1.6.5. Sigma 54 recognises the -12,-24 type promoter.

σ^{54}-RNA polymerase holoenzyme ($E\sigma^{54}$) is responsible for the expression of genes whose products are involved in diverse metabolic processes, such as dicarboxylic acid transport, nitrogen assimilation and fixation, toluene and xylene catabolism, pilin and flagellin synthesis, and hydrogen metabolism (Merrick, 1993; Morett & Segovia, 1993; Shingler, 1996).

The sigma 54 subunit (encoded by the *rpoN* gene in *E.coli*) enables the core RNA polymerase to recognise and initiate transcription from a distinct class of -12, -24 bacterial promoters that differ considerably from the usual -10,-35 type promoter which is recognised by the $E\sigma^{70}$. $E\sigma^{54}$ binds to the promoter with the consensus sequence at –24, -12, of TGGCAC-N5-TTGCAT. This consensus has an invariant-24 GG motif, and an almost universally conserved -12 GC motif (Fig. 1.6.2.a). Unlike the housekeeping σ^{70}, σ^{54} has been shown to exhibit promoter specific binding in the absence of the core indicating that σ^{54} may have a more active role in recruiting the polymerase to the promoter (Buck & Cannon, 1992). The σ^{54} is able to bind to the promoter alone compared to the σ^{70} type which only binds to the promoter in the form of $E\sigma^{70}$. The affinity of the promoter bound σ^{54} for DNA is greatly increased when complexed with the holoenzyme (Buck & Cannon, 1992). However, following the binding of the $E\sigma^{54}$ to the promoter to form a stable complex, the $E\sigma^{54}$ is isomerisation incompetent and is thus incapable of forming an open complex, unless it is activated by a positive regulator protein. Hence, all -12,-24 type promoters are dependant on transcriptional activator proteins that promote the isomerisation process. It is the transcriptional activator proteins which provide the required ATPase activity for ATP hydrolysis which thus provides the necessary energy for the isomerisation process and open complex formation (Shingler, 1996).

The amino acid sequence of the sigma 54 protein has been deduced and was shown to contain defined regions with specific functions (Fig. 1.6.1.b).

Inverted Repeat	50-200b.p	TGGCAC	CTGC	A→
Upstream Activator Sequence (UAS)	**IHF binding site or Poly A region**	-24	-12	**Start of transcription**
		Sigma54 binding site		

Fig.1.6.2.a. The sigma 54-type promoter region. The start of transcription (+1) is indicated and normally occurs at an adenine residue. Upstream of the the start is the -12,-24 conserved region which is the site at which the sigma54 binds. The Upstream Activator Sequence (UAS) is the region where the regulator protein binds and looping out of the DNA by the IHF binding site or polyA region brings the regulator to activate the holoenzyme (Shingler, 1996).

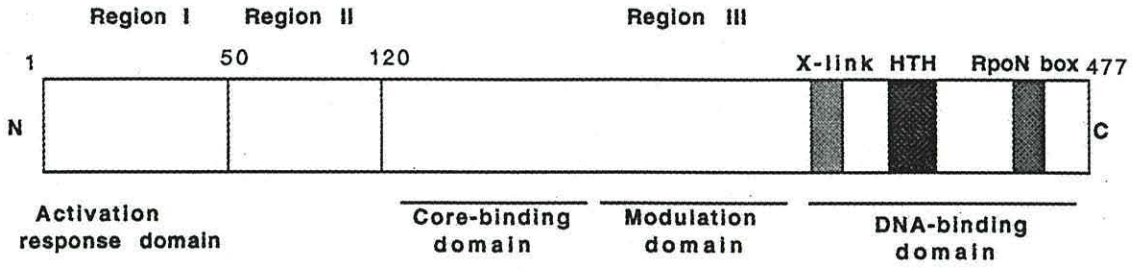


Fig. 1.6.2.b. Structure of a typical sigma54 transcription factor. The protein contains two functionally conserved regions (regions I and III) separated by a region that varies in length among species (Merrick, 1993). Region I (the activation region) is thought to contain a 'leucine patch' which is thought to prevent the holoenzyme from proceeding past the closed complex stage (Wang & Gralla, 1996). Region I has also been implicated in the activator-dependant formation of an open complex. Region III contains a putative DNA-binding helix-turn-helix motif thought to be involved in DNA binding. It also contains the RpoN box which contains the sequence ARRTVAKYRE consensus sequence that is the signature motif of sigma54 proteins. The X-link region has not been designated a function but it is thought to have a function of cross linking to promoter DNA (Wang & Gralla, 1996).

1.7. Regulator proteins control the expression of σ^{70} type promoters.

1.7.1. The LysR family of transcriptional regulators.

Regulator proteins are involved in the control of the expression of the σ^{70} type transcriptional promoters. One family of this type of regulator is known as the LysR-type transcriptional regulators (LTTRs). LTTRs may be the most common types of positive regulators in prokaryotes. They are involved in the regulation of a diverse range of genes and operons in various genera (Schell, 1993). They are nearly all positive regulators, with a few exceptions which are negative regulators, which are responsive to small effector molecules. The LTTRs are usually between 276 to 324 amino acids in length and share >20% identity with other proteins of the same family. There is particular high homology at the amino terminal end (N-terminal). This is derived from a conserved helix-turn-helix motif that is present in the first 60 residues, and is a signature sequence for the identification of LTTRs. Another feature of the LTTR family is that they are normally transcribed divergently from the genes they regulate (Fig. 1.7.1.) (Schell, 1993).

LTTRs are responsive to chemical effectors, but remain bound to the promoter target DNA sequences, located between -80 and -60 from the transcriptional start, regardless of the presence or absence of inducer. Protein-DNA binding in this region overlaps the -10,-35 promoter region and thus prevents the $E\sigma^{70}$ from binding. Effector binding and transcription activation occurs through the relief of the DNA bending caused by LTTRs binding under non-inducing conditions. It has been shown that the LTTRs exist *in vivo* as multimers and bind to the regulator-binding site (RBS) in a multimeric form (Schell, 1993).

Members of the LTTR family have a conserved domain structure (Fig. 1.7.1.). Studies where amino acid substitutions have altered the function of LTTRs have been done to designate a function to each domain (Schell *et al.*1990). The region of greatest sequence homology is the 66 N-terminal residues, and in particular, between residues 23-42, which is the region that contains the helix-turn-helix (HTH) motif that has been implicated in DNA binding (Suzuki & Brenner, 1995). Residues outside of the 66 terminal amino acids show much less similarity and limited significant conservation. Nevertheless, the effector-binding pocket has been identified at the centre of the protein sequence, and another region has been identified as being

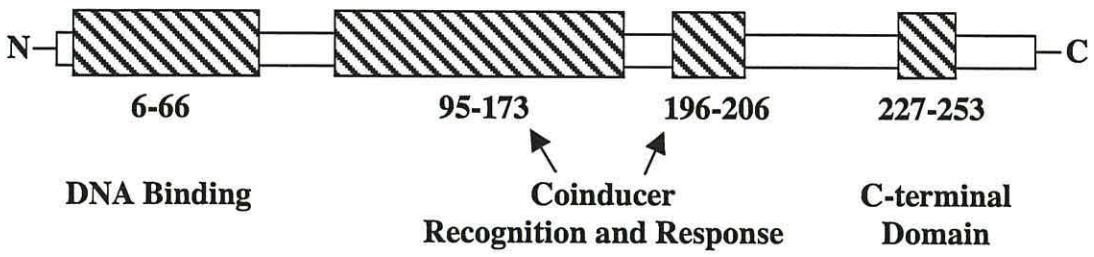


Fig. 1.7.1. Model for domain organisation of a LysR Type Transcriptional Regulator (LTTR). Residue numbers and domains are defined in text. The DNA binding domain contains a helix-turn-helix in the region between residues 20 to 40. Two regions have been found to be critical for coinducer recognition and response, these are from residues 95 to 173, and from residues 196-206. The C-terminal domain contains part of a β-turn- β ligand-binding crevice (Schell, 1993).

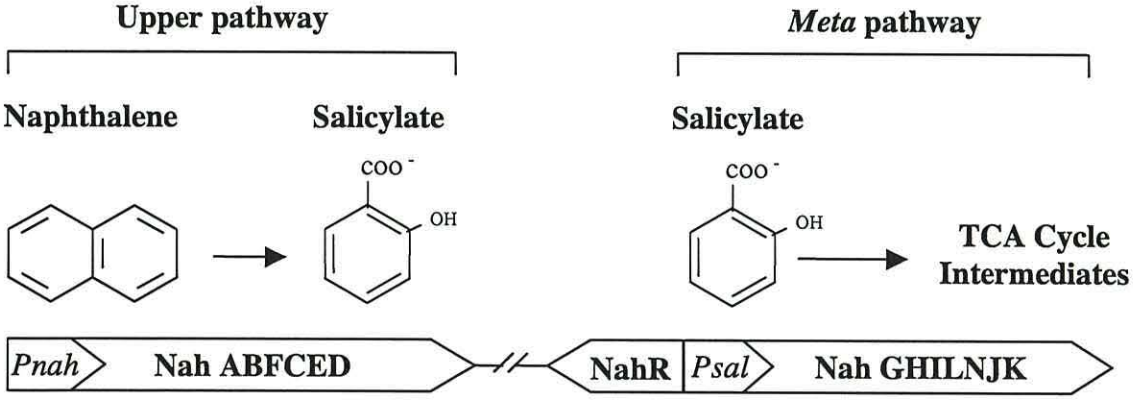


Fig. 1.7.2. The naphthalene degrading genes of *P.putida* NAH7. Map of the genes that are responsible for naphthalene degradation in plasmid NAH7 (83 kbp). Transcription of the genes is indicated by the direction of the arrow. The genes encoding the enzymes of the upper pathway and the *meta* pathway are separated by 7kbp on the plasmid. The NahR protein which is divergently transcribed from the *meta* pathway, has been shown to be responsible for the positive regulation of both upper and *meta* pathways (You *et al.* 1988).

necessary for effector mediated activation, which has been proposed to function as the interdomain hinge for effector mediated conformational change. Another region at the C-terminal end has been identified as functioning in the multimerisation process (Schell, 1993). However, the precise mechanism by which inducer binding activates transcription is largely unknown.

Some experiments exchanging C-terminal positions of LTTRs with those of different coinducer specificity have shown that recognition function lies between residues 122 and 270 leading to the assumption that the C-terminal is involved in coinducer recognition (Hovarth *et al.* 1987; Huang & Schell, 1991).

There are examples of LTTRs regulating aromatic catabolic operons. A group of regulators within the LysR family that includes CatR, CatM, and ClcR has been identified by sequence homology. All these LTTR type regulators regulate aromatic catabolising operons. The CatR protein activates the *catBC* operon of the β-ketoadipate pathway in *P.putida* in response to *cis,cis*-muconate (Parsek *et al.* 1992). The same catabolic alleles of the β-ketoadipate pathway in *Acinetobacter* sp. ADP1 are regulated by CatM in response to *cis,cis*-muconate (Neidle *et al.* 1989). The modified β-ketoadipate pathway for the metabolism of 3-chlorobenzoate is regulated by the ClcR, the third member of this group (Coco *et al.* 1993).

The catabolism of naphthalene by genes encoded within the NAH7 plasmid are regulated by NahR, which is another type of LTTR protein. Naphthalene catabolism will be focussed on in the next section of the introduction.

1.7.2. Naphthalene catabolism by genes encoded within the NAH7 plasmid.

Pathways for the catabolism of aromatic compounds have been studied and characterised in relation to the catabolism of environmental pollutants. An example of such a study is the elucidation of the pathway for the complete catabolism of naphthalene. The pathways found are plasmid encoded, of which examples are the pWW60-1 (Cane & Williams, 1986) and the NAH7 plasmid (Dunn & Gunsalus, 1973). Both plasmids encode all the genes necessary to convert naphthalene to amphibolic metabolites. The catabolism of naphthalene can be divided into an upper pathway that converts naphthalene to salicylate, and a lower pathway that converts salicylate to amphibolic metabolites (Fig. 1.7.2.). The naphthalene pathway genes

have been cloned from NAH7 and were found to exist in two separate operons. The operons are localised in 10kbp segments, separated by a 7kbp regulatory region (You *et al.* 1988).

1.7.3. Regulation of the *nah* operons.

A gene, designated *nahR*, separates the upper and the lower *nah* operons (Fig. 1.7.2.). The NahR protein is a member of the LysR family of σ^{70}-dependent regulatory proteins. It has been determined that NahR activates the expression of both the upper and lower *nah* operons in response to the inducer salicylate. Gel retardation and DNaseI protection assays showed that NahR binds specifically to DNA fragments containing either the upper operon promoter, or the lower operon promoter. The binding of NahR occured in the promoter region between –82 to –47 of both operons, by salicylate independent direct interaction (Schell & Poser, 1989). Further, it was determined that NahR binds to the promoter region by direct interaction with a number of conserved bases which have the consensus TTCA-N_6-TGA located between -75 and -65 of both upper and lower pathway operons (Schell, 1993).

1.8. Proteins of the NtrC family control the expression of σ^{54}-type transcriptional promoters.

1.8.1. The σ^{54}-dependent family of transcriptional activators.

The members of the σ^{54}-dependent family of transcriptional activators only regulate genes that have a σ^{54}-type promoter sequence. Regulator proteins of this family bind to DNA that is a comparatively long distance away from the promoter. This distance is commonly 100 to 200 base pairs upstream of the promoter, and is thus called an upstream activator sequence (UAS) (Fig. 1.6.2.a.) (North *et al.* 1993). The UAS is often referred to as a 'bacterial enhancer sequence' and is made up of an inverted repeat sequence. It has been known for the UAS to be still functional in activation when moved 3kb away. However, the inverted repeat sequence at the UAS still requires positioning on the same face of the DNA helix to facilitate the oligomerisation process of the $E\sigma^{54}$ initiation complex (Wyman *et al.* 1997). Because of this distance, the regulators require looping out of the DNA to achieve close

proximity to the promoter to activate transcription through contacts with $E\sigma^{54}$. The looping out is facilitated by a DNA bending protein known as the Integration Host Factor (IHF) (Hoover *et al.* 1990), or by an intrinsic polyA induced bend (Perez-Martin *et al.* 1994a; Perez-Martin *et al.* 1994b).

1.8.2. Structure of the $E\sigma^{54}$ activator protein.

All of the $E\sigma^{54}$ proteins that bind to -12,-24 type promoters require activation by a regulator protein of the NtrC family. These proteins have a characteristic domain structure (Fig. 1.8.1.). The central activation C-domain provides the ATP binding and hydrolysis activity (Weiss *et al.* 1991). It contacts the σ^{54} subunit of the polymerase holoenzyme during activation (Wang *et al.* 1997). The C-domain is the most conserved domain of this family of proteins. This domain can be subdivided into seven well-conserved sub-regions that are designated C1-C7. The C1 sub-region contains a motif that is similar to the glycine rich phosphate binding loop common in ATP- and GTP- binding proteins. Sub-region C4 contains a motif that is similar to Walker-type b sequences found in ATP- and GTP-binding proteins (Kelly & Hoover, 1999). Region C3 has the consensus ESELFGHxxGAFTGA and is diagnostic for activators of σ^{54}-holoenzyme. The C3 region contacts the σ^{54}-holoenzyme, and also couples ATP hydrolysis to open complex formation (Wang *et al.* 1997).

The D-domain shows less homology among family members compared to the C-domain. This domain contains a putative helix-turn -helix motif that is responsible for the DNA binding between the regulator protein and the UAS. Commonly, the first helix and the turn are conserved among group members, but the second helix in not conserved. This is probably due to the second helix being involved in recognising specific enhancer sequences (Shingler, 1996).

The B-linker region is a tether between the A-domain and the C-domain, assumed to be a flexible loop. Mutation of the B-domain does not affect the protein function (Wooton & Drumond, 1989). However, mutations in the B-domain of the members of the family, which are directly activated by small effector molecules, can affect function (Perez-Martin & de Lorenzo, 1995). Sequence analysis has also suggested the presence of a coiled coil motif within the B-domain (E. O'Neill, Umea University, Sweden, personal communication).

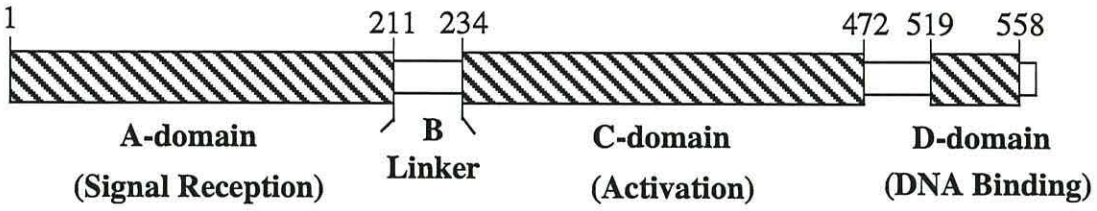


Fig. 1.8.1. Domain structure of sigma54 dependent regulators. Schematic diagram of the functional domains of sigma54 dependent regulators. The A-domain is responsible for signal reception, the C-domain contains the ATPase activity that activates the Eσ-holoenzyme, and the D-domain is responsible for the binding of the protein to the Upstream Activation Sequence(UAS). The amino acid residues presented are a representation of the positions of the domains as they occur in the *P.putida* sigma54 dependent regulator DmpR (Shingler, 1996).

The N-terminal A-domain is responsible for the activation of the regulator protein. This domain has the greatest divergence in protein sequence, which reflects the distinct activation mechanisms of the family. These include activation by phosphorylation, removal of inhibitory protein-protein contacts, and direct binding of small effector molecules (Shingler, 1996). These activation mechanisms of the NtrC family regulator proteins are further introduced in the next three sections.

1.8.3. Activation of the σ^{54}-type regulatory proteins by phosphorylation.

The σ^{54}-type regulatory proteins controlled by phosphorylation are often referred to as the 'response proteins' of two component regulatory systems. In this type of system, a change in the environment is detected via a separate sensory histidine autokinase which is often located in the membrane. Transfer of a phosphate group to a conserved Asp residue within the A-domain serves to activate this type of NtrC family regulator protein. This type of activation is involved in the assimilation and fixation of nitrogen in *Klebsiella pneumoniae*, where a membrane bound autokinase, designated NtrB, controls the phosphorylation of NtrC at residue Asp54, in response to limited nitrogen in the environment (Klose *et al.* 1993). When there is excess nitrogen in the environment, the phosphatase activity of NtrB results in no activation of the nitrogen assimilation genes (Drummond *et al.* 1990). It has been shown that the NtrC protein can bind to DNA without phosphorylation of the Asp54 residue. The phosphorylation step is required for NtrC to be able to bind and hydrolyse ATP, and to form the open $E\sigma^{54}$ complex (Rippe *et al.* 1998).

1.8.4. Activation of the σ^{54}-type regulatory proteins by protein-protein contacts.

These types of σ^{54}-type regulatory protein mediate their regulatory function by protein-protein interactions. An example of this type of regulation is the NifA protein of *Azotobacter vinelandii*. The NifA protein regulates the activation of nitrogen-fixing nitrogenases in this bacterium. Unlike the phosphorylation-controlled NtrC transcriptional enhancer proteins, this class of protein does not have a residue that receives the phosphorylation event. Here, NifA is negatively regulated by protein-protein interactions with NifL, whose genes are both co-transcribed. The *nifLA* operon

is not constitutively expressed like other regulatory proteins, but is itself regulated by NtrC. The NifL protein contains a flavin adenine dinucleotide (FAD) as a prosthetic group. The flavin group is oxidised in the presence of oxygen and under nitrogen limiting conditions. This subsequently regulates the NifA activity by NifL binding to NifA and mediating its inhibition upon NifA by a 1:1 stoichiometric protein-protein contact (Dixon, 1998).

1.8.5. Activation of the σ^{54}-type regulatory proteins by direct effector activation.

The proteins of the direct effector group of σ^{54}-type regulatory proteins have the ability to respond directly to their activating signal, i.e the presence of specific ligands. Examples of this class of proteins are: the FhlA protein from *E.coli* which activates genes involved in formate and hydrogen metabolism (Hopper & Bock, 1995), XylR, which regulates the catabolism of toluene and xylenes in the TOL plasmids of *P.putida* (Abril *et al.* 1989), and DmpR, which regulates phenol catabolism in *P.putida* (Shingler *et al.* 1993). XylR and DmpR are activated in response to distinct aromatic substrates of the catabolic pathways they control. Extensive homology can be seen in the A-domains of the regulators of XylR and DmpR, which are responsive to aromatic ligands. The functions of these two aromatic responsive regulators will be focussed on in the next section of this chapter.

1.9. Aromatic responsive NtrC type transcriptional regulators.

1.9.1. The catabolism of phenol in *P.putida* sp. CF600.

The *dmp* operon of *P.putida* sp. CF600 consists of 15 genes encoding 9 enzymes necessary for the complete mineralisation of phenol and methylated phenols (Shingler *et al.* 1992) (Fig. 1.9.1.). The whole pathway is encoded on one transcriptional unit that is under the regulation of a single σ^{54}-type regulator designated DmpR. The *dmpR* gene is divergently transcribed from the catabolic *dmp* genes, and in addition, the transcription is constitutive at a sufficient level to respond to effector ligands (Shingler *et al.* 1993). DmpR binds to the enhancer-like elements located at the intergenic region between *dmpR* and the operon promoter. In the presence of phenol DmpR initiates transcription of the *dmp* operon (Shingler & Moore, 1994). It has been

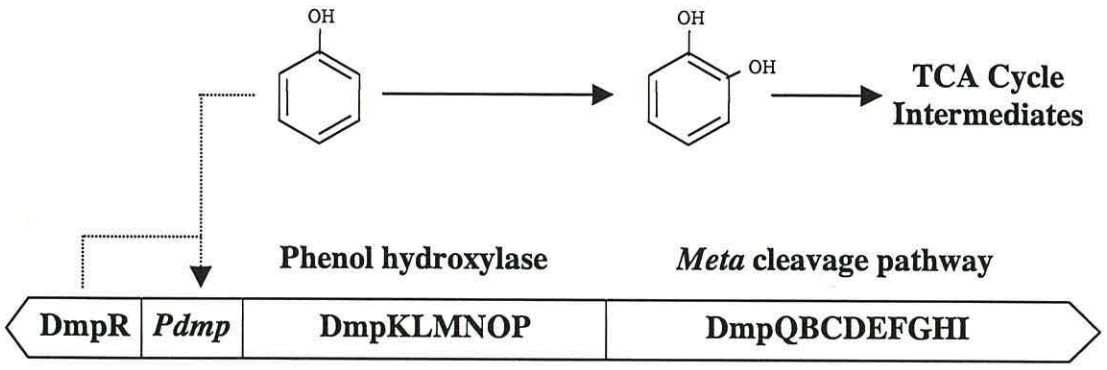


Fig. 1.9.1. Genetic organisation of the plasmid encoded *dmp* genes for the catabolism of phenol in *P.putida* CF600. The phenol hydroxylase catalyses the conversion of phenol to catechol, which is further catabolised to amphibolic intermediates by the *meta* pathway. Both pathways are combined on one transcriptional unit under regulation from the promoter *Pdmp*. The *dmp* genes are regulated by DmpR which is responsive to phenol as a positive effector (Shingler, 1996).

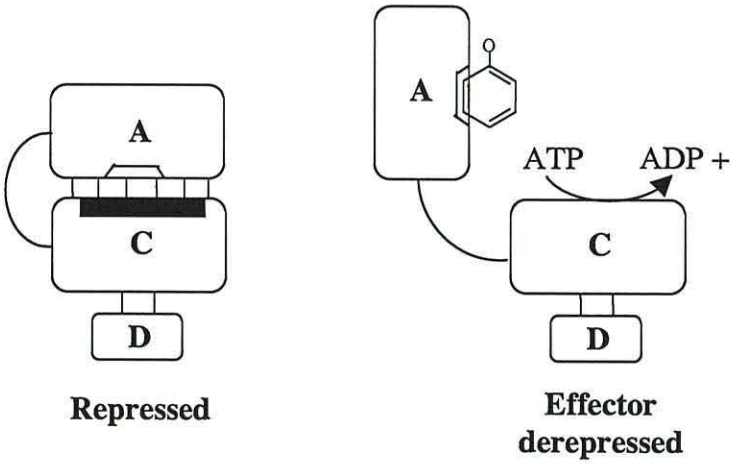


Fig. 1.9.2. General model for activation of DmpR by an aromatic effector. The DmpR is repressed when the A-domain has specific residue-residue interactions with the C-domain. When phenol is present, it binds to the A-domain and the A/C-domain interaction is broken. The C-domain ATPase is then free to bind and hydrolyseATP- which serves to activate the initiation of transcription (O'Neill *et al.* 1998).

shown that the A-domain of DmpR has affinity for phenol. It was also shown that binding occurs at one molecule of phenol per A-domain monomer. In addition, it has also been shown that the A-domain has affinity for the C-domain. When the A-domain is physically uncoupled from the rest of the protein, it results in a system that is unresponsive to phenol. This analysis of DmpR has led to a model in which the A-domain acts as an interdomain repressor of the DmpR C-domain ATPase and transcription promoting activity. The C-domain is repressed until specific aromatic effectors are bound, in this case phenol (Fig. 1.9.2.) (O'Neill *et al.* 1998).

1.9.2. The TOL pathway.

The TOL pathway is responsible for the complete catabolism of toluene or methyl toluenes (xylenes) (Assinder & Williams, 1990). Toluene is first converted to benzoate by the enzymes of the upper TOL pathway. The benzoate is then converted to catechol, which is further catabolised by *meta* cleavage. The genes encoding the enzymes of the TOL pathway are arranged into two distinct operons on the TOL plasmid. The archetypal TOL plasmid was isolated from *P.putida* mt-2 and designated pWW0 (Williams & Murray, 1974).

The upper pathway operon of the TOL pathway is 8kbp in length and contains the genes *xylUWCMABN* whose gene products catalyse the conversion of toluene to benzoate via benzyl alcohol and benzaldehyde (Harayama *et al.* 1989). The lower operon of the TOL pathway is 10kbp in length and contains the genes *xylXYZLTEGFJQKIH* whose gene products catalyse the conversion of benzoate to TCA intermediates by the *meta*-pathway. In contrast to the β-ketoadipate pathway, the lower *meta*-pathway of the TOL plasmid can metabolise alkyl-substituted intermediates (Fig. 1.9.3).

1.9.3. Regulation of the TOL pathway.

TOL pathway gene expression is inducible and regulated by two regulator genes located adjacent to the lower pathway genes, and designated *xylS* and *xylR*. The XylR protein is a member of the NtrC σ^{54}- dependent family of transcriptional regulators, and here it is responsible for the regulation of the expression of the *xylUWCMABN*

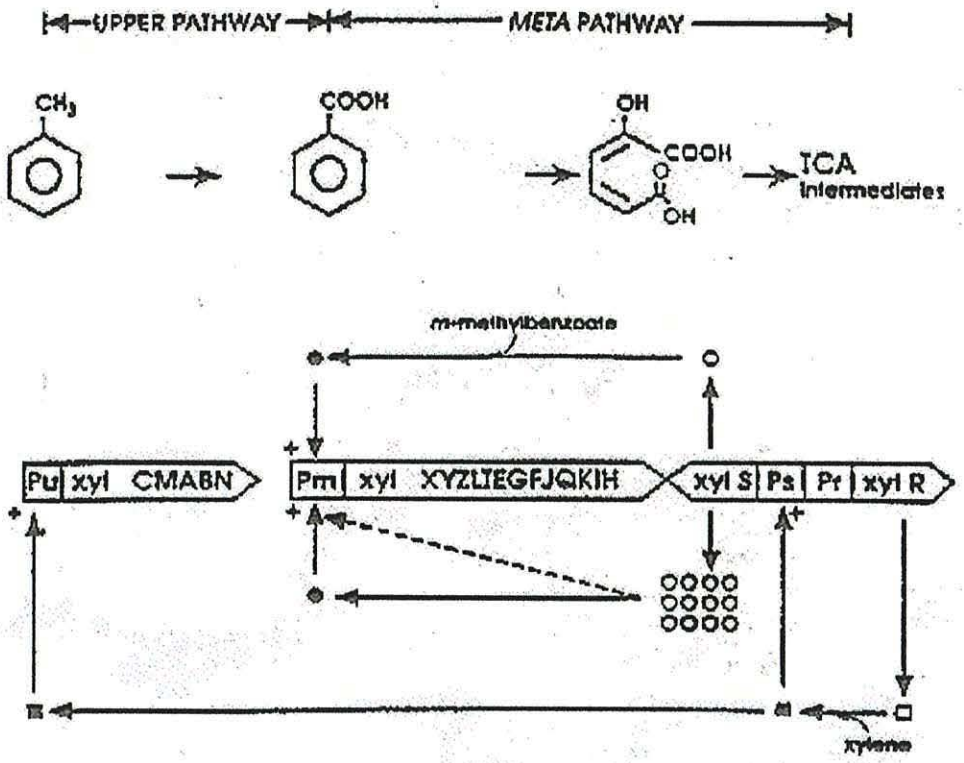


Fig. 1.9.3. Genetic organisation of the *xyl* operons of the pWW0 plasmid from *P.putida* mt-2. The upper pathway operon encodes the enzymes required to convert toluene to benzoate. The *meta* pathway encodes the enzymes for the conversion of benzoate to amphibolic intermediates. The transcription of these operons is regulated by XylS and XylR. The *xylS* gene is expressed at low levels from a sigma70 type promoter (Ps). When the XylS protein is in the presence of benzoates, it induces the *meta* pathway genes at promoter Pm. The *xylR* gene is expressed at low levels from a sigma70 type promoter (Pr). When in the presence of upper pathway substrates, XylR induces expression of the upper pathway from Pu in a process that also requires a sigma54-holoenzyme and IHF. XylR also stimulates high level expression of XylS by acting on a sigma54 type promoter in Ps, which results in benzoate-independent expression of XylS and stimulation of Pm (Gallegos *et al.* 1993).

upper pathway genes (Fig. 1.9.3.). For expression of the lower pathway genes, XylR must first promote the expression of the second regulatory gene, *xylS*, the gene product of which is a member of the σ^{70} dependent AraC family of transcriptional regulators (Gallegos *et al.* 1993). In the presence of inducers, XylR stimulates high-level expression of XylS. Thus, transcription activation by XylS is effector independent when it is present in high concentrations elicited by XylR. Here, there is a regulatory cascade where XylR co-ordinates the regulation of both upper and lower TOL pathway operons (Fig. 1.9.3.).

However, when the cell is presented with substrates that only require the enzymes of the *meta* pathway, XylS itself promotes transcription of the lower pathway operon while present at low levels in the cell. In this case, XylS is dependent on the substrates as effectors, which include benzoate and particularly 3-methylbenzoate (*m*-toluate) which is the best inducer of XylS (Kessler *et al.* 1994; Ramos *et al.* 1987).

The promoter regions of the TOL pathway have been studied and the upstream binding sites of XylR and XylS have been identified. These sites are typical of those of binding site for σ^{54} or σ^{70} dependent regulatory proteins respectively (Marques *et al.* 1994). Thus, there is a complex multi-protein mechanism involved in the regulation of the TOL pathway enzymes.

As with the DmpR protein, experiments have shown that XylR is activated by effector molecules binding to the A-domain of the protein and relieving suppression of ATPase activity (Delgado & Ramos, 1994).

1.10. Benzaldehyde and benzyl alcohol dehydrogenases from *Acinetobacter.*

1.10.1. Homologous genes to *xylB* and *xylC* of the TOL plasmid have been found in *Acinetobacter.*

By analogy, the corresponding pathway in *Acinetobacter* sp. ADP1 to the lower *meta* pathways of the TOL plasmid is the β-ketoadipate pathway. These two pathways serve to funnel benzoate to amphibolic intermediates via *meta* or *ortho* cleavage respectively. Moreover, the TOL plasmid has an upper pathway that funnels toluene to benzoate via benzyl alcohol and benzaldehyde (Assinder & Williams, 1990). No

pathways have yet been found in *Acinetobacter* sp. ADP1 that funnel metabolites to benzoate, although it is known that benzyl alcohol and benzaldehyde are growth substrates of ADP1. Nevertheless, genes have been found in *Acinetobacter calcoaceticus* NCIB 8250 that are homologous to the TOL pathway *xylB* that encodes a benzyl alcohol dehydrogenase (BADH), and *xylC* that encodes a benzaldehyde dehydrogenase (BZDH). The BADH and BZDHII of *A. calcoaceticus* NCBI 8250 are also designated *xylB* and *xylC* respectively and also convert benzyl alcohol to benzoate. The XylB and XylC proteins from both bacteria have been extensively characterised.

In *A. calcoaceticus* NCIB 8250 benzyl alcohol is oxidised to benzoate by chromosomally encoded BADH and by BZDHII; (BZDHI is quite a different enzyme and is involved in the metabolism of mandelate). The two enzymes have been purified and characterised and shown to be NAD-dependent enzymes (Macintosh & Fewson, 1988a; 1988b). The genes encoding the proteins, i.e. *xylB* and *xylC* have been cloned and characterised (Gillooly *et al.* 1998) and in addition, the genes for BADH and BZDH have also been cloned from the TOL plasmid pWW0 (Harayama *et al.* 1989).

1.10.2. The Benzyl alcohol dehydrogenase from *Acinetobacter calcoaceticus* NCIB 8250 and from the TOL plasmid pWWO.

The BADHs from *A. calcoaceticus* NCIB 8250 and from the TOL plasmid pWWO share 54% identity to each other and are both members of the family of zinc-dependent long chain alcohol dehydrogenases. The two bacterial BADHs were found to have a catalytic mechanism that is different from that believed to operate in horse-liver alcohol dehydrogenase (ADH), the archetypal enzyme of this family (Shaw *et al.* 1993; Gillooly *et al.* 1998; Inoue *et al.*1998). In horse liver ADH, the His51 has been proposed to act as a general base that abstracts a proton from the alcohol substrate through a proton-relay system (Eklund *et al.* 1981). The residue that plays the same role as His51 of ADH in BADH of pWW0 is His47. When His47 of the pWW0 BADH was mutated to a Gln (His47Gln), the BADH activity for benzyl alcohol of the mutant was over a hundred times lower. At position 51 of the pWW0 BADH is a valine residue. When a double mutant of His47Gln / Val51His was created, this

double mutant had ten times less activity than the single His47Gln mutant (Inoue *et al.* 1998).

1.10.3. The Benzaldehyde dehydrogenase from *Acinetobacter calcoaceticus* NCIB 8250 and from the TOL plasmid pWW0.

BZDHII from *A. calcoaceticus* NCIB 8250 shares 56% amino acid sequence identity with BZDH encoded by *P. putida* pWW0. Studies on rat liver aldehyde dehydrogenase have shown that residues Cys-302 and Glu-268 are essential for its catalytic activity (Farres *et al.* 1995). These residues are conserved in BZDH II, which suggested that the enzyme is a member of the large family of NAD^+-linked aldehyde dehydrogenases. BZDH II also has 22 of the 23 conserved residues that are invariant in the family, and contains the consensus sequence FTGSTxVG proposed to constitute the NAD^+-binding domain of aldehyde dehydrogenases (Gillooly *et al.* 1998; Hempel *et al.* 1993).

1.10.4. The substrate profiles of Benzyl alcohol dehydrogenase and Benzaldehyde dehydrogenase from *Acinetobacter calcoaceticus* NCIB 8250.

The substrate profile of BADH and BZDH from *A. calcoaceticus* NCIB 8250 has been determined. The BADH was found to be active against a range of substartes. Benzyl alcohol was the best substrate, while *para*- substituted hydroxy- or methoxy-derivatives of benzyl alcohol were better substrates than their *ortho* or *meta* equivalents. BADH also had activity against coniferyl alcohol and cinnamyl alcohol (Chalmers *et al.* 1990).

The most effective substrate for BZDH II was found to be benzaldehyde. Also, it was found that benzaldehydes with a single small substituted group on the *meta* or *para* position were better substrates than any other benzaldehyde derivative. BZDH II but not BADH was found to have esterase activity against 4-nitrophenyl acetate (Macintosh & Fewson, 1988b).

In another study, the BADH and BZDH from *A. calcoaceticus* NCIB 8250 were compared to the BADH and BZDH from another TOL plasmid designated pWW53, which was present in *Pseudomonas putida* MT53 (Assinder & Williams,

1990). Both BADHs and both BZDHs from the two bacteria had similar substrate profiles, showing similar activity against benzyl alcohols and benzaldehydes with a single methyl substitution at any of the three positions, with the exception of the BZDHs which had no activity against 2-methylbenzaldehyde. It has also been found that the BADH and the BZDH from pWW0 have a similar substrate profiles to those of *A. calcoaceticus* NCIB 8250 (Chalmers *et al.* 1990).

1.11. Esterases and Lipases.

1.11.1. Esterases and lipases in Microorganisms.

Bacteria produce esterases and lipases that are lipolytic enzymes that hydrolyse ester bonds in carboxyl esters. Through the action of lipolytic enzymes, organic acids glycerols and other alcohols are liberated from the ester compound. Esterases and lipases belong to the class of ester hydrolases (EC 3.1.1). Specifically, esterases belong to the sub-class (EC 3.1.1.1), which are a sub-class of enzymes that hydrolyse water-soluble carboxyl esters with short fatty acid chains. Lipases belong to the sub-class (EC 3.1.1.3). The enzymes of this sub-class are specifically activated at lipid water/ lipid interfaces (Brzozowski *et al.* 1991). These enzymes hydrolyse water-insoluble substrates, which have long acyl chains.

Analysis of esterases and lipases by sequence alignments and crystallography has revealed their structure. Also, the organisation of the catalytic site residue triad has been found. The active sites are similar to those of serine proteases, which has subsequently lead to lipases and esterases being referred to as 'serine hydrolases'. The basal structure of the hydrolases is composed of alternating α-helices and β-strands that form an α/β-fold domain (Ollis *et al.* 1992). It is this fold domain that contains the catalytic triad that is the active site of the enzyme. The serine proteases also contain a similar catalytic triad as serine hydrolases, composed of a nucleophilic serine residue. The serine residue acts in conjunction with a histidine and an aspartic acid residue (Brady *et al.*1990). The nucleophilic residue in the serine proteases is located in the β-ε-Ser-α motif (Jaeger *et al.* 1994). Alignments of this region have shown that the serine residue invariably is in the consensus sequence Gly-X_1-Ser-X_2-Gly (Brady *et al.* 1990).

The deduced amino acid sequence at the serine active site according to PROSITE (Bairoch *et al.*1997) which searches a protein sequence against a profile database, shows that the signature characteristic of serine lipases is [LIV]-x-[LIVFY]-[LIVMST]-G-[HYWV]-S-x-G-[GSTAC], where residues at a point can be any of the ones stated in the parenthesis.

1.11.2. Esterases and lipases from *Acinetobacter* strains.

Several Acinetobacters have been shown to produce hydrolytic enzymes. One of the earliest reports was from the psychrophilic strain O_{16}, and from the mesophylic strain O_4, which are both soil isolates (Breuil & Kushner, 1975). Production of lipase was found when *A.calcoaceticus* (strain 69-V) was grown on alkenes. This lipase was purified and its location in the cell during growth was studied (Fischer & Kleber, 1987). Lipolytic activity has been detected in *Acinetobacter* species isolated from wool scour (Brahimi-Horn *et al.* 1991). An outer membrane esterase has been detected in the oil degrading strain *A.calcoaceticus* RAG-1 (Alon & Gutnick, 1993; Reddy *et al.* 1989). Esterase activity was demonstrated when *Acinetobacter lwoffi* was grown in synthetic lubricating oil as a sole carbon source (Amund, 1997). A lipase that efficiently hydrolyses oleyl benzoate was found in the culture supernatant of *Acinetobacter* nov, sp. strain KM109, and has been purified and characterised . This enzyme preferentially hydrolyzed p-nitrophenyl benzoate over p-nitrophenyl acetate (Mitsuhashi *et al.* 1999).

1.11.3. Esterases and lipases in *Acinetobacter* sp. ADP1.

The first report of hydrolytic activity to be detected in *Acinetobacter* sp. ADP1, (*A.calcoaceticus* BD413) was cell-bound esterase (Gutnick *et al.* 1991). The first detailed study of an esterase in ADP1 was done on a cloned gene termed *estA* (Kok *et al.* 1993). This gene encodes a 40KDa protein and the amino acid sequence showed homology to other esterases. EstA was found to have activity against short acyl chain *p*-nitrophenol esters (pNP), with the optimal activity against a 4 carbon acyl chain. Deletion of the *estA* gene in the wild type only partially abolished cell bound esterase activity in ADP1. The same group subsequently cloned and characterised an extracellular lipase, termed *lipA* (Kok *et al.* 1995). LipA is a 32KDa protein with high

sequence similarity to lipases which are exported from the cell in a two step process that requires a chaperone. LipA showed activity against long acyl chain *p*-nitrophenol esters with the optimal activity against an 18 carbon acyl chain. A second esterase termed *estB* has been characterised from ADP1 (Geißdörfer *et al.* 1999). The e*stB* is located in the same operon together with *rubA* and *rub*B. The latter two genes are involved in alkane degradation in ADP1. However, it was shown that the functional esterase is not necessary for alkane degradation. EstB showed activity with pNP esters with acyl chain lengths from 2 to 12 carbon atoms, with a preference around 6 and 8 carbon atoms. An additional esterase has been cloned from ADP1. Termed EstC, it has activity for Tween 80, and is secreted via the general secretory pathway (Parche *et al.* 1997)

1.12. The Scope of this Thesis.

The pathways for catabolism of the aromatic substrates that feed into the protocatechuate branch of the β-ketoadipate pathway in *Acinetobacter* sp. ADP1 have been, and are currently being studied by L.N. Ornston's laboratory (L.N. Ornston, personal communication). Moreover, the genes encoding the enzymes of these pathways are located adjacent on the chromosome to the genes encoding the enzymes of the protocatechuate branch of the β-ketoadipate pathway. The chromosomal island of catabolic diversity, encompassing the genes for the catabolism of protocatechuate, extends over 50kbp.

In comparison, sequence data from the benzoate-catechol branch of the β-ketoadipate pathway currently extends to 20kbp, and includes all the genes necessary for the complete catabolism of benzoate, via catechol, to amphibolic intermediates. No pathway has been characterised in ADP1 that funnels aromatic substrates through benzoate. By analogy with the TOL upper pathway from *P.putida*, and more importantly, with the pathway for benzyl alcohol degradation in *Acinetobacter calcoaceticus* NCIB 8250, both these pathways funnel benzyl alcohol to benzoate via benzaldehyde. It has been known for some time that both benzyl alcohol and benzaldehyde are growth substrates for ADP1. The aims of this study will be:

1. To clone and characterise the genes involved in benzyl alcohol catabolism in ADP1.

2. To study the regulation of the genes involved in benzyl alcohol catabolism in ADP1.

3. To widen the current shores of the benzoate-catechol chromosomal catabolic island by cloning contiguous fragments of DNA, upstream of the genes involved in the catabolism of benzoate.

CHAPTER 2

Materials and Methods

2.1. Materials and Equipment.

2.1.1. DNA Manipulation Molecules and Enzymes.

Restriction enzymes, Taq polymerase, T4 DNA polymerase, and T4DNA ligase were supplied by Promega Ltd., Southampton, UK). Lambda *Hind*III and 1-Kb ladder DNA size markers were obtained from Gibco BRL Lifetechnologies Ltd. (Glasgow, UK). Custom oligonucleotide synthesis was done by Gibco BRL or from MWG-Biotech Ltd (Ebersberg, Germany).

2.1.2. Chemicals.

Most chemicals and aromatic substrates were obtained from Sigma-Aldrich Co. or from Lancaster Synthesis Ltd. (Morecombe, Lancashire). 2-Hydroxybenzyl acetate and 4-hydroxybenzyl acetate were gifts from Dr. A. Boyes, Department of Chemistry, University of Wales, Bangor. Growth media components were obtained from LabM (Bury, Lancashire).

2.1.3. Equipment.

A model J2-21 Beckman (Beckman Instruments, High Wycomb, Buckinghamshire) preparative centrifuge were used for large volume culture harvesting, and an MSE Centaur 1 bench-top centrifuge (Duxford, Cambridge) was used for centrifugation of volumes of 1.5 ml or less. A model L8-80M ultracentrifuge with a TY65 fixed angle rotor was used throughout the work. Spectrophotometric analysis was done on a Uvikon 860 and an Uvicon 923 UV-visible spectrophotometer (Kontron Instruments, Watford, Hertfordshire).

2.2. Growth Media.

Acinetobacter sp. ADP1 and *E.coli* were routinely grown in LB media (10g Tryptone, 5g Yeast extract, and 10g NaCl, made up to 1 litre with distilled H_2O). For LB agar plates, 16g (1.5%) of agar (LabM, Bury, England.) was added to 1 litre of LB broth.

Acinetobacter sp. ADP1 was also maintained on minimal medium (MM) plus a carbon substrate. Minimal medium salts contained the following quantities of salt per litre: 1.0g $(NH_4)_2SO_4$; 2.5g KH_2PO_4; 0.1g $MgSO_4.7H_2O$; 0.005g $FeSO_4.7H_2O$, 0.25g nitrilotriacetic acid; 0.55g NaOH, and 1 ml of Bauchop and Elsden solution (Bauchop & Elsden, 1960). The Bauchop and Elsden solution contains the following amounts of salt per litre: 10.75g $MgSO_4$, 4.5g $FeSO_4.7H_2O$, 2.0g $CaCo_3$, 1.44g $ZnSO_4.7H_2O$; 1.12g $MnSO_4.4H_2O$; 0.25g $CuSO_4.5H_2O$; 0.28g $CoSO_4.7H_2O$; 0.06g H_3BO_3, and 51.3ml concentrated HCl.

For MM agar plates, 16g (1.5%) of agar (LabM, Bury, England.) was added to 1 litre of MM liquid. Where appropriate, ampicillin was incorporated at 100μg/ml, kanamycin at 50μg/ml for *Escherichia coli* and at 10μg/ml for *Acinetobacter*.

Carbon substrates were added to the minimal medium at the following concentrations: benzyl acetate, 2-Hydroxybenzyl acetate, 4-hydroxybenzyl acetate, benzyl alcohol, benzaldehyde and 4-hydroxybenzyl alcohol, ethyl salicylate, salicylate, benzoate, and 4-hydroxybenzoate at 2.5 mM and succinate at 10 mM. Volatile aromatic esters were provided as vapour by placing the liquid on the lid of the inverted Petri dish.

2.3. DNA Extraction and Quantification methods.

2.3.1. Preparation of genomic DNA from *Acinetobacter* sp. ADP1.

Genomic DNA was prepared as described by Ausabel *et al.* (1987). 30ml of culture was transferred to a sterile 50ml tube and spun at 7,000 rpm for 10 minutes. The pellet was resuspended in 9.5ml TE pH 8.0 containing 0.5% SDS, 100μg ml^{-1} proteinase K and incubated at 37°C for 1 hour. 1.8 ml of 5.0 M NaCl was added and mixed thoroughly, followed by 1.5 ml of 8% (w/V) hexadecyltrimethyl ammonium bromide (CTAB), 0.6 M NaCl. This mixture was incubated at 65°C for 20 minutes. The lysate was extracted with

an equal volume of chloroform/isoamyl alcohol and then centrifuged at 6,000 rpm for 10 minutes. The aqueous layer (upper layer) was transferred to a clean, sterile tube. 0.6 Volumes of isopropanol were added to the supernatant. The precipitated DNA was transferred to 70% (v/v) ethanol by spooling onto a glass rod. The DNA was washed with 70% (v/v) ethanol, allowed to air-dry and dissolved in 2 ml of TE pH8.0.

2.3.2. Preparation of plasmid DNA from *E.coli* strains .

Plasmid DNA was prepared either by Alkaline Lysis (Sambrook *et al.* 1989) or by QIAprep spin miniprep kits (QIAGEN Ltd., Crawley, West Sussex) following the manufacturer's instructions.

2.3.3. Preparation of purified DNA molecules from Agarose gels.

Band extractions were done using the QIAquick Gel Extraction Kit (QIAGEN Ltd., Crawley, West Sussex).

2.3.4. Quantification of DNA.

DNA concentrations were determined by measurement of absorbance at 260 and 280 nm. An absorbance of 1.0 at 260 nm corresponds to approximately 50µg of dsDNA. An A_{260}/A_{280} ratio of greater than 1.8 was taken to indicate purity (Sambrook *et al.* 1989).

2.4 DNA Amplification and Cloning Methods.

2.4.1. PCR amplification.

PCR amplifications were carried out in 50µl reaction volumes containing 10ng of template DNA, 100 pmol of each primer, 2.5 nmoles of each deoxynucleoside triphosphate, 300 nmoles $MgSO_4$, and 1U Vent polymerase (New England Biolabs) in the reaction buffer supplied by the manufacturer. In some reactions, 200 nmoles $MgCl_2$ and 1 U of Taq polymerase were used in lieu of 300 nmoles $MgSO_4$ and 1 U Vent

polymerase. The mixtures were subjected to a 4 min hot start at 94°C, and then to 30 cycles of 1 min at 94°C, 1 min at 56°C, and 2 min at 74°C.

2.4.2. Preparation of Competent *E.coli* cells.

Highly competent *E.coli* cells were prepared using a mixture of calcium, manganese and rubidium chloride salts.

Buffer I: K acetate (0.294g), RbCl (1.21g), $CaCl_2$ (0.147g), $MnCl_2$ (0.99g). Adjust to pH 5.8 with acetic acid. Make up to 100ml with distilled water. Sterilise by filtration and store at 4°C.

Buffer II: MOPS or PIPES (0.209g), $CaCl_2$ (1.025g), RbCl (0.121g), Glycerol (15ml). Adjust to pH 6.5 with potassium hydroxide. Make up to 100ml with distilled water. Sterilise by filtration and store at 4°C.

A 200ml LB culture was inoculated with overnight (o/n) grown *E. coli*. This culture was incubated to an OD_{600} of 0.4, put on ice for 5 minutes, and then harvested at 4°C. The cell pellet was resuspended in 80 ml of buffer I and incubated on ice for 5 minutes. The cells were then repelleted and resuspended in 4 ml of buffer II. This suspension was left on ice for 15 minutes, before 200µl volumes were subsampled into pre-chilled microfuge tubes which were immediately frozen in liquid nitrogen, and stored at -70°C.

2.4.3. Transformation of DNA into *Acinetobacter* sp.ADP1.

Linear DNA was added to an *Acinetobacter* sp.ADP1 log phase culture in LB broth. The culture was incubated at 37°C overnight where DNA is transformed into *Acinetobacter* sp.ADP1 by its natural transformation system. The transformed culture was then plated on LB agar containing the appropriate antibiotic.

2.4.4. Transformation of Plasmid DNA into *E.coli.*

Transformations were done using *E.coli* XL1-Blue MRF (Stratagene, UK) following the manufacturer's instructions or done using LiCl-prepared competent *E. coli.* No more than 5µl (1µg) of DNA was added to 100µl of competent cells which were incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 50 seconds and returned to the ice for 2 minutes. Approximately 900µl of LB broth was added to the transformation mixture, which was incubated at 37°C for 1 hour. Aliquots of the suspension were plated out on selective media, and incubated o/n at 37°C.

2.4.5. Ligation of DNA.

Ligations were done with ratios of vector to insert ranging from 1:2 to 1:5. The reactions were performed in a volume of 10µl at 16°C o/n consisting of ligase (Promega Ltd) which was added at the recommended 5 units per reaction.

2.4.6. Disruption of a Restriction site by T4 DNA polymerase.

T4 DNA polymerase can be used to fill 5' protruding ends with dNTPs, or for the generation of blunt ends from DNA molecules with 3' overhangs due to its 3' to 5' exonucelase activity. T4 DNA polymerase was used to disrupt the *Kpn*I site in the multi-cloning site of pUC18. The digestion of a DNA molecule with *Kpn*I enzyme produces a DNA molecule with 3' overhangs.

pUC18 (2µg) was digested with *Kpn*I restriction enzyme in a final volume of 50µl for 1 hour at 37°C. Then 10U of T4 DNA polymerase and 100µM of each dNTP was added to the restriction mixture, and incubated at 37°C for 5 minutes. The T4 DNA polymerase reaction was stopped by the addition of 2µl of 0.5M EDTA and heating to 75°C for 10 minutes. The DNA was then ethanol precipitated, dried, and resuspended in 8.5µl dH_2O, 1µl ligase buffer, and 0.5µl T4 DNA ligase. Following overnight ligation, the ligation mixture was digested with *Kpn*I for 1 hour to eliminate any reformed

plasmid. The digestion mixture was transformed into *E.coli* and transformants screened for the loss of the *Kpn*I restriction site in pUC18.

2.4.7. Dephosphorylation of vector DNA.

The use of dephosphorylated vector favours the formation of recombinant molecules as opposed to religated vector DNA. Normally, 1µg of vector DNA was digested with restriction enzyme. Then, 0.01U of alkaline phosphatase (Promega), 10µl of 10X buffer (supplied) was added to the DNA and made up to a total volume of 100µl with water. The mixture was incubated at 37°C for 30 min. A further 0.01u of alkaline phosphatase was added to the DNA mixture and incubated for a further 30 min. The enzyme was then inactivated by the addition of 2µl of 0.5M EDTA, and incubated at 65°C for 20 min. The digested DNA was extracted with phenol/chloroform/isoamyl alcohol, and precipitated at -70°C with 300 mM sodium acetate (pH5.2) and 2 volumes of cold ethanol. The precipitate was centrifuged at 15 000 rpm, before being washed with cold 70% ethanol, and resuspended in 10 mM TRIS (pH 8.0).

2.4.8. Phenol/chloroform extraction of DNA.

Phenol/chloroform was used to remove proteins from DNA preparations. An equal volume of phenol was added to the DNA solution and mixed to form an emulsion. The emulsion was centrifuged at 15 000 rpm for 3 min. The aqueous (top) layer was removed and added to an equal volume of chloroform/isoamyl alcohol (24:1), and mixed to an emulsion. The emulsion was centrifuged at 15 000 rpm, and the aqueous (top) layer removed. The extracted DNA was precipitated at -70°C with 300 mM sodium acetate (pH5.2) and 2 volumes of cold ethanol. The precipitate was centrifuged at 15 000 rpm, before being washed with cold 70% ethanol, and resuspended in 10 mM TRIS (pH 8.0).

2.5. Southern blotting and DNA-DNA hybridisation.

Southern blots were produced using Zeta-probe GT Blotting Membrane (Bio-Rad, Hemel Hempstead, UK). The DNA labelling was performed with the ECL (Enhanced Chemiluminesence) Direct Nucleic Acid Labelling System from Amersham Inernational plc. (Little Chalfont, Buckighamshire). Capillary transfer of DNA from gel to membrane was done by immersing the gel in 0.25M hydrochloric acid for 10 minutes with gentle agitation to depurinate the DNA. Then the gel was denatured in 0.5M NaOH/1M NaCl for 30 minutes before rinsing in distilled water. The gel was then placed on three 3MM chromatography paper, which was in contact with (but not immersed in) a tray of 0.4M NaOH. The gel and tray were covered with Clingfilm and a hole was cut to expose the area of the gel to be blotted. The blotting membrane was soaked in 0.4M NaOH and placed over the gel and any air bubbles removed. Three 3MM chromatography papers were placed over the membrane, and then an absorbent pad was placed on the chromatography papers. A weight was placed on the pad, and transfer was allowed to occur for 2 hours for plasmid DNA or o/n for genomic DNA. Following transfer, the membrane was washed in 2X SSC for 5 minutes, and then baked at 80°C for 2 hours.

2.6. Nucleotide Sequence Determination, and Analysis.

2.6.1. Nucleotide Sequencing.

DNA sequences were determined by primer walking of fragments cloned in pUC18. In some cases, sequencing primers (Promega) were used that recognise the pUC18 cloning vector. Sequencing was carried out by Alta Biosciences (University of Birmingham, Birmingham, U.K.), and by MWG-Biotech Ltd (Ebersberg, Germany). Manual sequencing was done using ^{T7}Sequencing kits obtained from Pharmacia Biotech (St. Albans, Hertfordshire) according to instructor's manual.

2.6.2. Nucleotide Sequence Analysis.

Searches of the GenBank database were carried out by the BLASTN and BLASTP programs from the National Center for Biotechnology Information, Bethesda, Md.

(Altschul *et al.* 1990). Sequence data were aligned and edited by using the Lasergene software package (DNAStar, Inc., Madison ,Wis.). Amino acid sequence alignments were performed by ClustalW (PAM350 matrix) (Thompson *et al.* 1994). The Wisconsin Genetics Computer Group programs were also used (Devereaux *et al.* 1984).

2.7. Electrophoresis.

2.7.1. DNA electrophoresis.

DNA molecules were separated and visualised by agarose gel electrophoresis. 0.5X TBE buffer (45mM Tris-HCl, 45mM Boric acid and 1mM EDTA) was used to prepare and submerge the gel. 1X TAE buffer (40mM Tris-acetate, 1mMEDTA) was used for the separation of DNA for Southern blotting. Ethidium bromide was added to the agarose prior to casting at a final concentration of 0.4μg/ml.

2.7.2. Protein electrophoresis.

Proteins were separated and visualised by a discontinuous 12% sodium dodecyl sulphate (SDS) polyacrylamide electrophoresis gel using a Mini-Protean II electrophoresis cell (Bio-Rad Laboratories, Ltd., U.K.). The separating polyacrylamide gel was constituted of 0.375M Tris-HCl pH8.8 buffer, and 0.1% SDS. The stacking gel was constituted of 4% polyacrylamide, 0.125M Tris-HCl pH6.8 buffer, and 0.1% SDS. The gel was submerged in electrode buffer (Tris base 3g/l, Glycine 14.4g/l, SDS 1g/l) pH8.3.

Protein samples were denatured by heating to 95°C for 5 minutes in Sample buffer (62mM Tris-HCl pH6.8, 1.25% SDS, 0.05% β-mercaptoethanol, 10% glycerol, 0.05% Bromophenol blue). Gels were run at 200V until the dye front had emigrated out of the gel and into the lower buffer tank. Low Range molecular weight markers (Bio-Rad Laboratories, Ltd., U.K.) were used as standards. Gels were stained with Coomassie Blue dye for 1 hour and destained overnight in 20% industrial methylated spirits (IMS) and 10% acetic acid.

2.8. RNA manipulation methods.

2.8.1. Isolation of Total RNA from *Acinetobacter* sp. ADP1.

Cells were grown on minimal medium plus a carbon source until they reached a cell density of about 10^8 cells/ml. Total RNA was prepared from 10 ml of the culture with RNeasy Mini columns (Qiagen) following the manufacturer's instructions. The RNA was eluted from the column with 50 µl of water.

2.8.2. Reverse Transcriptase (RT)-PCR.

Total RNA isolated from *Acinetobacter* sp.ADP1 was treated with DNase I to remove any genomic DNA contamination by incubation with 1 U of RNase-free DNase (Promega) and 1 U of RNasin (Promega) in 40 mM Tris-HCl (pH 7.9) containing NaCl (10 mM), $CaCl_2$ (10 mM) and $MgSO_4$ (6 mM) for 30 min at 37°C. The RNA was cleaned by passage through an RNeasy Mini column prior to use in Reverse Transcriptase PCR (RT-PCR). RT-PCR was carried out with the Access RT-PCR kit (Promega). PCR amplifications were carried out in a 50µl volume containing 0.5µg of total RNA, 50 pmol of each primer, 50 µM of each dNTP, 1 mM $MgSO_4$, 5 U of avian myeloblastosis virus reverse transcriptase, and 5 U of *Tfl* DNA polymerase in the reaction buffer supplied by the manufacturer. After reverse transcription at 48°C for 1 h, the reaction mixtures were heated to 94°C for 2 min and given 40 cycles of 30 s at 94°C, 1 min at 55°C, and 2 min at 68°C. Negative control reactions to eliminate the possibility of amplifying residual genomic DNA were performed in the same way, except that the reverse transcriptase was omitted from the reaction mixtures.

2.8.3. Primer Extension.

The start point of the mRNA transcript may be determined by designing an oligonucleotide which anneals to the mRNA and which can be extended to the 5' end of the transcript. By labelling the oligonucleotide 5' carbon with ^{32}P the extended product

can be visualised and its size calculated by running it on a sequence gel in adjacent lanes to sequencing reactions with the same oligonucleotide.

A primer is designed that is approximately 100 bp away from the hypothesised transcriptional start site. The primer is labelled as follows:

100μM Primer (0.5μl), 10X Kinase buffer (2.0μl), DEPC-treated H_2O (15.5μl), γ^{32}P-ATP (1.0μl), T4 polynucleotide kinase (1.0μl). The mixture was incubated at 37°C for 30 min and then inactivated by incubating at 68°C for 10 min.

RNA was isolated from *Acinetobacter* sp. ADP1 cells as described above. The RNA isolated was treated with DNase to remove any DNA that may cause mispriming during the formation of cDNA.

The first step in the synthesis of the cDNA is annealing of the labelled primer. The primer is extended using reverse transcriptase, and the RNA digested leaving labelled DNA. 1μl of labelled primer was added to 9μl of Dnase-treated RNA. The solution was precipitated by 1μl 3M sodium acetate (pH7.0) and 25μl ethanol, incubated at -70°C for 10 minutes, and spun at 4°C for 10 minutes. Following drying of the precipitate, the pellet was resuspended in 30μl hybridisation buffer (20mM HEPES, 0.4M NaCl, 80% formamide, made in DEPC treated H_2O). The pellet was vortexed vigorously and heated at 50°C for 5 minutes to be fully dissolved. The dissolved pellet was incubated at 75°C for 15 minutes, and then at 45°C for 1 hour to hybridise the primer to the RNA. The DNA-RNA hybrids were precipitated by addition of 75μl ethanol, incubated at -70°C for 20 minutes and dried in a vacuum. The pellet was dissolved in 31 μl DEPC treated H_2O. Then the extension mix was added which contains:

5X Reverse transcriptase buffer (10μl); 50mM DTT (1μl); Reverse Transcripatase (25U/μl)(2.5μl); 10mM dNTPs (5.0μl); RNAsin (40U/μl) (0.6μl)

and incubated at 37°C for 1 hour. The mixture was then heated to 72°C for 10 minutes to inactivate the reverse transcriptase. Then 1μl of 10mg/ml RNAase was added and incubated at 37°C for 30 minutes. The resulting cDNA was precipitated by addition of 5μl 4M ammonium acetate (pH4.8) and 125μl ethanol and incubated at -70°C for 20

minutes. The precipitate was centrifuged at 4°C, and the resulting pellet washed in 70% ethanol and respun for 5 minutes. The pellet was dried in a vacuum and resuspended in 5µl of loading buffer (80% formamide, 10mM EDTA (pH8.0), 1mg/ml xylene cyanol FF, 1mg/ml bromophenol blue). The cDNA was then run on a 6% sequencing gel, along with a DNA sequencing lane for estimating the size of the transcript.

2.9. Biochemical Enzyme Assays of AreABC proteins.

2.9.1. Preparation of cell extracts.

Cells were harvested by centrifugation, washed with 100 mM phosphate buffer (pH 7.4) and stored as pellets at -20°C. Cell extracts were prepared by suspending frozen pellets in ice cold 100 mM phosphate buffer pH 7.4, disrupting by a pass through a French pressure cell (SLM Instruments, Inc., Urbana, Ill.), and centrifuging at 120,000 x g for 30 min at 4°C. The supernatant was stored frozen at -20°C as 1 ml aliquots.

2.9.2. Benzaldehyde dehydrogenase (BZDH) Assays.

Benzaldehyde dehydrogenase (BZDH) activity was measured in 3 ml assays containing 100 mM Bicine pH 9.5, 2 mM NAD^+, 100 µM substrate. The reaction was initiated with 10 µl of cell extract and the production of NADH was followed at 340 nm. All three hydroxybenzaldehydes absorb light at 340 nm with the following estimated molar extinction coefficients: 2-hydroxybenzaldehyde (5000 mol^{-1} cm^{-1}), 3-hydroxybenzaldehyde (1500 $mol^{-1}cm^{-1}$), 4-hydroxybenzaldehyde (18400 $mol^{-1}cm^{-1}$). The difference between these values and the molar extinction coefficient of NADH at 340 nm (6220 $mol^{-1}cm^{-1}$) was used to calculate activity with these substrates.

2.9.3. Benzyl alcohol dehydrogenase (BADH) Assays.

Benzyl alcohol dehydrogenase (BADH) activity was measured in a linked assay in 3 ml containing 100 mM CAPS pH 10.5, 2 mM NAD^+, 200 µM substrate and 10 µl of undiluted extract of *E. coli* (pADPW29) overexpressing BZDH. The reaction was

initiated with 10 µl of enzyme extract, and the production of NADH was followed spectrophotometrically at 340 nm. The BZDH was added in excess such that the activity of BADH was rate limiting and that the rate of reaction was linear with respect to the amount of BADH added. Calculated rates took into account that two NADH are produced per benzyl alcohol molecule oxidised except in the case of 2-methylbenzyl alcohol since 2-methylbenzaldehyde is not further oxidised by BZDH. A small correction was made to the rate of 2- and 3-hydroxybenzyl alcohol oxidation to allow for the absorbances of 2-hydroxybenzoate (molar extinction coefficient 40 $mol^{-1}cm^{-1}$) and 3-hydroxybenzoate (250 $mol^{-1}cm^{-1}$) at 340 nm.

2.9.4. 4-Nitrophenyl Ester Assays for Esterase Activity.

Benzyl esterase was assayed with 4-nitrophenyl ester substrates in 1 ml assay mixtures containing 980 µl of 100 mM phosphate buffer pH 8, 10 µl containing varying amounts of 4-nitrophenyl ester in methanol, and the reaction was initiated by the addition of 10 µl of enzyme.

2.9.5. Linked Assay for Benzyl Esterase Activity.

The activity of the esterase AreA with benzyl aromatic ester substrates was determined in a linked assay. The rate of increase of absorbance at 340 nm was measured in 1 ml assay mixtures containing 100 mM phosphate buffer, pH8, 2 mM NAD^+, varying amounts of substrate, and 10 µl of BZDH and BADH from extracts of *E. coli* containing pADPW29 and pADPW30 respectively. The reaction was initiated by the addition of esterase. A linear response of rate to added esterase verified that the esterase-catalysed reaction was the rate-limiting step. The assay produces two NADH per one molecule of benzyl alcohol produced. This is taken into account when determining K_m and V_{max} values.

2.10. Biochemical assays of the SalAE proteins.

2.10.1. Transformation of metabolites.

Transformations of ethyl salicylate into salicylate, and of salicylate to catechol and *cis,cis*-muconate by cell extracts were monitored spectrophotometrically. The measuring cell contained 100µM substrate, 100µM Tris pH7.5 and 10µl of cell extract while the reference cell contained only buffer and enzyme.

2.10.2. Salicylate Hydroxylase assay.

Salicylate hydroxylase (SalA) activity was measured in 3 ml assays containing 50 mM Tris pH 7.5, 100µM NADH, and 100 µM salicylate. The reaction was initiated with 20 µl of cell extract and the rate of oxidation of NADH was determined at 340 nm (6220 $mol^{-1}cm^{-1}$).

2.10.3. Salicylate Esterase Assay.

The activity of the esterase against ethyl salicylate as substrate was determined in a linked assay. Ethyl salicylate was redistilled under reduced pressure to remove small amounts of contaminating ethanol. The rate of increase of absorbance at 340 nm was measured in 1 ml assay mixtures containing 100 mM phosphate buffer, pH8, 2 mM NAD^+, 100µM of substrate, and 10 units of yeast alcohol dehydrogenase (Sigma-Aldrich Co). The reaction was initiated by the addition of esterase. A linear response of rate to added esterase verified that the esterase-catalysed reaction was the rate-limiting step. The assay produced one mole of NADH per mole of ethyl salicylate utilised.

2.11. β-galactosidase assay.

The β-galactosidase activity resulting from the expression of an orf::*lacZ* fusion can be used to study the induction of a gene by effector molecules. This activity was determined by the standard method (Miller, 1972).

Five millilitres of culture was grown overnight in LB broth with or without the presence of inducers. The OD_{600} of the cell cultures were measured. Fifty and 100µl of cell cultures were added to 950 and 900µl of assay buffer respectively. The assay buffer consists of $Na_2HPO_4.7H_2O$ (60mM), $NaH_2PO_4.H_2O$ (40mM), KCl (10mM), $MgSO_4.7H_2O$ (1mM), and β-mercaptoethanol (50mM) (added to the buffer immediately before use). Chloroform (50µl) was added to each solution and the sample vortexed to lyse the cells. The sample was then equilibrated at 30°C for 10 minutes. Then 200µl of 4mg/ml ONPG (*o*-nitrophenyl β-D-galactopyranoside) was added to the samples at regular intervals and vortexed. β-galactosidase activity produces a yellow colour which develops during this stage of the procedure. Colour development was stopped after 5 or 10 minutes incubation by adding 500µl of Na_2CO_3 (1M) at the same regular intervals as the ONPG was added. The samples were centrifuged for 3 minutes to pellet the debris at the chloroform interface. This eliminates the need to take a reading at Abs_{550} to correct for cell debris. The Abs_{420} of the reaction mix was measured and the β-galactosidase activity is then calculated by the formula:

$$\text{Miller Units} = 1000 \times \frac{Abs_{420}}{t \times V \times Abs_{600}}$$

where: Abs_{600} is the cell density before assay, Abs_{420} is the reading from the reaction, t is time of incubation (min), V is the volume of culture added (ml).

CHAPTER 3

Characterisation of the DNA adjacent to the *ben* genes in *Acinetobacter* sp. ADP1

3.1. Introduction.

In *Acinetobacter* sp. ADP1, the chromosomal *cat* and *ben* genes form a distinct supraoperonic cluster of genes involved in the conversion of benzoate into central metabolites. The *cat* genes are involved in the conversion of catechol to central metabolites, and the adjacent *ben* genes are involved in the conversion of benzoate to catechol (Neidle *et al.* 1987). This chapter will present the work undertaken to clone and characterise the DNA adjacent to the *ben* genes, and the investigations carried out to determine the function of the open reading frames present.

3.2. Cloning and identification of the genes upstream of the *ben* genes in *Acinetobacter* sp. ADP1.

3.2.1. Cloning of *Acinetobacter* sp. ADP1 DNA.

The plasmid pBAC98 was a gift from Ellen L. Neidle, Department of Microbiology, University of Georgia, Athens, Georgia. The plasmid was constructed as a part of the study to map the *Acinetobacter* sp. ADP1 chromosome (Gralton *et al.* 1997). It was cloned as follows. The DNA adjacent to the *ben-cat* cluster was isolated by using the kanamycin chromosomal drug-resistance cassette in *benM* of strain ISA36 (Table 3.2.1) (Fig. 3.2.1.). Chromosomal DNA from this strain was used to make a plasmid library in pUC19 in *E.coli* and plasmid pBAC68 (*Bgl*II to *Nsi*I in Fig.3.2.1.) was isolated by the selection for the acquisition of kanamycin resistance. A pBAC68 subclone, pBAC78 (*Bgl*II to *Kpn*I in Fig. 3.2.1.), was disrupted by a Km^r cassette to form pBAC87. Using the natural transformation system of ADP1, the corresponding chromosomal region was replaced by the disrupted *Acinetobacter* DNA of pBAC87, thereby generating ACN73 (Table 3.2.1.). A chromosomal fragment from ACN73 (Table 3.2.1) was similarly isolated on pBAC98 (*Nsi*I to *Kpn*I in Fig. 3.2.1.).

Nucleotide sequencing of pBAC68 revealed an open reading frame *benP* (Fig.3.2.1.) whose predicted protein product was found to be similar to various porins, some of which are associated with genes for aromatic compound degradation (Jones *et al.* 1999). The close proximity of *benP* to *benK*, with only 37 nucleotides separating the predicted coding regions, indicates that they may be co-transcribed. The possibility that *benP*, like *benK*, is part of a regulon that metabolises aromatic compounds through the β-ketoadipate pathway (Collier *et al.* 1997), has prompted further investigation of *benP* in the Neidle laboratory.

Further analysis of the pBAC68 sequence data found an open reading frame, designated *areR*, upstream of, and divergently transcribed from *benP*. The open reading frame is homologous to genes encoding NtrC-type transcriptional activators (Morett & Segovia, 1993; Osuna *et al.* 1997). Sequence analyses were consistent with the possibility that the *areR*-encoded protein acts in a σ54-dependent fashion to control the expression of the genes immediately downstream (Fig. 3.2.1.). The *areR* gene function will be studied in detail in Chapter 6 of this thesis.

3.2.2. Identification of the *areCB* genes.

The nucleotide sequence of pBAC98 was determined by primer walking by Alta Biosciences (University of Birmingham, Birmingham, U.K.). Sequence analysis revealed two full open reading frames and a partial open reading frame located downstream of *areR* (Fig. 3.2.1.). The two full open reading frames have predicted protein products homologous to bacterial dehydrogenases active against benzaldehyde and benzyl alcohol respectively. The genes have been designated *areCB* in the order of their transcription. From GenBank database searches the deduced amino acid sequence of AreC has 88% similarity and 77% identity to XylC, benzaldehyde dehydrogenase II from *A. calcoaceticus* NCIB8250 (Gillooly *et al.* 1998) (Table 3.2.2.), and to a lesser extent has 64% similarity, and 44% identity to the benzaldehyde dehydrogenase encoded by the *xylC* gene on the *Pseudomonas* TOL plasmid pWW0 (Inoue *et al.* 1998) (Table 3.2.2.). Similarly, the deduced AreB amino acid sequence has 91% similarity and 82% identity to XylB, benzyl alcohol dehydrogenase from *A.calcoaceticus* NCIB8250 (Gillooly *et al.*

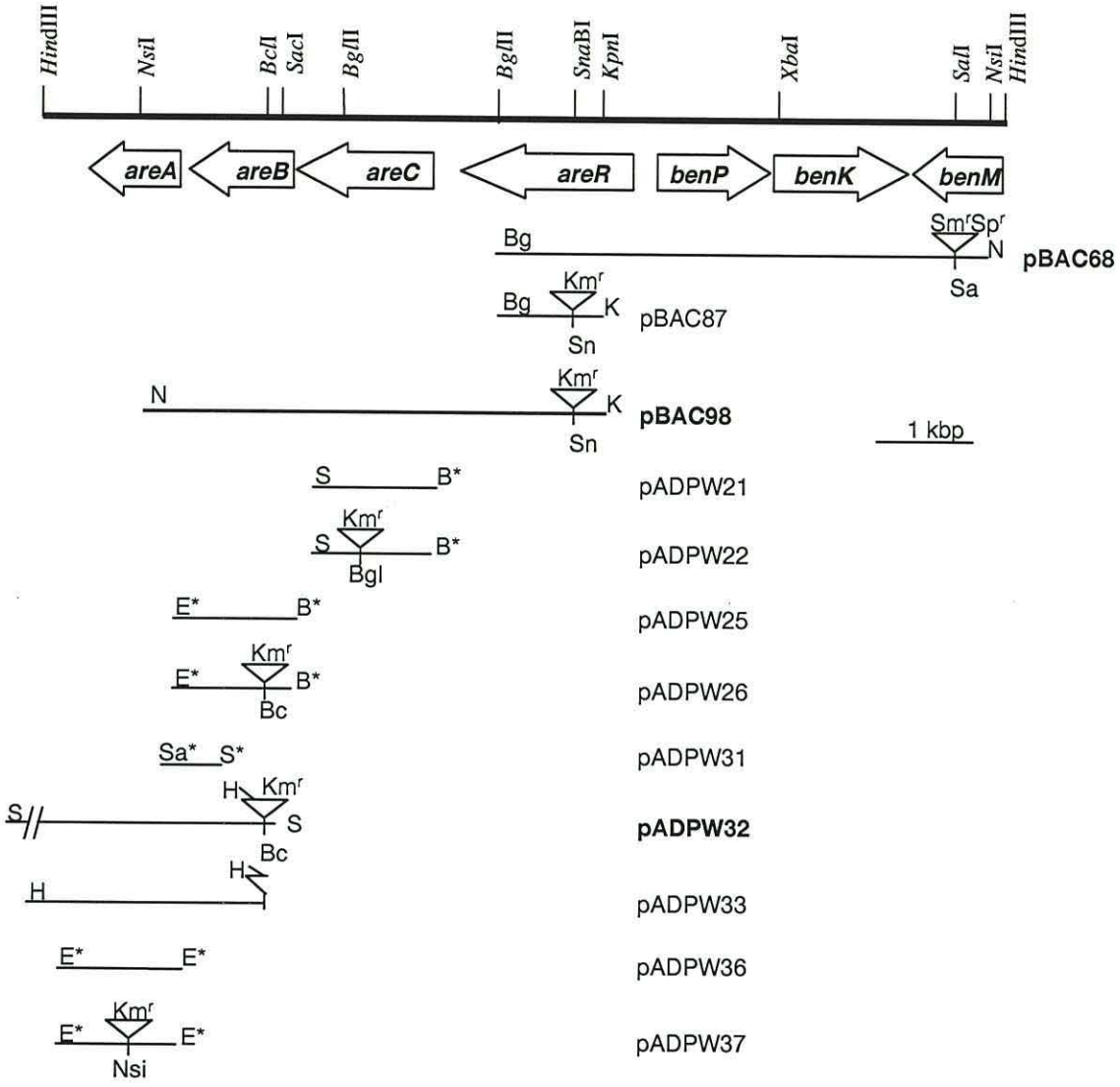


Fig. 3.2.1. Physical map of the DNA adjacent to the *ben* genes in ADP1. Physical map of the *areCBA* genes and their location relative to one end of the supraoperonic *ben-cat* cluster. The various inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 3.2.1. Plasmids named in boldface type contain inserts that were cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The Km^r cassette insertions are not to scale. The abbreviations for the restriction sites are: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Nsi*I; Nd,*Nde*I; S, *Sac*I; Sa, *Sal*I; Sn, *Sna*BI.

TABLE 3.2.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter		
ADP1 (BD413)	Wild type	(Juni, 1969)
ADPW56	*AreC*::Km^r; transformation of ADP1 with pADPW22	This study
ADPW57	*AreB*::Km^r; transformation of ADP1 with pADPW26	This study
ADPW58	*AreA*::Km^r; transformation of ADP1 with pADPW37	This study
ADPW68	*AreB*::Km^r; transformation of ADP1 with pADPW75	This study
ISA36	*BenM*::Sm^{r}Sp^r	Collier, 1998
ACN73	*AreR*::Km^r; transformation of ADP1 with pBAC87	Jones, 1999
E. coli		
DH5α	F$^-$, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	(Gibco BRL)
XL1-Blue MRF'	Δ (*mcrA*)*183*, Δ(*mcrCB-hsdSMR-mrr*)*173*, *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, *lac* [F' *proAB lacI^{q}Z* ΔM15 Tn*10* (Tet^r)]	(Stratagene)
pUC18/19	Ap^r, cloning vector	(Yanisch-Perron, 1985)
pUC4K	Ap^r, Km^r, source plasmid for Km^r cassette	(Vieira, 1982)
pUI1637	Ap^r, Km^r, source plasmid for Km^r cassette	(Eraso, 1994)
pBAC68	7.0-kbp *Bgl*II/*Nsi*I fragment cloned from ISA36 containing Sm^{r}Sp^r cassette and the whole of *benP* and part of *areR* in pUC19	(Collier 1997)
pBAC87	pBAC78 with Km^r cassette excised from pUI1637 with *Eco*RV and cloned into *Sna*BI site in *areR*	(Jones, 1999)

pBAC98	6.7-kbp *Nsi*I/*Kpn*I fragment cloned from ACN73 containing Km^r cassette and the whole of *areCB* and part of *areR* and *areA*	Gift from E. Neidle
pADPW21	1.6-kbp *Bam*HI*/*Sac*I fragment containing *areC* in pUC18	This study
pADPW22	pADPW21 with Km^r cassette from pUC4K cloned into *Bgl*II site in *areC*	This study
pADPW25	1.4-kbp *Bam*HI*/*Eco*RI* fragment containing *areB* in pUC18	This study
pADPW26	pADPW25 with Km^r cassette from pUC4K cloned into *Bcl*I site in *areB*	This study
pADPW31	0.65-kbp *Sac*I*/*Sal*I* fragment containing part of *areA* in pUC18	This study
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and the whole of *areA* in pUC18	This study
pADPW33	3.6-kbp *Hin*dIII subclone of pADPW32 in pUC18	This study
pADPW36	1.2-kbp *Eco*RI* fragment containing *areA* in pUC18	This study
pADPW37	pADPW36 with Km^r cassette from pUC4K cloned into *Nsi*I site in *areA*	This study
pADPW72	1.0-kbp *Eco*RI*/*Hin*dIII* fragment containing part of *areB* in pUC18	This study
pADPW75	pADPW72 with Km^r cassette from pUI1637 cloned into *Sac*I site in *areB*	This study
	* denotes restriction site added by PCR	

TABLE 3.2.2. *Acinetobacter* sp. ADP1 genes and gene products.

Gene designation	Putative function of gene product	Size of gene (bp)	%(A+T)	Size of gene product		Most similar gene products (% amino acid identity/similarity) (Acc. no. or reference)
				Residues	kDa	
areR	Regulatory protein	1803	60	600	68.4	AcoR (*P.putida*) (25/48%) (3688510) AcoR (*C.magnum*) (21/43%) (Kruger *et al.* 1994)
areC	Benzaldehyde dehydrogenase	1455	53.2	484	51.9	XylC(*A.calcoaceticus*)(77/88%) (Gillooly *et al.* 1998) XylC (*P.putida*) (44/64%) (D63341)
areB	Benzyl alcohol dehydrogenase	1116	53.1	371	38.9	XylB(*A.calcoaceticus*)(82/91%) (Gillooly *et al.* 1998) XylB (*P.putida*) (56/70%) (D63341)
areA	Benzyl esterase	981	63	326	37.1	Esterase (*A. fulgidus*) (31/48%) (2648837) Esterase (*P.putida*) (28/50%) (Choo *et al.* 1998)

1998) (Table 3.2.2.), and 70% similarity and 56% identity to the benzyl alcohol dehydrogenase encoded by the *xylB* gene on the *Pseudomonas* TOL plasmid pWW0 (Shaw *et al.* 1993) (Table 3.2.2.).

3.2.3. Identification of the *areA* gene.

Due to the close proximity of the partial open reading frame to *areB*, and its same direction of transcription, it was hypothesised that the two genes may be co-transcribed. To clone the additional chromosomal DNA downsteam of the partial open reading frame, the *areB* gene was subcloned by PCR from pBAC98 into pUC18, creating pADPW25. This was done using the following PCR primers where the *Eco*RI or *Bam*HI restriction sites used for cloning into pUC18 are underlined and the bases which differ from the wild-type sequence are in bold: areB1 (forward) 5'-ACAGGATGGATCCATATGACAAAGTTTACC-3'; areB1 (reverse) 5'-TTCAATGAATTCCCCCGCTACATGAG-3'. The cloned *areB* gene was then disrupted by the insertion of a Km^r cassette from pUC4K (Vieira & Messing, 1982) creating pADPW26 (*areB*::Km). The cassette was inserted into the *BclI* restriction enzyme site that is approximately 200bp into *areB*. This was achieved since *Bam*HI, (by which the Km^r cassette fragment from pUC4K was excised) has compatible ends to *BclI*. (Fig 3.2.1.). pADPW26 was linearised and transformed into ADP1 creating ADPW57 (Table 3.2.1.). Southern hybridisation analysis of ADPW57 confirmed the integration of the Km^r cassette into the ADP1 genome (Fig. 3.2.2.). The developed Southern Blot also showed that ADPW57 had an 8.0kbp *Sac*I fragment that hybridised with *areB*. This fragment would be useful in cloning the remaining DNA of the partial open reading frame and DNA further upstream of it (Fig.3.2.2; lanes 11 and 12.). Thus, chromosomal DNA from ADPW57 was cut with *Sac*I, ligated to *Sac*I-cut pUC18, and transformed into *E. coli*. Selection for Km^rAp^r transformants produced the plasmid pADPW32, which has an 8.0 kbp insert in pUC18 (Fig.3.2.1.). Southern hybridisation of pADPW32 using a probe for the N-terminal end of *areA* confirmed that pBAC98 and pADPW32 contained overlapping chromosomal DNA fragments (data not shown).

Fig. 3.2.2. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *areB* to create ADPW57.

Fig. 3.2.2.a. Map of the DNA sequence upstream and downstream of *areB* in ADP1, pADPW25 (1.4kbp), and pADPW26 (*areB*::Km^r 2.6kbp). ADPW57 was created by transformation of pADPW26 into ADP1. ADPW57 has a chromosomal rearrangement where the 1.2kbp Km^r cassette is integrated into the genome at the *BclI* restriction site that is adjacent to the *Xba*I site. Restriction sites marked with an asterisk are on the plasmid multi-cloning site.

Fig. 3.2.2.b. The agarose gel blot was hybridised with an *areB* probe purified by band extraction from *Eco*RI /*Bam*HI digested pADPW25. The analysis of ADPW57 digested with *SacI* showed the expected size of a fragment containing the DNA downstream of *areB* that would also contain the kanamycin resistance marker. Lanes S contain 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *areB* are shown (kbp):

Lane 1: pADPW25 (*Bam*HI /*Eco*RI) 1.4.
Lane 2: pADPW26 (*Bam*HI /*Eco*RI) 2.6.
Lane 3: ADP1 (*Xba*I /*Bgl*II) 1.9.
Lane 4: ADPW57 (*Xba*I /*Bgl*II) 3.3.

Lane 5: pADPW25 (*Hin*dIII) 4.1.
Lane 6: pADPW26 (*Hin*dIII) 3.15, 0.95.
Lane 7: ADP1 (*Hin*dIII) 4.1.
Lane 8: ADPW57 (*Hin*dIII) 3.5,1.8.

Lane 9: pADPW25 (*Sac*I) 4.1.
Lane 10: pADPW25 (*Sac*I) 5.3.
Lane 11: ADP1 (*Sac*I) 7.0.
Lane 12: ADPW57 (*Sac*I) 8.2.

Fig. 3.2.2.a.

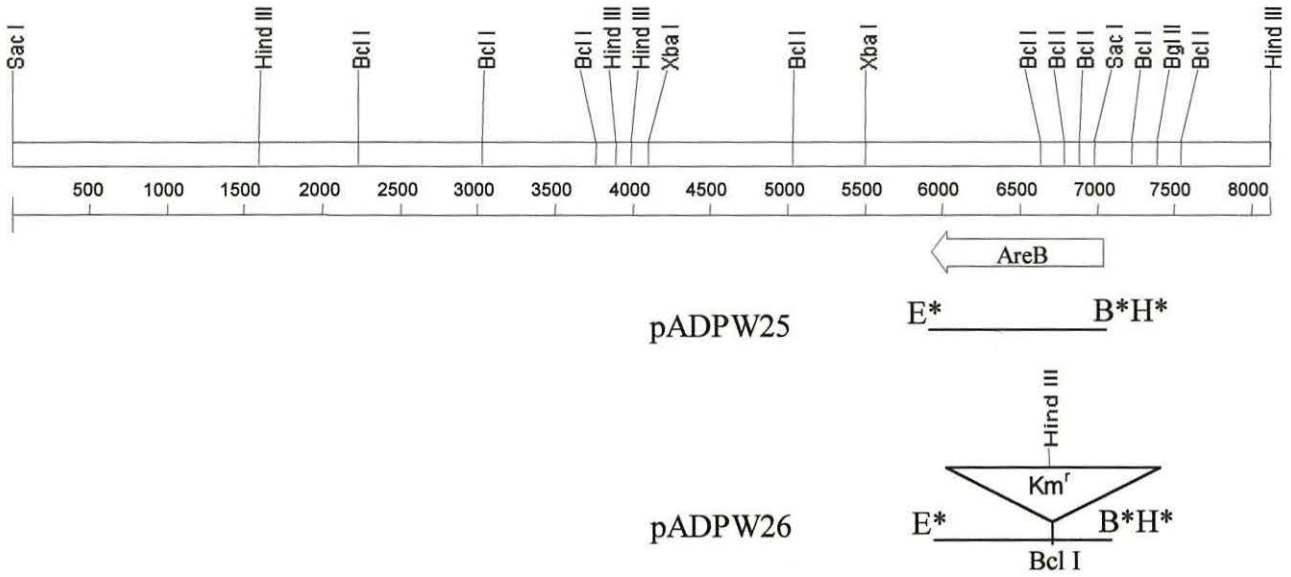
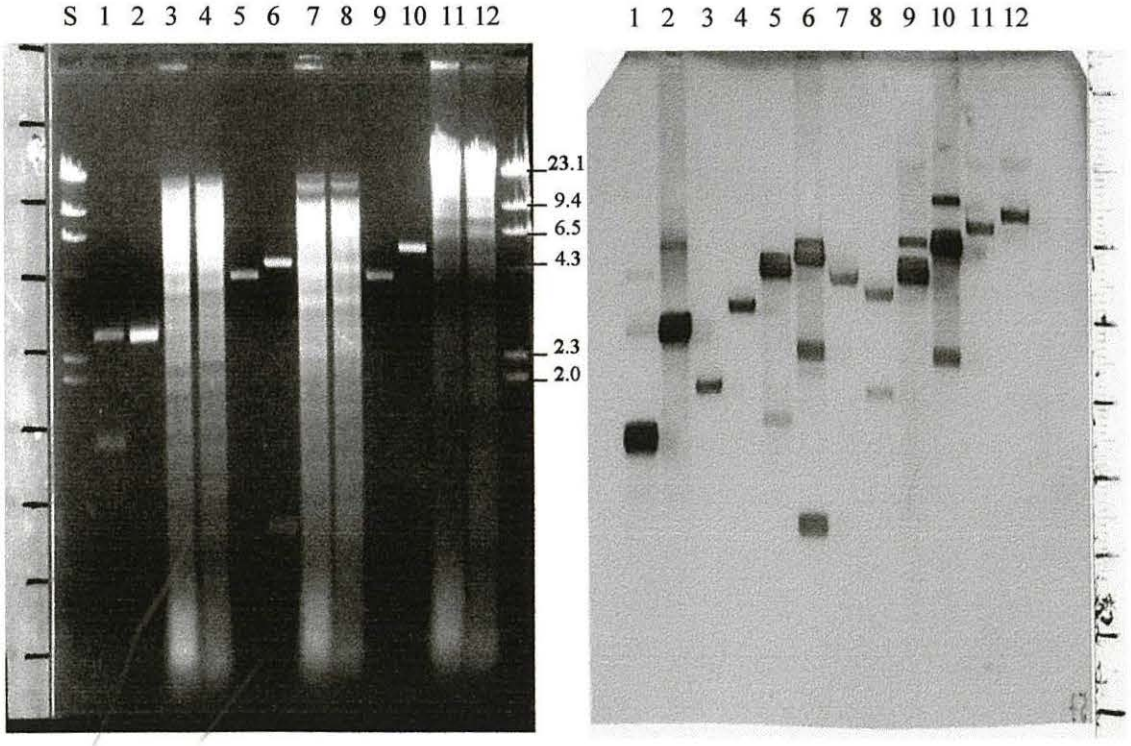


Fig. 3.2.2.b.



The probe used for the hybridisation was a PCR-generated fragment of the N-terminal end of *areA*, which was cloned into pUC18 creating pADPW31 (Fig 3.1.1.). The pBAC98 sequence was amplified with the primer sequences areA1 (forward) 5'-ATGTTATGTCGACTTTGGTTTC-3', and areA1 (reverse) 5'-GCATCATGAGGAGCTCGAGGGTAT-3'. The *Sal*I and *Sac*I sites used for cloning the amplified DNA are underlined and the bases altered from the chromosome are in boldface.

pADPW33 was then sub-cloned from pADPW32 (Fig.3.2.1.) (Table. 3.2.1.) to use as a DNA sequencing template. Sequencing of pADPW33 was done by MWG-Biotech Ltd (Ebersberg, Germany). Sequence alignment of pBAC98 and pADPW33 confirmed that the fragments overlapped. The complete sequence for the partial open reading frame, designated *areA*, was obtained.

From GenBank database searches, the predicted AreA protein has homology to a number of esterases particularly 48% similarity and 31% identity to an esterase from *Archaeoglobus fulgidus* and 50% similarity, and 28% identity to an esterase from *Pseudomonas* sp. (Choo *et al.* 1998) (Table 3.2.2.).

3.3.1. Chromosomal disruption of *areA*, *areB*, and *areC* in *Acinetobacter* sp. ADP1.

Strains ADPW58 (*areA*::Km), ADPW68 (*areB*::Km), and ADPW56 (*areC*::Km) (Table3.2.1.) were created by similar methods to those used to create ADPW57 (see section 3.2.3.). It was necessary to make another *areB*::Km^r strain i.e. ADPW68, since ADPW57 showed a general growth weakness for reasons which we have not yet diagnosed. In comparative growth tests ADPW57 grew much more poorly than did ADP1 even on succinate as the sole carbon source.

The gene knock-outs for *areA*, *areB*, and *areC* were done as follows. pUC18-derived plasmids were made that carried one of the *are* genes disrupted by a Km^r cassette. In the following list of the PCR primers used to make these plasmids, the *Eco*RI or *Bam*HI restriction sites used for cloning into pUC18 are underlined and the bases which differ from the wild-type sequence are in bold.

The whole of the *areC* gene was PCR cloned from pBAC98 into pUC18 on a *Bam*HI/*Sac*I fragment, creating pADPW21 (Fig. 3.2.1.). The reverse primer was chosen from downstream of the native *Sac*I site (Fig. 3.2.1.), and the forward primer contains an engineered *Bam*HI site: areC (forward) 5'-AGGACGTTACGGGATCCATATGACATTACT-3'; areC (reverse) 5'- ATGCCCATCTGGATCTCCACCACTGAAGTT-3'. The *areC* gene was disrupted by the insertion of the cassette from pUC4K (Vieira & Messing, 1982) into the unique *Bgl*II site of pADPW21 creating pADPW22 (Fig.3.2.1).

A plasmid-borne *areB* disruption was made after first constructing pADPW72, which has a PCR-generated *Eco*RI-*Hin*dIII insert in pUC18. This 1.0 kbp fragment, with 0.5kbp flanking the *Sac*I site in *areB*, was amplified from pBAC98. Primer sequences, with the *Eco*RI and *Hin*dIII sites underlined and the altered bases in bold, were as follows: areB2 (forward) 5'-TTGAATTCGGCGCGACCCTTGAAATTGGAG-3', and areB2 (reverse) 5'-TTGAAAGCTTCACCTGCACCAGTTTGTATGCC-3'. The central *Sac*I site in the resultant pADPW72 (Fig. 3.2.1.) was used as the insertion site for a Km^r cassette from pUI1637 (Eraso & Kaplan, 1994) to create pADPW75 (Fig. 3.2.1.).

The whole of the *areA* gene was PCR cloned from pADPW32 into pUC18 creating pADPW36 (Fig. 3.2.1.) using the following primers: areA2 (forward) 5'-AAAGAATTCTTAACATATGTTACAGAGTTT-3'; areA2 (reverse) 5'-GGCATGCAGGGAATTCGAACCAGAGATA-3'. To interrupt *areA* on pADPW36, the pUC4K (Vieira & Messing, 1982) cassette was cloned into a unique *Nsi*I site, creating pADPW37 (Fig. 3.2.1.).

Plasmids pADPW22, pADPW75 and pADPW37 were linearised by digestion with a restriction enzyme and each was used to transform ADP1. Southern hybridization confirmed that in strains ADPW56 (*areC*::Km^r) (Fig.3.3.1), ADPW68 (*areB*::Km^r) (Fig. 3.3.2.), and ADPW58 (*areA*::Km^r) (Fig3.3.3.), the altered plasmid-borne allele had replaced the corresponding chromosomal wild-type region.

Fig. 3.3.1. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *areC* to create ADPW56.

Fig. 3.3.1.a. Map of the DNA sequence upstream and downstream of *areC*, pADPW21 (1.6kbp), and pADPW22 (*areC*::Km^r 2.8kbp). ADPW56 is created by transformation of pADPW22 into ADP1. ADPW56 has a chromosomal rearrangement where the 1.2kbp Km^r cassette is integrated into the ADP1 genome at the *Bgl*II restriction site. Restriction sites marked with an asterisk are from the plasmid multi cloning site.

Fig. 3.3.1.b. The agarose gel blot was hybridised with an *areC* probe purified by band extraction from pADPW21. Lane S contains a 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *areC* are shown (kbp):

Lane 1: pADPW21 (*Sac*I) 4.3.

Lane 2: pADPW22 (*Sac*I) 5.5.

Lane 3: ADP1 (*Sac*I /*Kpn*I) 3.2.

Lane 4: ADPW56 (*Sac*I /*Kpn*I) 4.4.

Fig.3.3.1.a

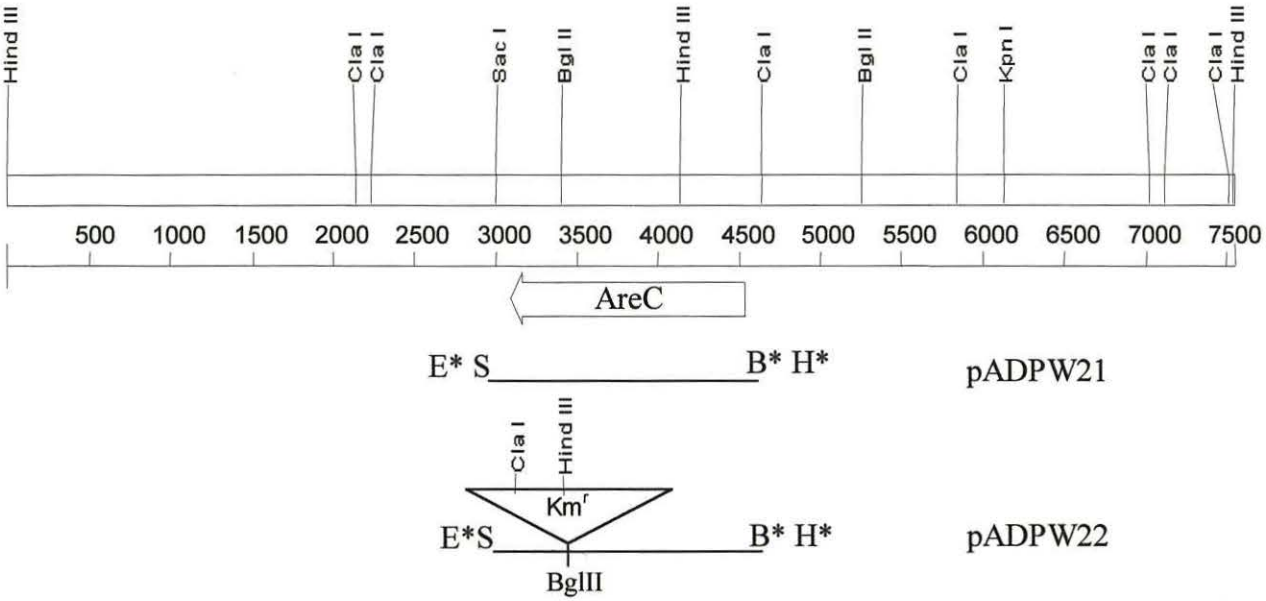


Fig.3.3.1.b

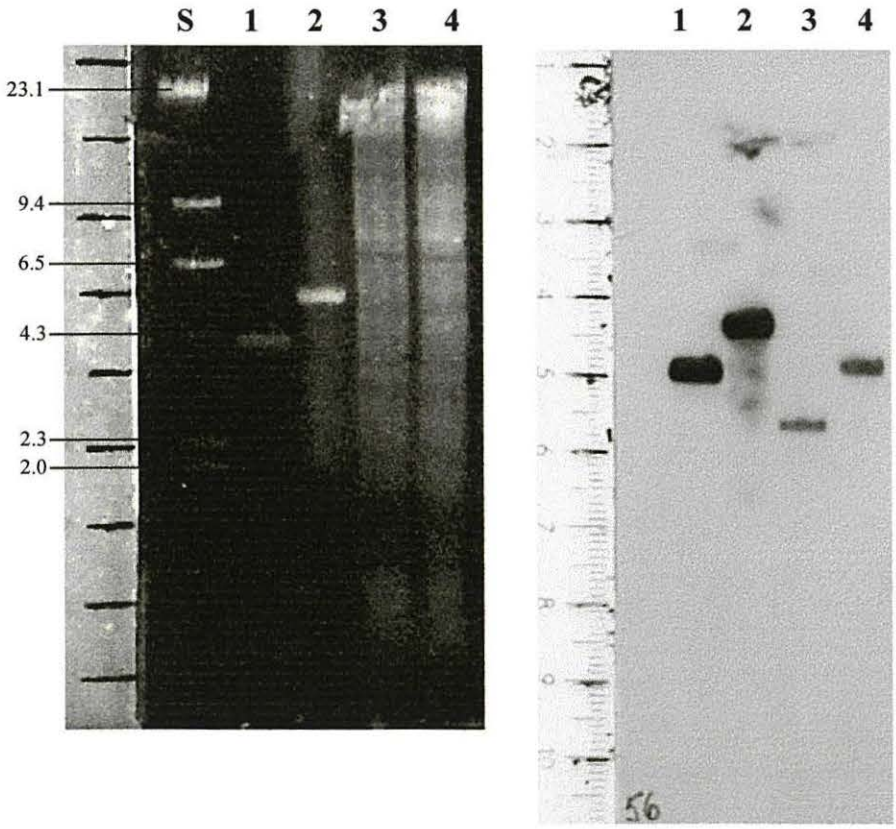


Fig. 3.3.2. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *areB* to create ADPW68.

Fig. 3.3.2.a. Map of the DNA sequence adjacent to *areB*, pADPW72 (1.0kbp), and pADPW75 (*areB*::Km^r, 3.2kbp). ADPW68 was created by transformation of pADPW75 into ADP1. ADPW68 has a chromosomal rearrangement where the 2.2kbp Km^r cassette from pUI1637 is integrated into the ADP1 genome at the *Sac*I restriction site. Restriction sites marked with an asterisk are from the plasmid multi cloning site.

Fig. 3.3.2.b. The agarose gel blot was hybridised with an *areB* probe purified by band extraction from *Eco*RI /*Hin*dIII digested pADPW72. Lanes S contain 8*Hin*dIII size marker (Promega) and the fragment sizes stated (kbp). The sizes of the bands produced following hybridisation to *areB* are shown (kbp):

Lane1: pADPW72 (*Hin*dIII /*Eco*RI) 1.0.
Lane2: pADPW75 (*Hin*dIII) 3.2, 0.5 (out of range).
Lane3: ADP1 (*Hin*dIII) 4.0.
Lane4: ADPW68 (*Hin*dIII) 3.0. 1.2 (very faint).

Lane5: pADPW72 (*Bgl*II/*Xba*I) 3.9.
Lane6: pADPW75 (*Bgl*II/*Xba*I) 4.6, 0.5 (out of range).
Lane7: ADP1 (*Bgl*II/*Xba*I) 1.9.
Lane8: ADPW68 (*Bgl*II/*Xba*I) 3.0, 0.4 (out of range).

Fig.3.3.2.a.

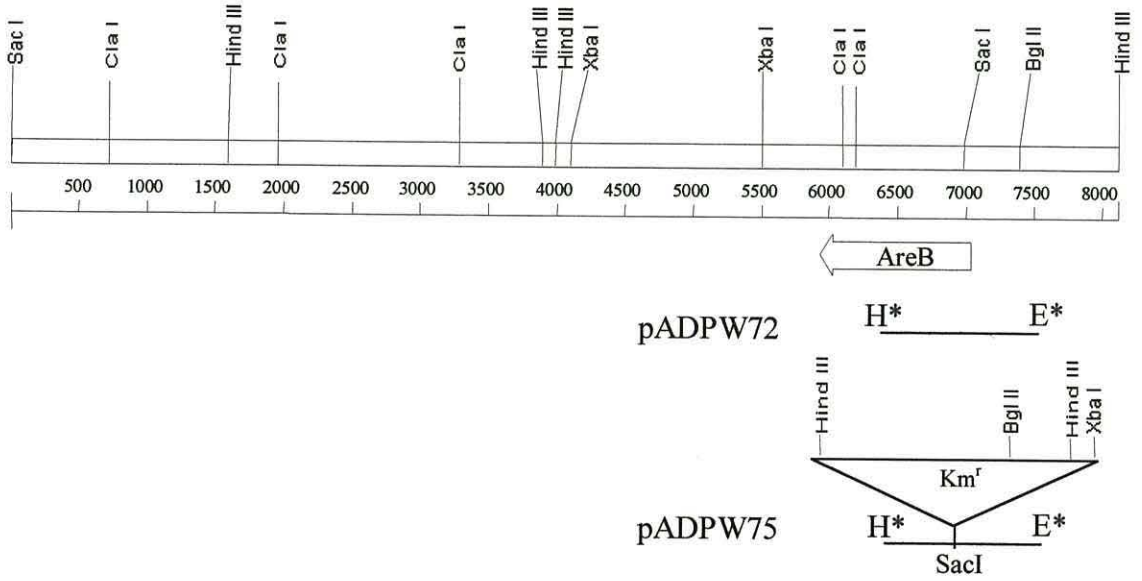


Fig.3.3.2.b.

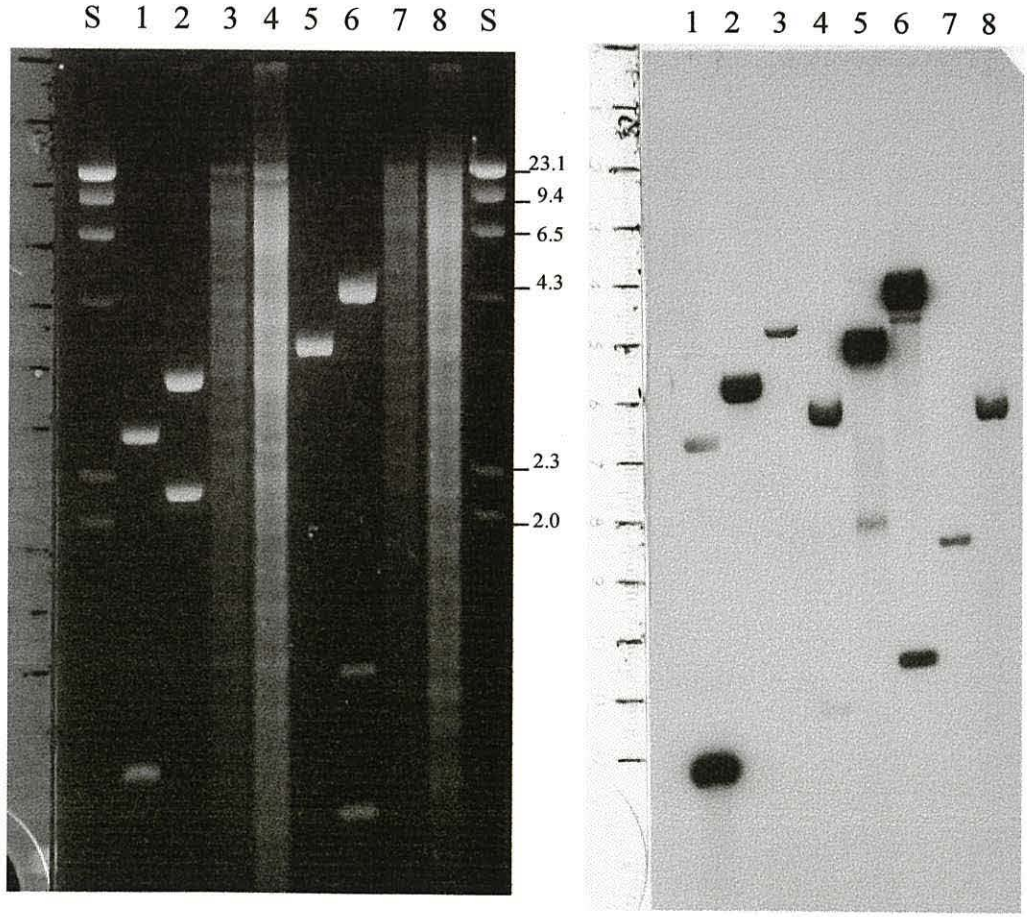


Fig. 3.3.3. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *areA* to create ADPW58.

Fig. 3.3.3.a. Map of the DNA sequence upstream and downstream of *areA*, pADPW36 (1.0kbp), and pADPW37 (*areA*::Km^r 2.2kbp). ADPW58 was created by transformation of pADPW37 into ADP1. ADPW58 has a chromosomal rearrangement where the 1.2kbp Km^r cassette from pUC4K is integrated into the genome at the *Nsi*I restriction site. Restriction sites marked with an asterisk are from the plasmid multi cloning site.

Fig. 3.3.3.b. The agarose gel blot was hybridised with an *areA* probe purified by band extraction from *Eco*RI digested pADPW36. Lane S contain 8*Hin*dIII size marker (Promega) and the fragment sizes are stated (kbp). The sizes of the bands produced following hybridisation to *areA* are shown (kbp):

Lane1: pADPW36 (*Hin*dIII) 3.7.

Lane2: pADPW37 (*Hin*dIII) 3.8, 1.1.

Lane3: ADP1 (*Hin*dIII) 4.1.

Lane4: ADPW58 (*Hin*dIII) 3.5, 1.8.

Fig. 3.3.3.a.

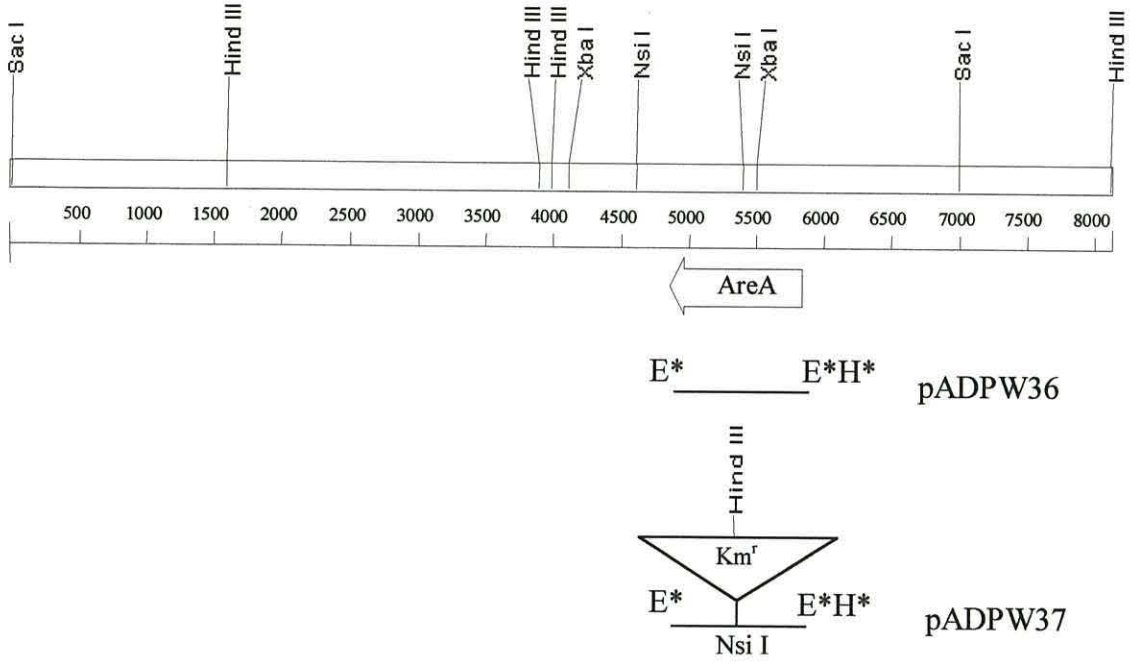
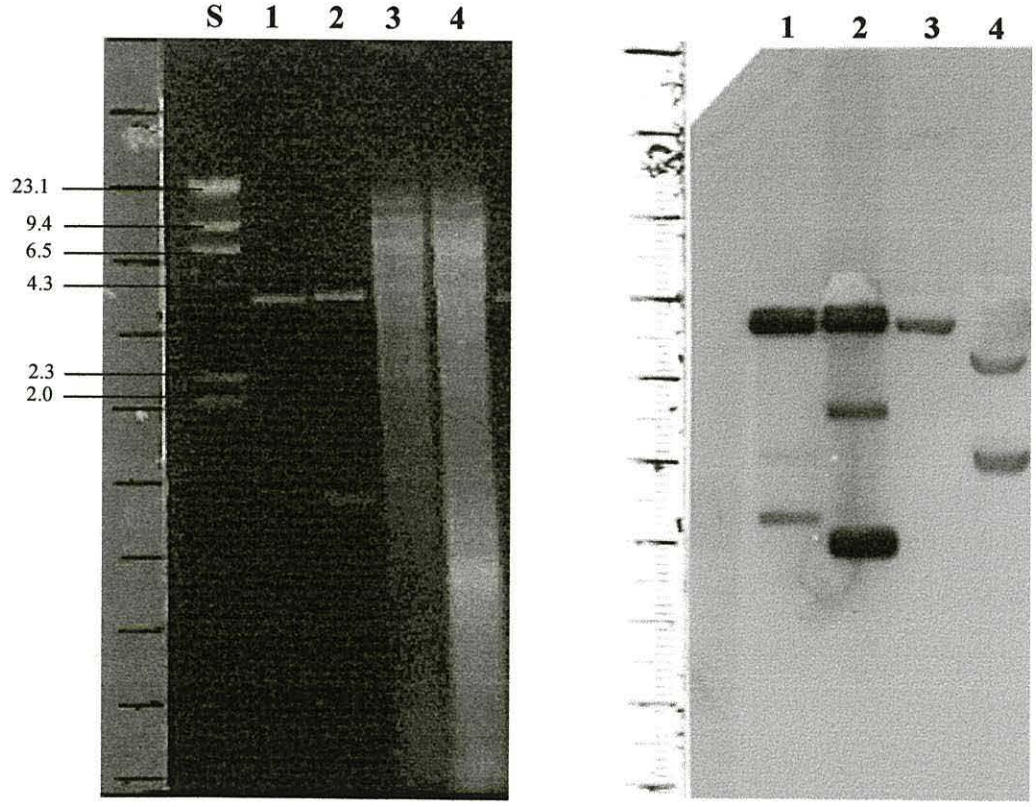


Fig. 3.3.3.b.



3.4. The phenotypes of *are* gene knockout mutants.

3.4.1 Growth tests to determine the phenotypes of ADPW58, ADPW68, and ADPW56.

A Km^r cassette in strains ADPW58, ADPW68 and ADPW56 respectively (Table 3.2.1.) specifically disrupted the chromosomal copy of *areA, areB* and *areC*. These were tested for plate growth on benzyl acetate, benzyl alcohol and benzaldehyde. Benzyl acetate was used because of the hypothesis that the function of AreA might be to cleave the ester bond in benzyl acetate, releasing benzyl alcohol and acetate.

With ADPW56 (*areC*::Km) the rate of growth of the mutant strains on all three substrates was severely reduced, but not completely eliminated. Patches took from 2-3 days to achieve significant size compared with overnight growth for ADP1. With ADPW68 (*areB*::Km) a reduced growth rate was observed on benzyl acetate and benzyl alcohol but the strain appeared to grow normally on benzaldehyde. The growth rate of ADPW58 (*areA*::Km) was reduced only on benzyl acetate. The results were therefore consistent with disruption of a pathway involving sequential action of AreA, AreB and AreC. However, the plate growth tests did not produce a clear-cut conclusive result of the effect of the chromosomal disruption on the mutant's growth abilities.

3.4.2. Competition for growth substrate between *Acinetobacter* wild-type and *are*-mutant strains.

Wild type ADP1 and the knock-out mutants ADPW56, ADPW58 and ADPW68 were all grown overnight in minimal media containing 10 mM succinate to a cell density of about 10^9 cells/ml. A 5 µl aliquot of a 10^{-2} dilution of the wild type and one of each of the mutants were inoculated into 4 tubes of 990 µl minimal media containing either succinate, benzaldehyde, benzyl alcohol or benzyl acetate as sole carbon source respectively. The mixed cultures were grown on each substrate at 37°C for between 12 and 14 generations. Dilutions of the mixed culture were plated on LB agar. 100 colonies were patched onto LB agar containing 10 µg/ml kanamycin, and onto LB agar as a positive control. The percentage of mutants in the culture was assessed as the number of

Km^r colonies. A 10 µl aliquot of each mixed culture was further inoculated into 990 µl of minimal medium containing the corresponding substrate. These cultures were grown for an additional 12 to 14 generations after which the percentage of mutants was again determined. Competition for growth between ADP1 and ADPW56 (*areC*::Km) (Fig. 3.4.1.a); between ADP1 and ADPW68 (*areB*::Km) (Fig. 3.4.1.b); and between ADP1 and ADPW58 (*areA*::Km) (Fig. 3.4.1.c) was performed on the four substrates.

Where the plate growth rate appeared to have been reduced, the percentage of mutant colonies dropped from approximately 50% to below 6% after the second subculturing. This was the case for ADPW56 on all three substrates, for ADPW68 on benzyl acetate and benzyl alcohol, and for ADPW58 on benzyl acetate. Neither ADPW68 nor ADPW58 was outcompeted by ADP1 during growth on benzaldehyde nor was the latter during growth on benzyl alcohol. Control experiments with succinate as the sole carbon source showed no out-competition of the mutants by ADP1.

These results are consistent with those from the plate growth tests in section 3.4.1, showing that disruption of the *are* genes causes a reduction in the cell's ability to use benzyl acetate as a sole carbon growth substrate, and indicating that *areA* encodes the protein that catalyses the ester cleavage of benzyl acetate. Hence, the genes are designated '*are*' for aromatic ester catabolism.

3.5. Comparisons of the Are proteins with similar enzymes.

3.5.1. Comparison of AreC with other aldehyde dehydrogenases.

AreC has an amino acid sequence that is typical of a large family of NAD^+-linked aldehyde dehydrogenases (classII). In particular it contains the consensus sequence FTGSTXVG (residues 228-235 FTGSTQVG) (Fig. 3.5.1). Also conserved are the residues Cys285 and Glu251, homologous to Cys302 and Glu268 in the human liver mitochondrial aldehyde dehydrogenase that have been shown to be involved in the catalytic reaction (Farres *et al.* 1995; Wang & Weiner, 1995). These AreC residues also align, respectively with the *A.calcoaceticus* NCIB8250 XylC (Gillooly *et al.* 1998), and the AreC from the TOL plasmid of *P.putida* (Shaw *et al.* 1993) (Fig. 3.5.1).

Fig. 3.4.1. Growth Competition Experiments between wild type ADP1 and ADP1 mutants.

Growth rate competition experiments in liquid medium on benzyl acetate, benzyl alcohol, and benzaldehyde were set up between each of the knockout mutants: Fig. 3.4.1.a; ADPW56 (*areC::*Km) and ADP1, Fig. 3.4.1.b ADPW68 (*areB::*Km) and ADP1, and Fig. 3.4.1.c ADPW58 (*areA::*Km) and ADP1. Initial cell densities of wild type and mutants were set to be approximately equal. The proportion of the mutants during two sequential subcultures each of about 12-14 generations were determined by plating out serial dilutions and determining the percentage of Km^r colonies. Control experiments with succinate as the sole carbon source showed no out-competition of the mutants by ADP1.

Fig. 3.4.1.a

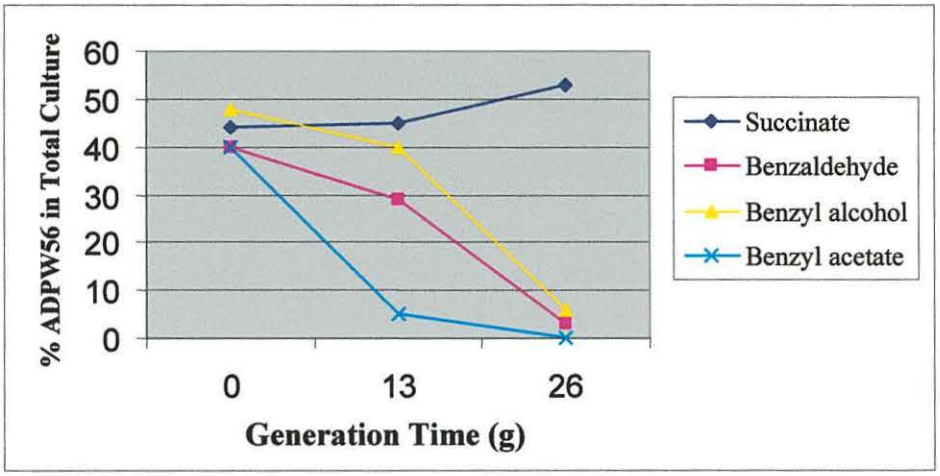


Fig. 3.4.1.b

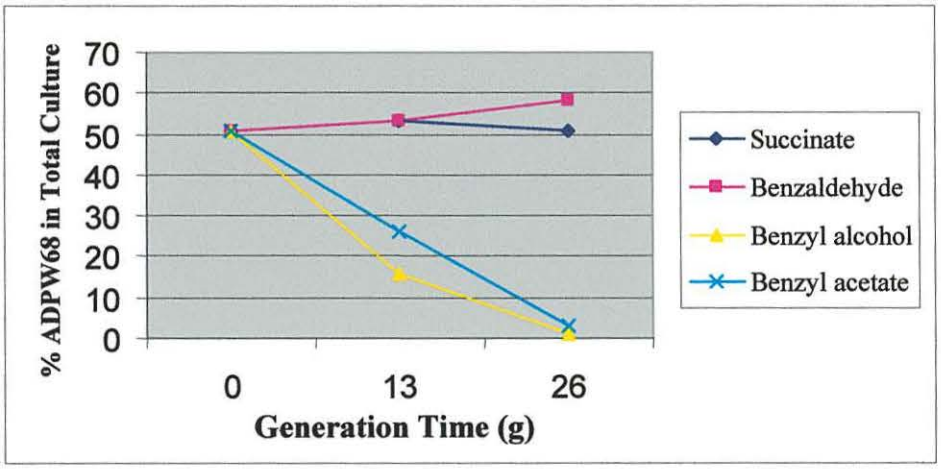
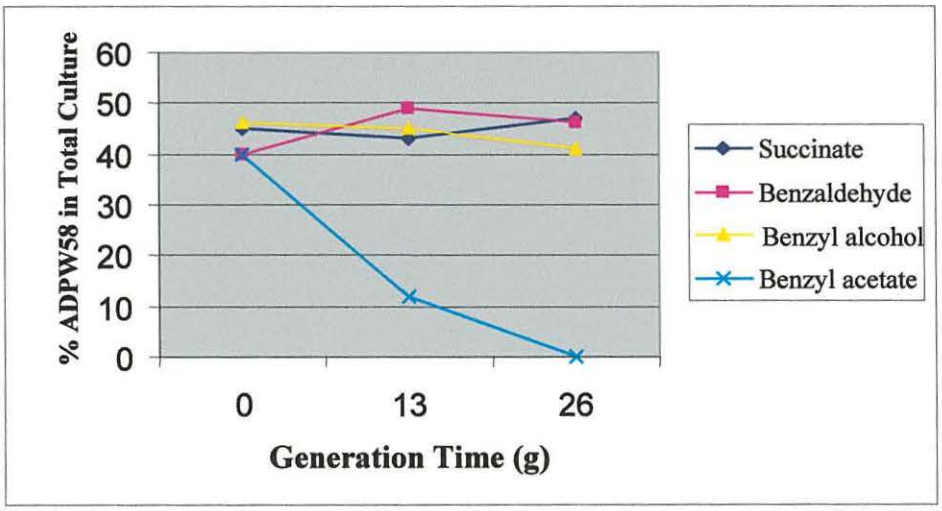


Fig. 3.4.1.c



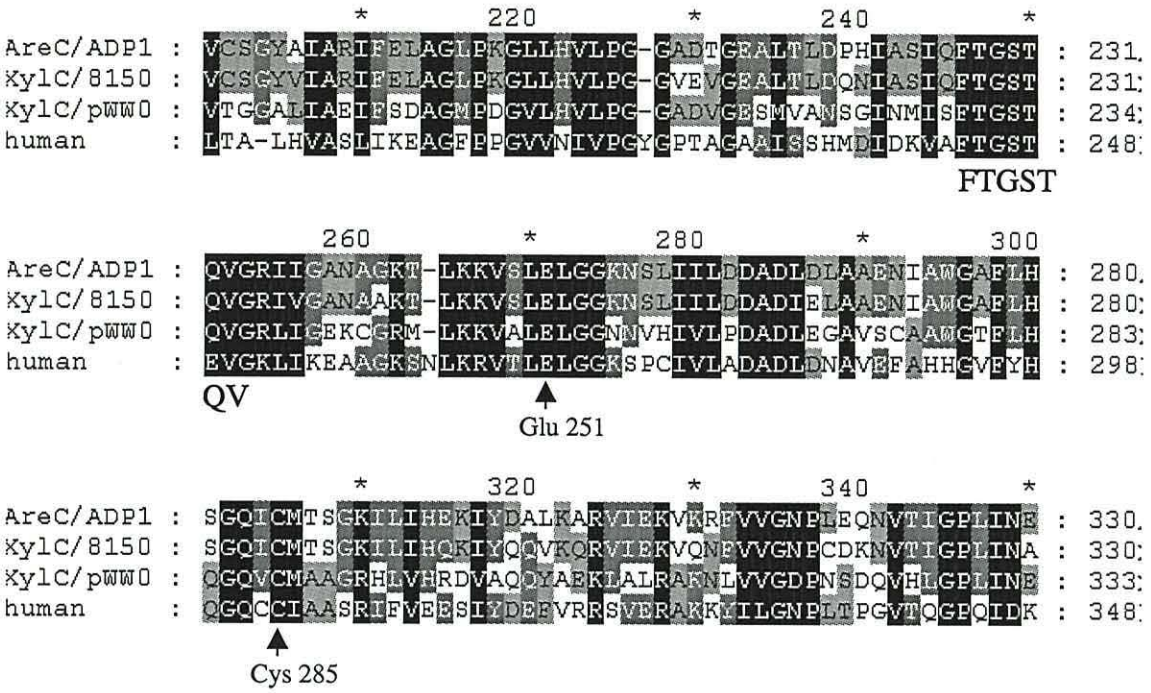


Fig. 3.5.1. Alignment of aldehyde dehydrogenases. A selective alignment of 150 residues from : ADP1/AreC: *Acinetobacter* sp. ADP1 AreC (this study); XylB/8150: *A. calcoaceticus* NCIB8250 XylB (Gillooly *et al.* 1998); XylB/pWW0: *P.putida* pWW0 XylB (Shaw *et al.* 1993); human: human liver aldehyde dehydrogenase (Farres *et al.* 1995). Residues indicated with an arrow correspond to their position in ADP1. Residues conserved in all four sequences are highlighted in black, residues conserved in three of the four sequences are highlighted in dark shade grey, and residues conserved in two of the four sequences are highlighted in light grey.

3.5.2. Comparison of AreB with other alcohol dehydrogenases.

The deduced amino acid sequence of AreB shows all the characteristics of Group 1 long chain zinc-dependent alcohol dehydrogenases (ADH) as typified by horse-liver alcohol dehydrogenase (Eklund *et al.* 1981). AreB has the conserved residues which bind the catalytic Zn^{2+} (Cys44, His65, Cys173) and the four Cys residues which bind the structural Zn^{2+} (at positions 94, 97, 100, 108) (Fig. 3.5.2). It also contains the characteristic motif GHEXXGXXXXXGXXV from residues 64 to 78 (Fig. 3.5.2). Two bacterial sequences that are similar to AreB include two benzyl alcohol dehydrogenases, both designated XylB, from *A.calcoaceticus* NCIB8250 (Gillooly *et al.* 1998) and from the TOL plasmid of *P.putida* (Shaw *et al.* 1993). In AreB, as in each of these bacterial sequences, there is a replacement of the His51 (as numbered in the horse-liver enzyme) by a hydrophobic residue, in this case Ile49. This His is conserved in almost all of the long chain zinc-dependent ADHs and is believed to act as a general base during the catalytic reaction. However, Inoue *et al.* 1998, have suggested that in the pWW0 XylB an adjacent His residue takes this role, and in AreB there is an equivalent adjacent residue (His45), as well as in the *A. calcoaceticus* NCIB8250 XylB (Fig. 3.5.2).

3.5.3. Comparison of AreA with other known esterases.

AreA appears to be a member of the α/β hydrolase fold family of enzymes (Alon *et al.*1995; Ollis *et al.* 1992; Schrag & Cygler, 1997). Despite a characteristic GXSXG consensus sequence, many members of this family display relatively low primary sequence conservation and are mainly identified using a method based on predicted structural similarities (Alon *et al.* 1995). In the case of AreA, the primary sequence can be well aligned with two other *Acinetobacter* esterases in this family (Fig. 3.5.3.). In the *Acinetobacter lwoffii* RAG-1 esterase, the residues that comprise the catalytic triad were suggested to be Ser149, either Asp196 or Glu244, and either His265, His274 or His298 (Alon & Gutnick, 1993). Three of these residues, Ser149, Glu244 and His274, align with those in AreA indicating that in AreA, Ser166, Asp266 and His296 may form the catalytic triad. These AreA residues also align, respectively, with Ser201, Asp295 and

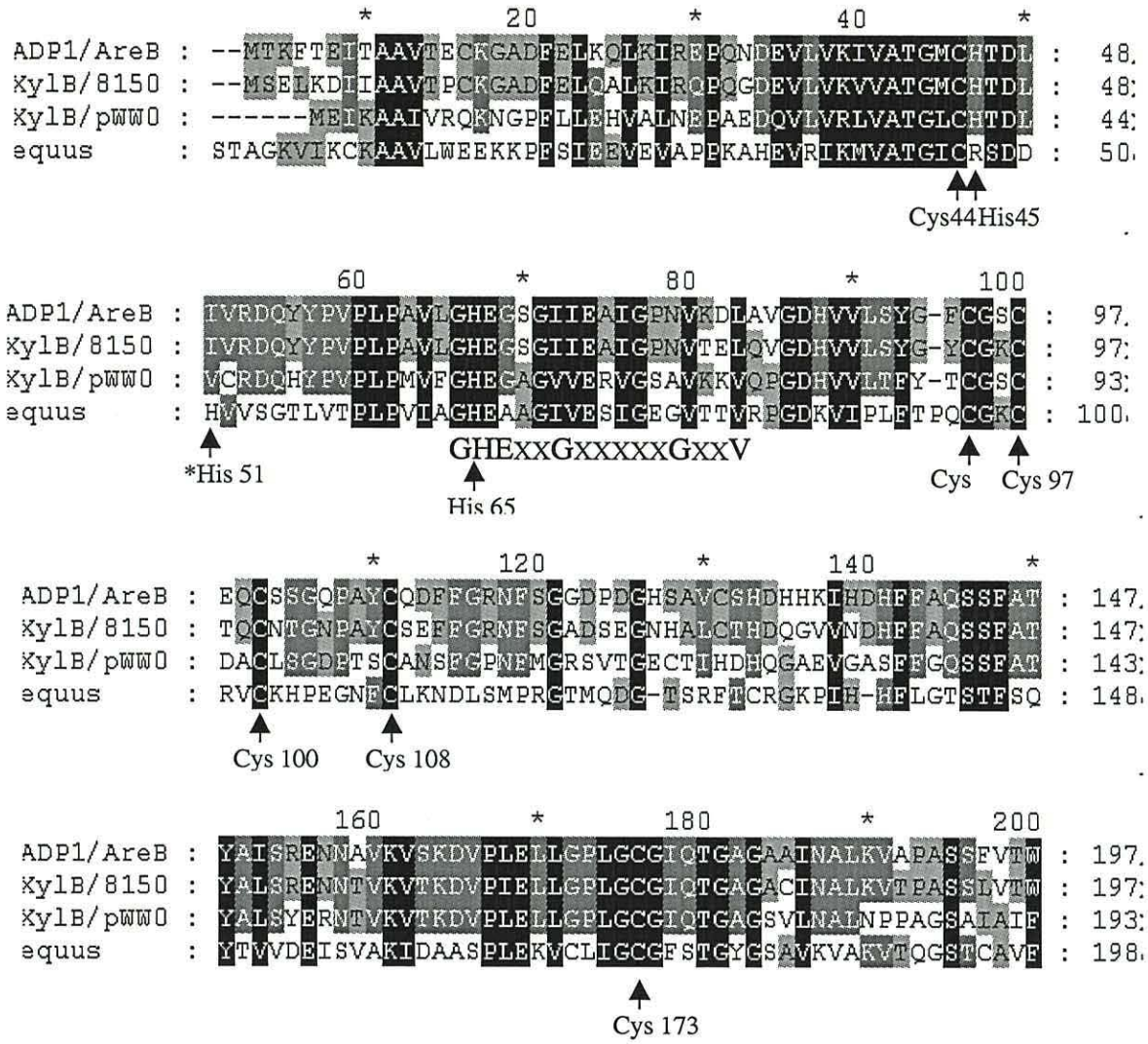


Fig. 3.5.2. Alignment of alcohol dehydrogenases. A selective alignment of 200 residues from: ADP1/AreB: *Acinetobacter* sp. ADP1 AreB (this study); XylB/8150:*A. calcoaceticus* NCIB8250 XylB (Gillooly *et al.* 1998); XylB/pWW0: *P.putida* pWW0 XylB (Shaw *et al.* 1993); equus: horse liver ADH (Eklund *et al.* 1981). Residues indicated with an arrow correspond to their position in ADP1, save the residue which is asterisked, which corresponds to its position in horse liver ADH. Residues conserved in all four sequences are highlighted in black, residues conserved in three of the four sequences are highlighted in dark shade grey, and residues conserved in two of the four sequences are highlighted in light grey.

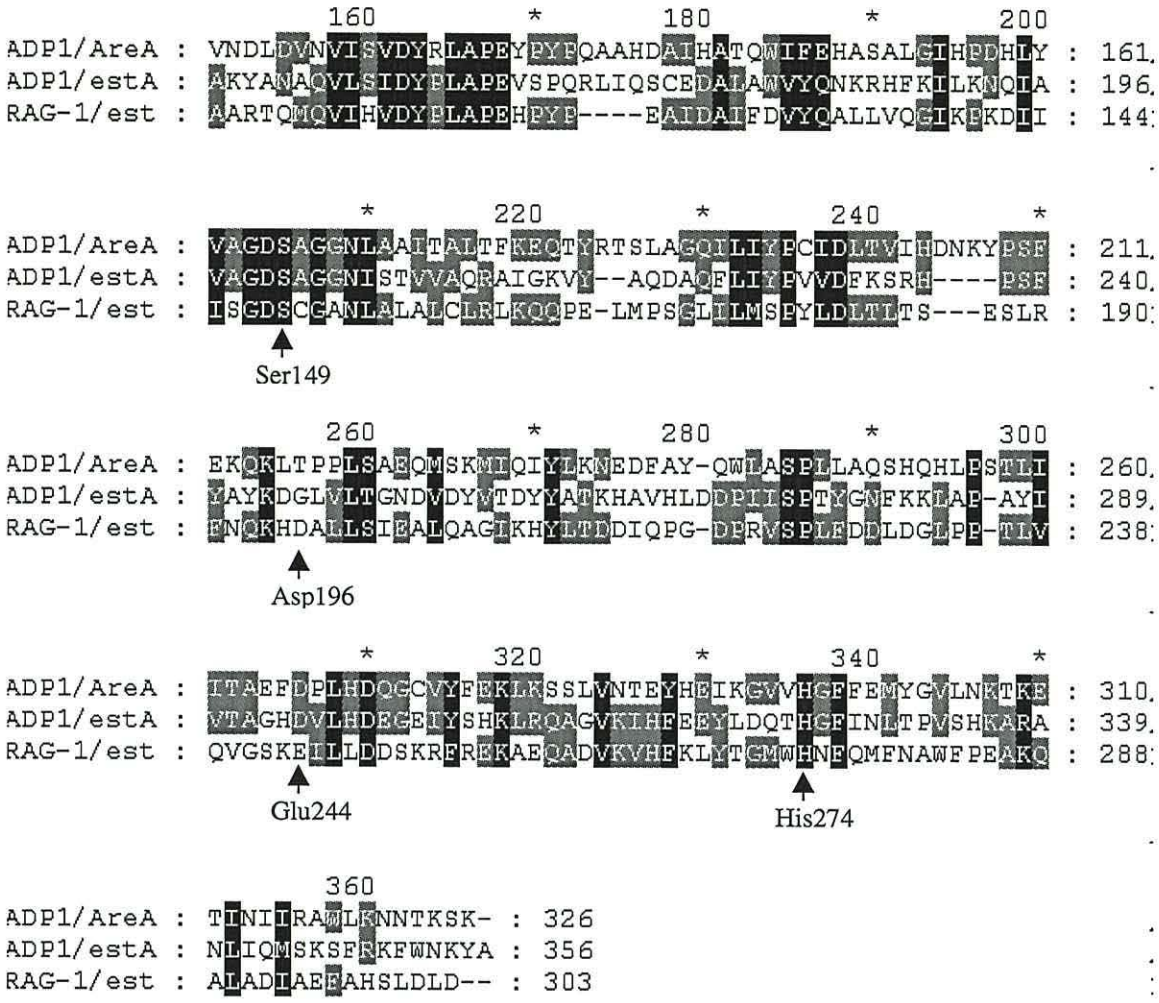


Fig. 3.5.3. Alignment of esterases. Selective alignment of 220 residues from: ADP1/AreA: *Acinetobacter* sp. ADP1 AreA (this study); ADP1/estA: *Acinetobacter* sp. ADP1 estA (Kok *et al.* 1993); RAG-1/est: *A.lwoffii* esterase (Alon & Gutnick, 1993). Conserved residues indicated with an arrow correspond to their position in RAG-1. Residues that are conserved in all three sequences are highlighted in black and residues conserved in two of the three sequences are highlighted in grey.

His325 of a carboxylesterase of *Acinetobacter* BD413 (ADP1) (Kok *et al.* 1993) (Fig. 3.5.3.).

3.6. Discussion.

3.6.1. Hypothesis for the catabolism of aryl esters by the *areABC* gene products.

The evidence presented in this Chapter strongly supports a hypothesis that the *areABC* genes of ADP1 encode enzymes for the sequential conversions of benzyl acetate to benzyl alcohol, benzaldehyde and benzoic acid (Fig. 3.6.1). The *areABC* genes would thereby direct the metabolism of these substrates into the catechol branch of the β-ketoadipate pathway that is encoded by the adjacent *ben* and *cat* genes.

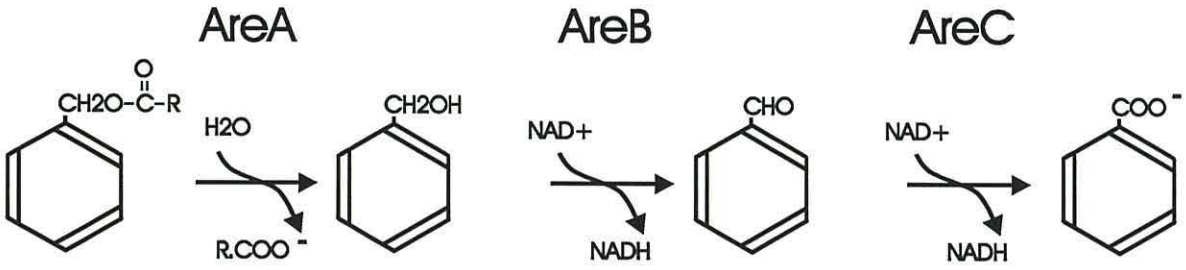


Fig. 3.6.1. Hypothesis for the catabolism of benzyl acetate by reactions catalysed by the *are* gene products.

CHAPTER 4

Biochemical assays of the AreA, AreB and AreC proteins

4.1. Introduction.

In Chapter 3 it was proposed that the *areABC* genes encode enzymes for the sequential conversions of esters of benzyl alcohol to benzyl alcohol, benzaldehyde and benzoic acid. To further prove this hypothesis, the biological functions of the three gene products were studied by enzyme assays of cell-free extracts of induced or uninduced *Acinetobacter* sp. ADP1 and of extracts of cells carrying cloned *are* genes expressed to high specific activities in *Escherichia coli.* The proteins were assayed against a range of unsubstituted and substituted aromatic esters, alcohols and aldehydes. The relative activities of the wild type, and high expression cell free extracts against the aromatic substrates were compared.

4.2. The preparation of cell free extracts containing the AreA, B, or C proteins.

4.2.1. The cloning of *areC*, *areB*, and *areA* in *E. coli.*

Oligonucleotide primers were designed to generate PCR fragments of each *are* gene with a novel *Nde*I site introduced at the putative ATG start codon of each reading frame as well as a second constructed restriction site upstream of the *Nde*I site to allow cloning into pUC18. A natural or constructed restriction site downstream of the gene was used to clone the PCR-generated fragments first into pUC18 (Chapter 3.2.3. & 3.3.1.). In this way, the *areC* and *areB* genes were amplified from pBAC98, and *areA* was amplified from pADPW32 to create pADPW21, pADPW25, and pADPW36 respectively (Table 4.2.1., Fig.4.2.1.). The constructed *Nde*I sites are italicised (CATATG) and the 3'- ATG of this site corresponds to the start codon of the respective *are* gene. The *Eco*RI or *Bam*HI restriction sites used for cloning into pUC18 are underlined and the bases which differ from the wild-type sequence are in bold: areB (forward) 5'-ACAGGATGGATC**C***A***T***A***T***G**ACAAAGTTTACC-3'; areB (reverse) 5'-TTCAATGAATTC**C**CCCCGCTACATGAG-3'; areA2 (forward) 5'-

AAAGAATTCTTAA*CATATGT*TACAGAGTTT-3'; areA2 (reverse) 5'-GGCATGCAGGGAATTCGAACCAGAGATA-3'. In the case of *areC*, the reverse primer was chosen from downstream of the native *Sac*I site (Fig. 4.2.1.): areC (forward) 5'-AGGACGTTACGGGATC*CATATG*ACATTACT-3'; areC (reverse) 5'-ATGCCCATCTGGATCTCCACCACTGAAGTT-3'. These three clones were sequenced on one strand to check that mutations had not been incorporated by the PCR reaction (MWG-Biotech Ltd, Ebersberg, Germany).

4.2.2. Cloning of individual *are* genes into pET5a high expression vector.

Fragments from each of pADPW21, pADPW25, and pADPW36 were excised using *Nde*I and *Eco*RI, religated into the expression vector pET5a, and then transformed into *E.coli* DH5α to produce plasmids pADPW29 (*areC*), pADPW30 (*areB*), and pADPW40 (*areA*) (Table 4.2.1.)(Fig. 4.2.1.). Each of the *are* genes was inserted such that gene expression would be optimally controlled by the pET5a T7 promoter on the plasmid when carried by an appropriate *E. coli* host.

4.2.3. Expression of individual *are* genes.

Plasmids pADPW29, pADPW30 and pADPW40 were transformed into *E.coli* BL21 (DE3) pLysS and individually expressed by growth in 500 ml of LB broth containing ampicillin, to an O.D. at 600 nm of 0.6. The cultures were induced by the addition of 0.4 mM IPTG and further incubated for 4 hours. Cell-free extracts of each culture were prepared as described in Chapter 2.9.

SDS-PAGE analysis of extracts of the induced *E.coli* hosts showed high levels of expression of each protein, each with the size predicted from the putative sequence. The over-expressed proteins were not observed in either uninduced strains or those lacking the recombinant plasmids (Fig. 4.2.2.).

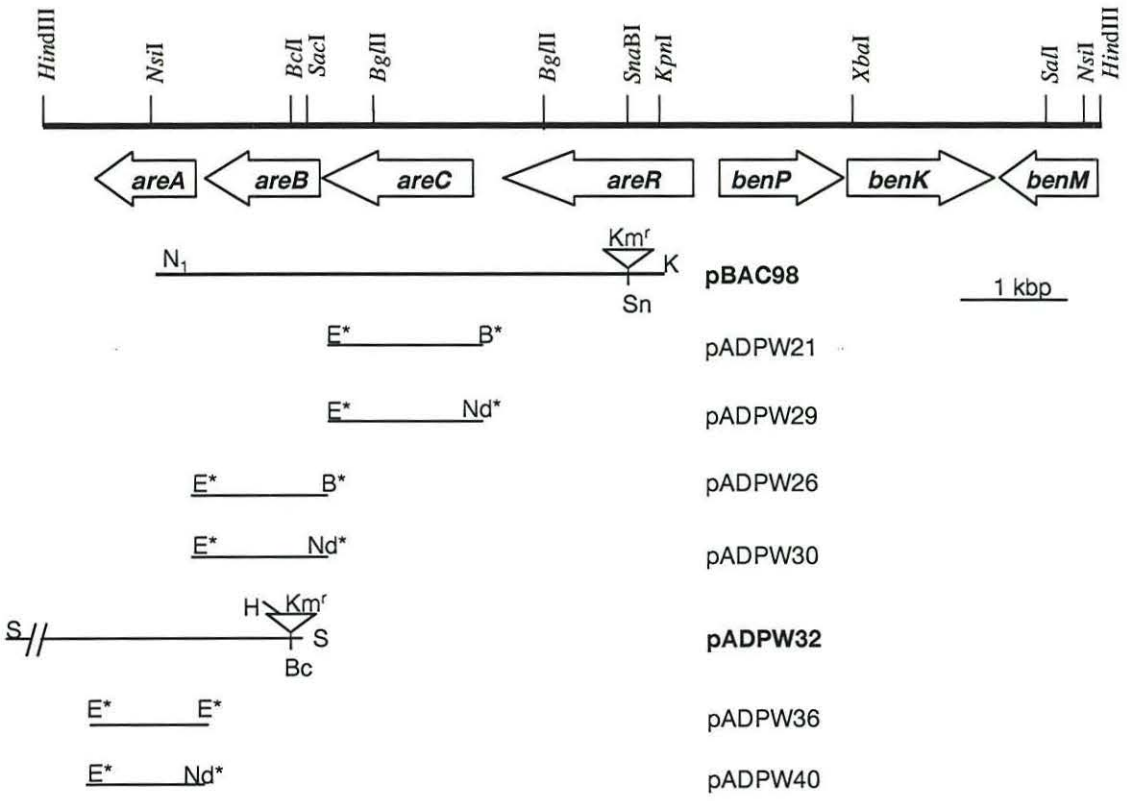


Fig. 4.2.1.. Physical map of the *areCBA* genes and their location relative to one end of the supraoperonic *ben-cat* cluster. The plasmids produced from cloning genomic DNA into vectors are specified in Table 4.2.1. Plasmids named in boldface type contain inserts that were cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The abbreviations for the restriction sites are: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Nsi*I; Nd, *Nde*I; S, *Sac*I; Sa, Sn, *Sna*BI.

TABLE 4.2.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter		
ADP1 (BD413)	Wild type	Juni, 1969
E. coli		
DH5α	F^-, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ^-, *thi-1*, *gyrA96*, *relA1*	Gibco BRL
BL21(DE3)pLysS	F^-, *omp*T, *hsd*S_B(r_B^-, m_B^-), *dcm*, *gal*, (DE3), pLysS, Cm^r	Promega
pET5a	Ap^r, T7 expression vector	Promega
pUC18	Ap^r, cloning vector	Yanisch-Perron, 1982
pBAC98	6.7-kbp *Nsi*I/*Kpn*I fragment cloned from ACN73 containing Km^r cassette and the whole of *areCB* and part of *areR* and *areA*	Jones *et. al.* 1999
pADPW21	1.6-kbp *Bam*HI*/*Sac*I fragment containing *areC* in pUC18	This study
pADPW25	1.4-kbp *Bam*HI*/*Eco*RI* fragment containing *areB* in pUC18	This study
pADPW29	*Nde*I*/*Eco*RI* fragment containing *areC* in pET5a	This study
pADPW30	*Nde*I*/*Eco*RI* fragment containing *areB* in pET5a	This study
pADPW31	0.65-kbp *Sac*I*/*Sal*I* fragment containing part of *areA* in pUC18	This study
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and the whole of *areA* in pUC18	This study
pADPW36	1.2-kbp *Eco*RI* fragment containing *areA* in pUC18	This study
pADPW40	*Nde*I*/*Eco*RI* fragment containing *areA* in pET5a * denotes restriction site added by PCR	This study

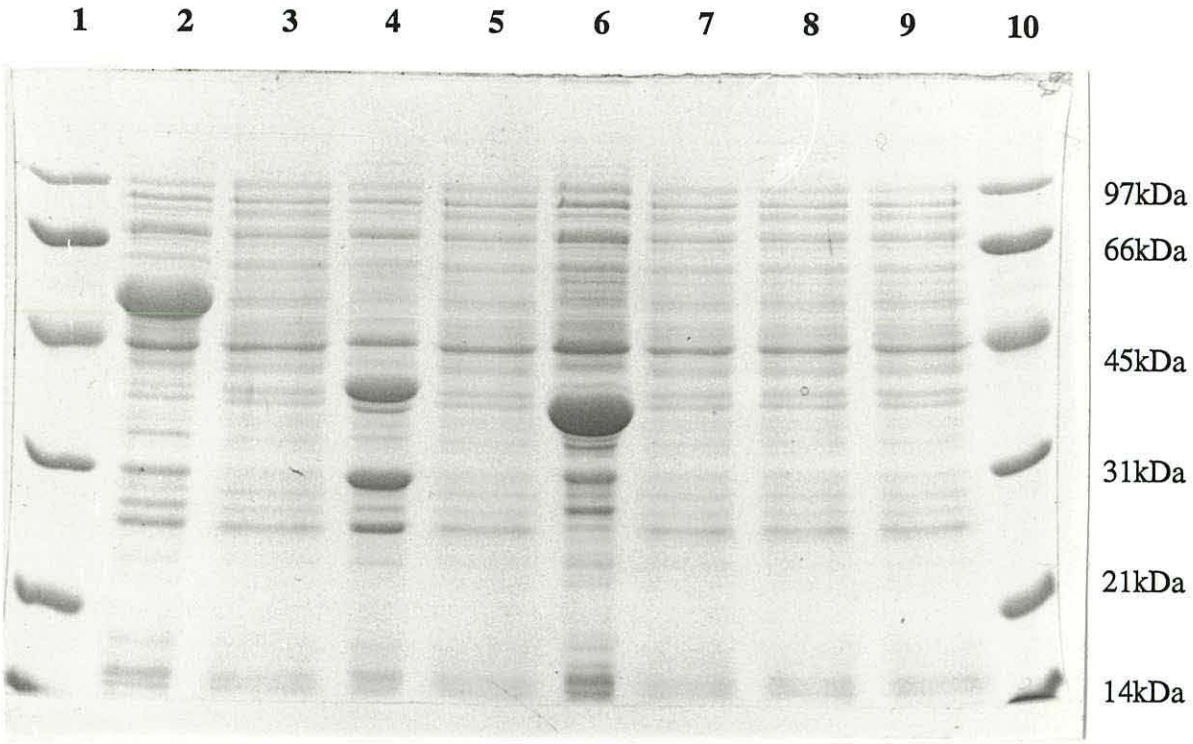


Fig. 4.2.2. SDS-PAGE of over-expressed *are* proteins. The lanes contain lysates from *E.coli* BL21(DE3)pLysS carrying the following plasmids, induced (I) or uninduced (U) with IPTG: Lanes 1, 10, molecular weight standards; Lane 2, pADPW29 (I); Lane 3, pADPW29 (U); Lane 4, pADPW30 (I); Lane 5, pADPW30 (U); Lane 6, pADPW40 (I); Lane 7, pADPW40 (U); Lane 8, pET5a; Lane 9, no plasmid; the molecular weight standards are (in kDa). The estimated M_r values for the overexpressed bands are AreC (lane 2), 52 kDa; AreB (lane 4), 39 kDa with minor components at 32, 27 kDa; AreA (lane 6) 37 kDa. Each lane contains approximately 50μg of protein.

4.2.4. Preparation of ADP1 wild type cell free extract with AreA, B or C activity.

500 ml of *Acinetobacter* sp. ADP1 culture was grown in 2 x 2 l baffled flasks (total of 1 litre culture) in minimal medium salts containing either succinate (10mM), benzaldehyde, benzyl alcohol, 4-hydroxybenzyl alcohol or benzyl acetate (all at 2.5mM) as the sole carbon sources. The cultures were grown at 37°C until mid log phase of growth when they were harvested at 4°C and washed with 0.1M phosphate buffer pH7.4. Cell-free extracts of each culture were prepared as described in Chapter 2.9.

4.3. AreABC enzyme assays.

4.3.1. Preliminary investigations of the benzaldehyde dehydrogenase, benzyl alcohol dehydrogenase, and benzyl esterase assays.

Preliminary tests of the biochemical assays were done as follows. All rates were determined in triplicate at two enzyme concentrations, with two independently prepared cell-free extracts. Test assays confirmed that the rate of NADH production was linear with respect to the enzyme concentrations. The stoichiometry of NADH produced at total conversion of the substrates (i.e. in the presence of excess NADH) was shown to be 1.0 mole/mole. For example, if 100μM benzaldehyde was present in the assay then a net change of 0.622 absorbance units would be expected since the molar extinction coefficient of NADH is 6220 mol^{-1} cm^{-1}. By this analysis method, all substrates assayed were found to be soluble at 1mM. All assay rates were taken as the initial assay velocity over a linear curve. Spectral scans of all assay substrates and products were done to check whether the aromatic compound's spectra absorbed at 340nm, which is the wavelength at which NADH production is followed. 2-, 3-, and 4-hydroxybenzaldehydes and 2- and 3-hydroxybenzoates did absorb at 340nm. Hence, their molar extinction coefficients were measured at 340nm and at pH 9.5 as: 2-hydroxybenzaldehyde (5000 mol^{-1} cm^{-1}), 3-hydroxybenzaldehyde (1500 mol^{-1} cm^{-1}), 4-hydroxybenzaldehyde (18400 mol^{-1} cm^{-1}), 2-hydroxybenzoate (40mol^{-1} cm^{-1}) and 3-hydroxybenzoate (250mol^{-1} cm^{-1}). These values were taken into account when calculating the initial reaction velocity involving these compounds. For example, when one mole of 2-hydroxybenzaldehyde is converted to 2-

hydroxybenzoate, and one mole of NADH produced, the net absorbance change is 6220 – 5000 + 40 = 1260. Hence, the measured reaction velocity (A_{340}/min) must be multiplied by 4.9 (= 6220/1260) to arrive at the actual rate velocity.

4.3.2. Stability of AreABC enzyme activity.

To determine the stability of each of AreA, B and C, a standard assay for each enzyme was set up. The Are enzyme activity was determined immediately following preparation of the cell free extract. Then the cell-free extract was stored at -20°C before being assayed again under the standard conditions. Biochemical assays were replicated within 48 hours. Rate determinations of the stored extracts were within <2% of the freshly prepared cell-free extracts for each Are protein.

4.4. Biochemical assay analysis of the AreB and C proteins from *Acinetobacter* sp. ADP1 and from the cloned enzymes.

4.4.1. Enzyme assays of cloned and the wild type AreC protein.

Cell-free extracts from ADP1 grown on succinate, benzaldehyde, benzyl alcohol, 4-hydroxybenzyl alcohol and benzyl acetate were assayed as described in Chapter 2.9.2. Each cell free extract, save the one prepared from succinate-grown cells showed NAD^+-dependent dehydrogenase activity against benzaldehyde. However the cell-free extract, prepared from mid-log phase cells grown on benzaldehyde produced a low protein yield as compared to the yield obtained when ADP1 was grown on benzyl alcohol or benzyl acetate. It also failed to give adequately reproducible assay results, probably because the cell free extract comprised 10% of the total assay volume. Hence, assays to determine the specificity of the wild type AreC were done only on ADP1 cell-free extracts grown on benzyl alcohol, 4-hydroxybenzyl alcohol and benzyl acetate.

The AreC protein, expressed from plasmid pADPW29, also showed NAD^+-dependent dehydrogenase activity against benzaldehyde. The wild type enzyme prepared from benzyl alcohol-grown cells, from benzyl acetate-grown cells, and from 4-hydroxybenzyl alcohol-grown cells, as well as the cloned and over-expressed AreC were assayed against

a range of methyl- and hydroxy- substituted benzaldehydes. The cell free extracts showed dehydrogenase activity against all the methyl- and hydroxy-substituted benzaldehydes except for 2-methylbenzaldehyde, against which there was no measurable activity (Table 4.4.1.). The best substrate tested was 2-hydroxybenzaldehyde (salicylaldehyde) which was oxidised at a rate 6-fold greater than was benzaldehyde or 3-hydroxybenzaldehyde. Moreover, in all cases, the relative activities of the ADP1 cell-free extracts against different substrates were essentially identical to that of the over-expressed AreC activity although the induced specific activities in ADP1 were lower than in the recombinant *E.coli* by a factor of 10. AreC was induced to considerably higher specific activity when grown on benzyl acetate or benzyl alcohol or 4-hydroxybenzyl alcohol than when grown on succinate.

4.4.2. Biochemical assays of cloned and the wild type AreB protein.

Cell-free extracts from ADP1 grown on benzyl alcohol, 4-hydroxybenzyl alcohol and benzyl acetate, and the AreB protein, expressed from IPTG induced pADPW30 were assayed as described in Chapter 2.9.3. All cell-free extracts, except the one prepared from succinate grown cells showed NAD^+-dependent dehydrogenase activity against benzyl alcohol. The AreB protein from the above cell free extracts was assayed against a range of methyl- and hydroxy- substituted benzyl alcohols. The NAD^+-dependent benzyl alcohol dehydrogenase activity was determined in a linked assay in the presence of excess AreC with compensation made for the twofold production of NADH. The production of two NADH moles/mole benzyl alcohol was taken into account for all substrates except 2-methylbenzyl alcohol, since AreC had no activity against 2-methylbenzaldehyde (section 4.4.1). By linking the alcohol dehydrogenase assay to benzaldehyde dehydrogenase, the initial rate of NADH production was linear over a much longer time period expediting accurate rate determination.

AreB showed significant activity against benzyl alcohol and the methyl- and hydroxy-substituted benzyl alcohols. But unlike AreC, AreB was able to oxidise the 2-methyl analogue at a rate comparable to the other substrates (Table 4.4.2.).

Table 4.4.1. Relative activities of AreC benzaldehyde dehydrogenase against different benzaldehydes in crude extracts of cells grown on different substrates.

Strain	Growth substrate	Relative activity[1] (specific activity in units/mg protein) against benzaldehyde with substituent:						
		None	2-methyl	3-methyl	4-methyl	2-OH	3-OH	4-OH
E.coli (pADPW29)	LB + Amp + IPTG	100 (3.0)	<1	67	29	600	102	80
ADP1	Benzyl alcohol	100 (0.29)	<1	80	33	600	101	117
ADP1	Benzyl acetate	100 (0.16)	<1	80	29	600	101	96
ADP1	4-OH benzyl alcohol	100 (0.27)	<1	80	28	600	105	107
ADP1	Succinate	(<0.001)	ND	ND	ND	ND	ND	ND

[1] Activity relative to benzaldehyde (=100). All reaction rates upon which the relative activities are based were the average of three measurements none of which varied by > 5%. ND, not determined.

Table 4.4.2. Relative activities of AreB benzyl alcohol dehydrogenase against various benzyl alcohols in crude extracts of cells grown on different substrates.

Strain	Growth substrate	Relative activity[1] (specific activity in units/mg protein) against benzyl alcohol with substituent:						
		None	2-methyl	3-methyl	4-methyl	2-OH	3-OH	4OH
E.coli (pADPW30)	LB + Amp + IPTG	100 (2.3)	78	105	107	47	41	18
ADP1	Benzyl alcohol	100 (0.94)	62	99	97	42	32	14
ADP1	Benzyl acetate	100 (0.90)	63	103	106	43	33	11
ADP1	4-OH benzyl alcohol	100 (0.92)	57	106	110	50	51	14
ADP1	Succinate	(<0.001)	ND	ND	ND	ND	ND	ND

[1] Activity relative to benzyl alcohol (=100). All reaction rates upon which the relative activities are based were the average of three measurements none of which varied by > 5%. ND, Not determined.

As was the case with AreC, the relative activities of AreB from the wild type against the substrates were essentially identical to that of the over-expressed AreB activity, and the specific activity in the wild type was lower than in the recombinant *E.coli* by a factor of 2.5. AreB was induced to a higher specific activity when grown on benzyl acetate or benzyl alcohol or 4-hydroxybenzyl alcohol than when grown on succinate where no activity was detected.

4.5. Biochemical assays of cloned AreA protein.

Wild-type ADP1 is able to grow upon benzyl acetate, benzyl propionate, benzyl butyrate, 2-hydroxybenzyl acetate, 4-hydroxybenzyl acetate but not benzyl benzoate. The esterase AreA, cloned on pADPW40 (Table 4.2.1, Fig.4.2.1.), was assayed in two ways: (i) against 4-nitrophenyl esters, using the absorbance at 405 nm of the 4-nitrophenolate anion product (Chapter 2.9.4.) and (ii) against benzyl esters by linkage to NADH production in the presence of excess of both AreB and AreC (Chapter 2.9.5.).

4.5.1. Biochemical assays of cloned AreA protein against 4-nitrophenol esters.

A cell free extract containing AreA from IPTG-induced pADPW40 was prepared as described in Chapter 4.2.3. The AreA protein was assayed at a range of concentrations of 4-nitrophenyl acetate (2C), 4-nitrophenyl butyrate (4C), and against 4-nitrophenyl caproate (6C) by the 4-nitrophenyl ester biochemical assay (Chapter 2.9.4.). The initial velocity was determined over a linear curve, and the rate data results were analysed to determine the kinetic parameters K_m and V_{max} using the Enzpack computer package (Williams & Zaba, 1985). The computed values (Table 4.5.1) show the value of K_m dropping as the chain length of the carboxylic acid moiety of the 4-nitrophenyl ester increases. 4-nitrophenyl acetate is the poorest substrate as can be judged by the higher K_m and lower V_{max} values. The 4-nitrophenyl butyrate was the best substrate tested with higher affinity and higher V_{max} value.

Table 4.5.1. Kinetic parameters for AreA benzyl esterase against p-nitrophenyl esters.

Substrate	Michaelis constant K_m (μM)	Maximum velocity V_{max} (μmoles/min/mg)	Relative specificity constant (V_{max}/K_m)
4-Nitrophenyl acetate	2410 (2260-2720)	220 (207-236)	0.11
4-Nitrophenyl butyrate	61 (53-68)	960 (912-1020)	16
4-Nitrophenyl hexanoate	11 (10.7-12.3)	430 (424-443)	39

*The range in parentheses after each value specifies the 68% confidence limits

4.5.2. Biochemical assays of cloned AreA protein against benzyl ester substrates.

Cell free extracts from ADP1 grown on benzyl alcohol, 4-hydroxybenzyl alcohol and benzyl acetate (prepared as described in Chapter 4.2.4), were tested for AreA activity in the presence of excess of both AreB and AreC linking the esterase to NADH production (Chapter 2.9.5.) against a range of concentrations of benzyl acetate, benzyl propionate, benzyl butyrate and benzyl benzoate. The assay consists of the substrate, NAD^+, and excess amounts of extracts of *E.coli* containing AreC and AreB expressed by IPTG induction of pADPW29 and pADPW30 respectively. The assay was started by the addition of AreA. Hydrolysis of the ester substrate by AreA results in the production of benzyl alcohol that is converted to benzoate by AreB and AreC. This is coupled with the production of two NADH moles/mole of benzyl ester which is taken into account when determining initial velocity. The rate data results were analysed to determine the kinetic parameters K_m and V_{max} using the Enzpack computer package (Williams & Zaba, 1985). The calculated values for K_m (Table 4.5.2.) decrease as the chain length of the carboxylic acid moiety of the ester substrate increases. Benzyl acetate is the poorest substrate as can be judged by the higher K_m and lower V_{max} values. Benzyl butyrate is the best substrate tested with higher affinity and higher V_{max} value.

4.6 Discussion.

Wild-type ADP1 is able to utilise a range of aromatic aldehydes, alcohols and esters as growth substrates. The Are enzymes are expressed during growth on some of these substrates. Assays of benzyl alcohol and benzaldehyde dehydrogenase activities against the same range of substrates were measured after growth of ADP1 on benzyl acetate, benzyl alcohol and 4-hydroxybenzyl alcohol. In all cases, the relative activities against different substrates were essentially identical to that of the over-expressed AreC and AreB activities although the induced specific activities in ADP1 were lower than in the recombinant *E.coli* by factors of 10 and 2.5, respectively. AreB and AreC were induced to considerably higher specific activity when grown on benzyl acetate or benzyl alcohol than when grown on succinate.

Table 4.5.2. Kinetic parameters for AreA benzyl esterase against benzyl esters.

Substrate	Michaelis constant K_m (μM)	Maximum velocity V_{max} (μmoles/min/mg)	Relative specificity constant (V_{max}/K_m)
Benzyl acetate	227 (213-238)	690 (670-712)	3
Benzyl propionate	67 (63-70)	2090 (2020-2140)	31
Benzyl butyrate	42.4 (40-45)	2050 (2010-2090)	47
Benzyl benzoate	47 (43.6-49)	1100 (1070-1130)	23

*The range in parentheses after each value specifies the 68% confidence limits

Esterase assays of the cloned AreA enzyme were done against 4-nitrophenyl esters and benzyl esters. For both sets of ester substrates, the value of K_m dropped as the chain length of the carboxylic acid moiety increased. The acetate esters were the poorest substrates in both classes as could be judged by the lower relative specificity constants (V_{max}/K_m) resulting from both higher K_m and lower V_{max} values. The butyrate esters were the best substrates tested in both classes with higher affinities and higher V_{max} values.

The Are gene function (hypothesised in Fig. 3.6.1.) now extends to include the catabolism of substituted benzyl alcohols and benzaldehydes. Growth tests also showed that ADP1 was able to grow on 2-hydroxybenzyl acetate and 4-hydroxybenzyl acetate (these two compounds were a gift from Alister Boyes, Dept of Chemistry, University of Wales Bangor). Evidence suggests that the *are* genes encode the enzymes for the sequential conversion of 2- or 4-hydroxybenzyl acetates through the corresponding alcohols and aldehydes to 2- or 4-hydroxybenzoates (Fig. 4.6.1.). The 4-hydroxybenzoate feeds the *pob/pca* branch of the β-ketoadipate pathway, while further catabolism of 2-hydroxybenzoate (salicylate) is the subject of chapter 7 of this thesis.

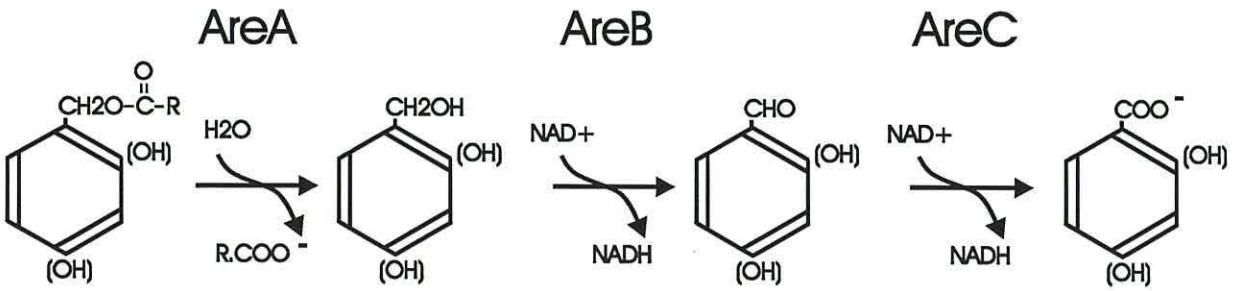


Fig. 4.6.1. Hypothesis for the catabolism of benzyl acetate by reactions catalysed by the *are* gene products.

CHAPTER 5.

The *areABC* genes are transcribed as an operon

5.1 Introduction.

In Chapters 3 and 4, it was shown that the *areCBA* genes encode the enzymes for the conversion of aryl esters or substituted aryl esters to benzoate or substituted benzoates. In this chapter, the regulation of the *areABC* genes will be described. It will also show that the *areCBA* genes are expressed on one transcript of mRNA. The start of transcription of *areCBA* will be reported, as well as a thorough analysis of the sequence upstream of *areC* for promoter regions. In addition, the compounds that are responsible for the induction of the *areCBA* genes will be described.

5.2. Sequence analysis of the *areABC* genes region.

5.2.1. Sequence analysis of the *areABC* intergenic regions.

From the nucleotide sequence, the three genes *areCBA* appear to be transcribed in the same direction and are separated by relatively short intergenic regions (27 bp and 83 bp respectively) (Fig. 5.2.1). The sequence from the ATG translation start of *areC* to 300 bp after the TAA stop codon of *areA* was analysed using the TERMINATOR computer programme from the Wisconsin Genetics Computer Group programs (Devereaux *et al* 1984). There is a small inverted repeat between *areB* and *areA* with a 1 in 7 mismatch:

AGTTTAG ATC CTCAACT

In addition, 11bp after the putative TAA stop codon of *areA* is a perfect inverted repeat with the sequence:

TTAAAAAGGTTCTT AAT AAGAACCTTTTTAA

that is a possible transcription terminator as might be expected at the end of an operon.

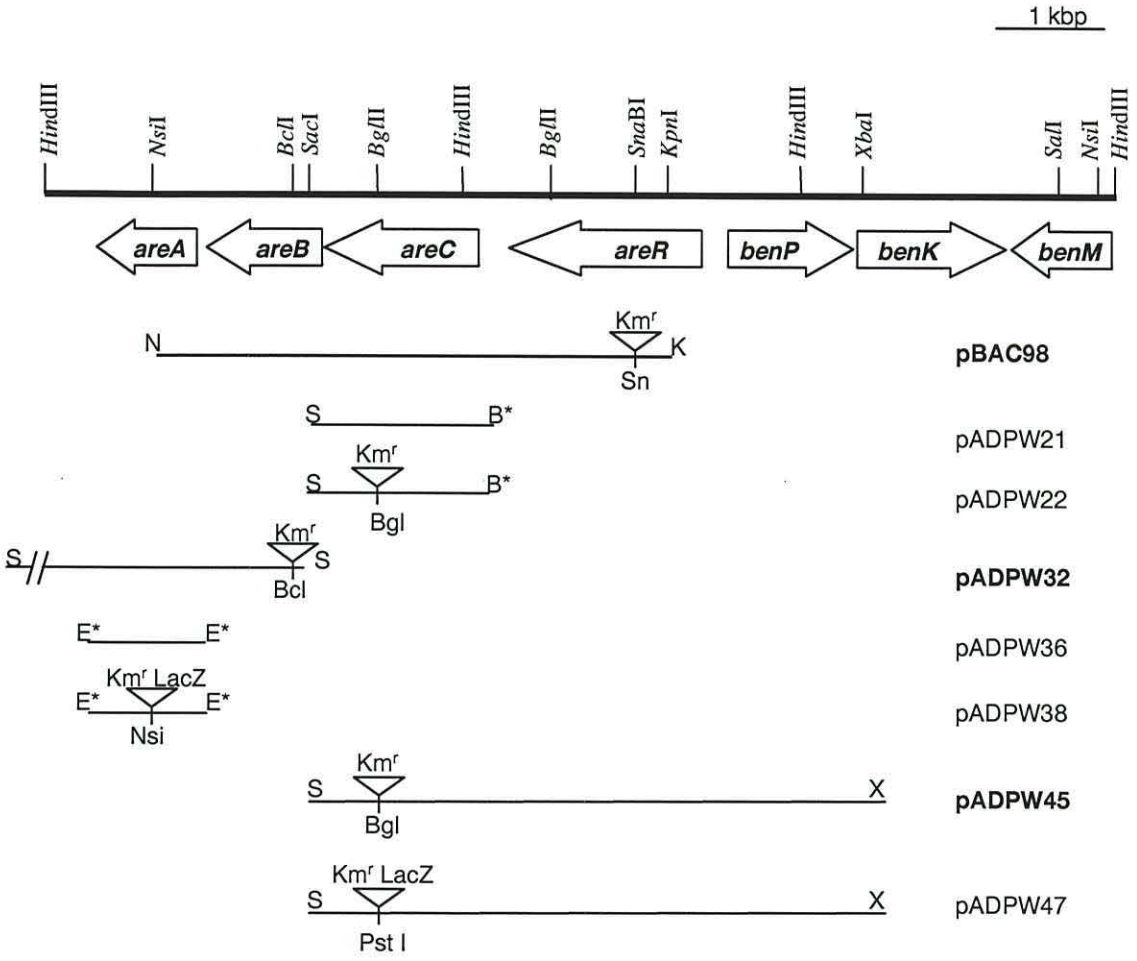


Fig. 5.2.1. Physical map of the DNA adjacent to the *ben* genes in ADP1. Physical map of the *areCBA* genes and their location relative to one end of the supraoperonic *ben-cat* cluster. The various inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 5.6.1. Plasmids named in boldface type contain inserts that were cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The Km^r cassette or lacZ-Km^r cassette insertions are not to scale. The abbreviations for the restriction sites are: B, *Bam*HI; Bc, *BclI*; Bg, *BglII*; E, *Eco*RI; H, *Hin*dIII; N, *Nsi*I; S, *Sac*I; Sn, *Sna*BI.

5.2.2. Sequence analysis of the *areC – areR* intergenic region.

The region between *areR* and *areC* (Fig 5.2.1.) was searched for promoter regions that might serve to promote the transcriptional unit. No promoter region was found for the binding of the primary σ^{70}-holoenzyme (DeHaseth *et al.* 1998). However, a promoter element was found that exactly matches the consensus for the binding of the σ^{54}-holoenzyme (Morett, & Segovia, 1993). From bases -123 to -109 upstream of the start codon of *areC* there is a possible σ^{54}-dependent -12, -24 promoter sequence (TGGCAC-N_5-CTGC) that is well matched to the consensus sequence TGGCAC-N_5-TTGC. This strongly suggests that the *are* genes are regulated by a σ^{54}-dependent transcriptional regulator protein.

Moreover, further upstream of the -12, -24 promoter sequence, from bases –220 to –184, there is a large inverted repeat with only a 1 in 17 mismatch. The inverted repeat sequence also has 7 bp repeat sequence within each arm with a 1 in 7 mismatch.

The repeat sequence is:

ATTGTACATTGTTCAGAATTCTGGACAATGTACAAT

ATTGTACATTGTTCAGA

This type of inverted repeat is characteristic of a regulatory binding site where the activator of the σ^{54}-holoenzyme binds cooperatively to form an oligomeric complex (Shingler, 1996). The sequence analysis is of particular interest since the upstream gene, *areR*, appears to encode a member of the σ^{54}-dependent NtrC type transcriptional regulator family. The deduced AreR amino acid sequence has the conserved regions found in members of this family that function in ATP binding, σ^{54} interaction and DNA binding (Shingler, 1996). The AreR sequence is analysed in detail in Chapter 6 of this thesis.

Sequence analysis evidence suggests that *areC* is at the start of a transcriptional unit that transcribes *areCBA* as a single message. Sequence analysis also suggests that *areCBA* genes are positively regulated by a σ^{54}-dependent transcriptional regulator, to which AreR (encoded by the adjacent open reading frame to *areC*) shares high sequence homology. It is also hypothesised that *areR* probably regulates the *areCBA* transcriptional

unit. The role of *areR* in transcription regulation will also be experimentally analysed in Chapter 6 of this thesis.

5.3. RT-PCR analysis of transcripts of the *areABC* genes.

5.3.1. Preparation of total RNA from *Acinetobacter* sp. ADP1.

In order to show whether the *are* genes are inducible and encoded on a single mRNA as implied by the gene organisation, RT-PCR analysis of *areCBA* was carried out. Total RNA was isolated from *Acinetobacter* sp. ADP1 cells grown on minimal medium containing either benzyl acetate, benzyl alcohol, or succinate using an RNeasy Mini column (Qiagen) as described in Chapter 2.8.1.

5.3.2. RT-PCR of *Acinetobacter* sp. ADP1 Total RNA after growth on benzyl substrates.

RT-PCR was carried out with the Access RT-PCR kit (Promega). Amplifications were carried out across the intergenic regions between *areC* and *areB* using primer pairs (*areCB* forward) 5'-TCAAAGCGCGTGTAATCGAAAAGGTCAAAC-3'; and (*areCB* reverse) 5'-ATGCCCATCTGGATCTCCACCACTGAAGT-3'; and for the intergenic region between *areA* and *areB* with primer pairs (*areBA* forward) 5'-CAGGCGGTGGTGTAAAGTTTGCTCTTGAAT-3'; (*areBA* reverse) 5'-ATTGCCCCCTGCGCTGTCTCCTG-3' (Fig. 5.3.1.a.). Negative control reactions to eliminate the possibility of amplifying residual genomic DNA were performed in the same way, except that the reverse transcriptase was omitted from the reaction mixtures.

The expected amplification product size for the *areC-areB* RT-PCR was 953 bp and for the *areB-areA* RT-PCR was 946 bp. The PCR products obtained, together with restriction digests chosen to confirm the presence of expected restriction sites, were analysed by agarose gel electrophoresis. The products obtained spanning *areCB* and spanning *areBA* from the total RNA of cells grown on both benzyl acetate and benzyl alcohol were of the expected sizes (Fig. 5.3.1.b). The presence of restriction sites in the

Fig. 5.3.1.a. RT-PCR of *are* genes. The position of the restriction sites and the primers used for the RT-PCR are relative to last base of *areA*.

Fig. 5.3.1.b. Agarose gel electrophoresis of RT-PCR products. RT-PCR amplification by primers from ADP1 grown on benzyl acetate and benzyl alcohol. The sizes of molecular size markers (in bp) in lanes S (HyperLadder I, from Bioline, London, UK) are indicated by arrows (1,500 1,000, 800, 600, 400 and 200 bp). Lanes: 1: *areCB* benzyl acetate-grown cells (expected size, 953 bp); 2: *areCB*, benzyl acetate-grown cells digested with *Sac*I (643 and 310 bp); 3: *areCB*, benzyl alcohol grown cells (expected size, 953 bps); 4: *areCB*, benzyl alcohol-grown cells digested with *Sac*I (643 and 310 bps); 5: *areBA*, benzyl acetate-grown cells (expected size, 946 bp); 6: *areBA*, benzyl acetate grown cells digested with *Acc*I (622 and 324 bp), 7: *areBA*, benzyl alcohol-grown cells (expected size, 946 bp); 8: *areBA*, benzyl alcohol-grown cells digested with *Acc*I (622 and 324 bp). No detectable products were obtained in control reactions, with each pair of primers, from which reverse transcriptase had been omitted or in reactions carried out on succinate-grown cells (data not shown).

Fig. 5.3.1.a

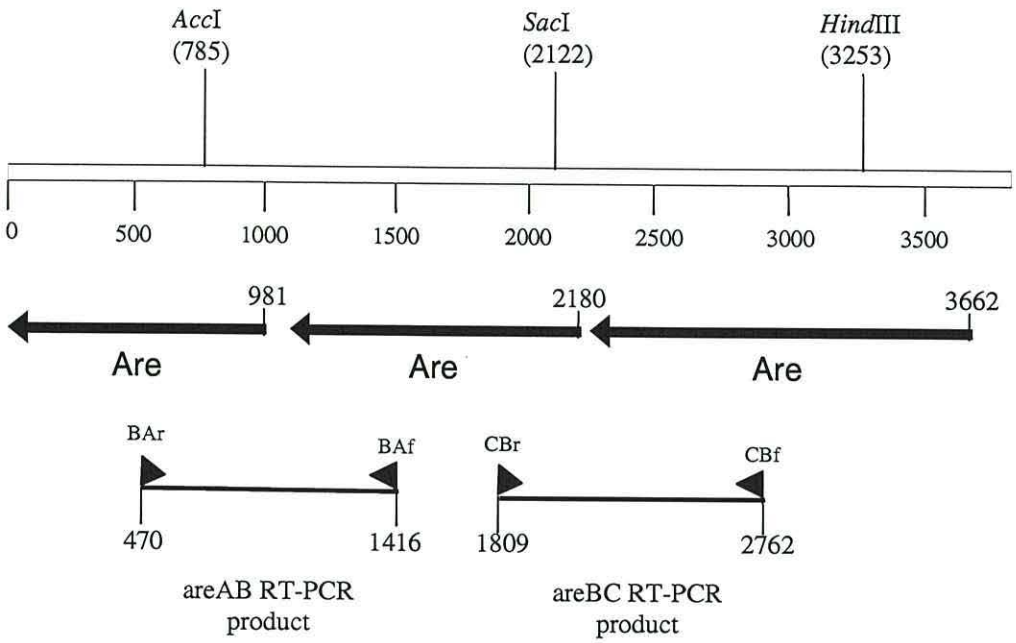
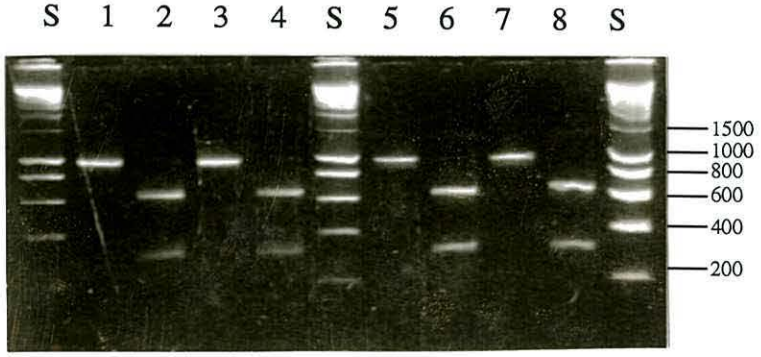


Fig.5.3.1.b



expected positions within the fragments was confirmed by digestion with *Sac*I (*areCB)* and *Acc*I (*areBA*). *areCB* digested with *Sac*I resulted in 643 bp and 310 bp fragments, and *areBA* digested with *Acc*I resulted in 631 bp and 315 bp fragments. No products were obtained from total RNA of succinate-grown cells or from reaction mixtures from which the reverse transcriptase had been omitted.

5.4. Primer extension to determine the transcription initiation site of the *are* genes.

5.4.1. Preparation of total RNA from *Acinetobacter* sp. ADP1.

Total RNA was prepared with RNeasy Mini columns (Qiagen). Ten micrograms of total RNA was prepared from ADP1 grown in minimal media containing benzyl acetate, benzyl alcohol or succinate respectively.

5.4.2. Primer extension of *Acinetobacter* sp. ADP1 Total RNA.

Primer extension reactions were performed as described in Chapter 2.8.3. The sequence of the 5'-labelled *areC* specific primer is 5'-ATCAAGTAATGT**CAT**ATAGACCTCGTA-3'. The primer-binding site overlaps the putative start codon of *areC* that is highlighted in bold type. Unfortunately the same primer was unsuccessful in obtaining the corresponding sequence ladder. Thus, a sequence ladder was prepared by sequencing pUC18 with a universal primer by the dideoxy chain termination method (Pharmacia).

Primer extension results show that the transcription start site is located 97 nucleotides upstream of the ATG (Figs. 5.4.1.a; 5.4.1.b.). The conserved -12, -24 consensus element for a σ^{54} promoter binding site is located further upstream from the transcription start. The genetic organisation of the *areR* – *areA* intragenic region including the start codon of *areC* and the putative ribosome binding site (RBS) of *areC* can also be deduced (Fig.5.4.2).

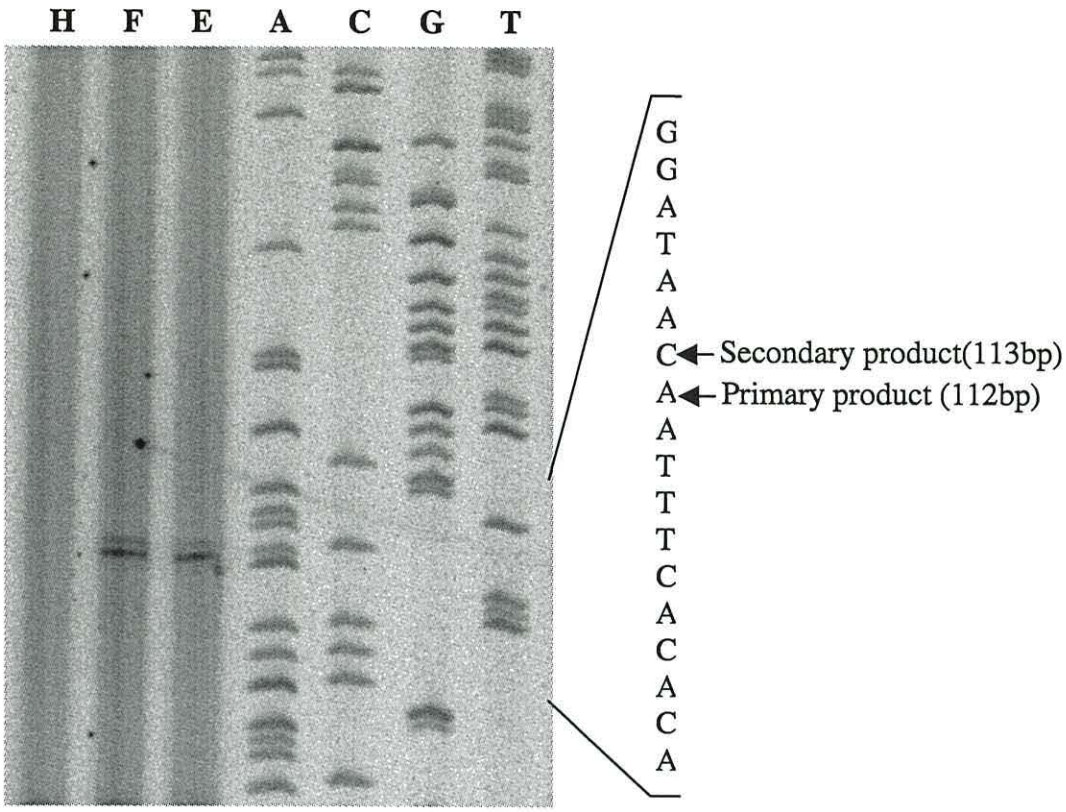


Fig. 5.4.1.a. Primer extension of *areC* mRNA from Acinetobacter sp. ADP1. The experiment was done with total RNA and a primer that overlaps the putative start codon of *areC*. Lanes A, C, G, and T show the respective sequencing products from pUC18 sequenced using an universal primer. Lane E, F, and H represent the signal obtained from the Primer Extension experiment using 10 µg of total RNA from *Acinetobacter* sp. ADP1 cells grown with benzyl acetate (E), benzyl alcohol (F) and succinate (H) as the sole carbon growth source. An arrow indicates the positions of the two signals produced.

Primary product (112bp)

AATTGTGAGCGGATAACAATTTCACACAGGAAACA-$(N)_{77}$- **ACTGGCCGTCGTTTTAC**-5`

Fig. 5.4.1.b. The pUC18 sequence template used as marker for primer extension. The universal primer is highlighted in bold text, and the base corresponding to the primary primer extension product is in bold text. The underlined sequence corresponds to the sequence in Fig. 5.4.1.a.

Fig. 5.4.2. The intergenic region between *areR* and *areC*. The start codon of *areC* is indicated, and the putative ribosome binding site (Shine-Delgano) of *areC* is indicated above the sequence(RBS). The nucleotide corresponding to the primer extension signal (Fig. 5.4.1.a.) is highlighted at the +1 site. The putative –12,-24 promoter elements are indicated. Arrows below the sequence mark an inverted repeat that is hypothesised to act as the binding site for the regulator protein.

5.5. Biochemical assays to determine the inducibility of the *areABC* genes.

5.5.1. Preparation of *Acinetobacter* sp. ADP1 grown on benzaldehyde, benzyl alcohol, or benzyl acetate.

Cell free extracts were prepared of *Acinetobacter* sp. ADP1 cells grown on minimal media containing benzyl aceate, benzyl alcohol, benzaldehyde, or succinate. The cultures were grown at 37°C until mid log phase of growth when they were harvested at 4°C and washed with 0.1M phosphate buffer pH7.4. Cell free extracts of each culture were prepared as described in Chapter 2.9.

5.5.2. Assays for AreC, AreB, and AreA activity in *Acinetobacter* sp. ADP1 grown on benzaldehyde, benzyl alcohol, or benzyl acetate.

The activities of AreA, AreB, and AreC were assayed by the NAD^+ linked assays described previously (Chapter 2.9). AreA, AreB, and AreC specific activities were measured in crude cell free extracts from ADP1 cells grown on benzyl acetate, benzyl alcohol, or benzaldehyde. Two independent preparations of each cell free extract were assayed in triplicate at two enzyme concentrations, against benzyl acetate, benzyl alcohol, or benzaldehyde as substrates (Table 5.5.1). No AreB or AreC activity was detected in succinate-grown cells. The AreA specific activity measured in succinate-grown cells was >500 fold lower than when grown on any of the aromatic substrates. There was increased AreA activity when cells were induced by benzyl alcohol or benzaldehyde, and similarly, there was increased AreB activity when cells were induced by benzaldehyde. These data indicate that all three aromatic substrates induce all three *are* genes, with benzaldehyde being the poorest inducer, and benzyl acetate the best.

5.6. The creation of *lacZ* transcriptional fusions of *areA* and *areC* in ADP1.

Expression analysis of *areA* and *areC* was performed by measuring the β-galactosidase activity expressed from a chromosomal *areA*::lacZ transcriptional fusion in mutant strain ADPW61, and a *areC*::lacZ transcriptional fusion in mutant strain ADPW63. The reporter strains were constructed by insertion of the *lacZ*-Km^r cassette from pKOK6.1

Table 5.5.1. Activities of the Are proteins from Acinetobacter sp. ADP1 against *are* operon substrates.

Specific activities of AreA (Benzyl esterase), AreB (Benzyl alcohol dehydrogenase), and AreC (Benzaldehyde dehydrogenase) against benzyl acetate, benzyl alcohol, and benzaldehyde respectively in crude extracts of *Acinetobacter* sp. ADP1 cells grown on benzyl acetate, benzyl alcohol, benzaldehyde, or succinate.

	Specific activity (U/mg of protein) against benzyl substrates		
ADP1 Growth Substrate	Benzaldehyde Dehydrogenase	Benzyl alcohol Dehydrogeanase	Benzyl Esterase
Succinate	<0.04	<0.04	<0.04
Benzaldehyde	0.13	0.24	2.55
Benzyl alcohol	0.10	2.27	6.10
Benzyl acetate	0.18	2.82	9.01

[a]All reaction rates were based on the average of three measurements , none of which varied by >5%.

into each gene. The promoterless lacZ gene is preceded by stop codons in all reading frames (Kokotek & Lotz, 1989).

5.6.1. Insertion of the *lacZ*-Km^r cassette into *areA* and *areC*.

To insert the promoterless lacZ-Km^r cassette into *areA*, pKOK6.1 was digested with *Pst*I, whose recognition sites flank the *lacZ*-Km^r cassette, and excises the whole cassette excised on a 4.7 Kb fragment. The cassette was then introduced into the single *Nsi*I site of pADPW36, to create pADPW38 (Fig. 5.2.1), since *Nsi*I produces overhanging ends compatible to those produced by *Pst*I.

For the insertion of the *lacZ*-Km^r cassette into *areC*, genomic DNA from ADPW56 (*areC*::Km^r) was digested with *Xba*I and *Sac*I and ligated into pUC18. Transformants were selected on ampicillin and kanamycin. Restriction digests proved that the plasmid, designated pADPW45, contained a 6.2 Kbp fragment that contained the whole of *areR* and an *areC*::Km^r fusion (Fig. 5.2.1). The Km^r cassette in *areC* was then replaced by *lacZ*-Km^r cassette on a *Pst*I fragment. This was done because the Km^r pUC4K cassette (Table 5.6.1.) in *areC* was inserted on *Bam*HI/*Bgl*II compatible overhangs (Chapter 3.3.1.). However, this results in the loss of the recognition sequence for both enzymes. Nevertheless, there remains the *Pst*I site adjacent to the altered *Bam*HI site, on which Km^r may be excised. However, there is a *Pst*I recognition site in the multi-cloning site of pADPW45. Hence, the *Xba*I/*Sac*I fragment from pADPW45 was cloned into pUC18NP (No *Pst*I) to create pADPW46 (Table 5.6.1.). Then, the *lacZ*-Km^r cassette from pKOK6.1 was also digested with *Pst*I and introduced into pADPW46 as replacement to the Km^r cassette in *areC* to create pADPW47 (Fig 5.2.1).

5.6.2. Transformation of the lacZ transcriptional fusions into *Acinetobacter* sp. ADP1.

Both pADPW38 and pADPW47 were linearised and transformed into ADP1 respectively to create ADPW61 (*areA*::*lacZ*-Km^r) and ADPW63 (*areC*::*lacZ*-Km^r). The integration of

TABLE 5.6.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter		
ADP1 (BD413)	Wild type	Juni, 1969
ADPW57	*areB*::Km^r; transformation of ADP1 with pADPW26	This study
ADPW61	*areA*::lacZ-Km^r; transformation of ADP1 with pADPW38	This study
ADPW63	*areC*::lacZ-Km^r; transformation of ADP1 with pADPW47	This study
E. coli DH5α	F$^-$, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	Gibco BRL
XL1-Blue MRF'	Δ (*mcrA*)*183*, Δ(*mcrCB-hsdSMR-mrr*)*173*, *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, *lac* [F' *proAB lacI^{q}Z* ΔM15 Tn*10* (Tet^r)]	Stratagene
pUC18	Ap^r, cloning vector	Yanisch-Perron, 1982
pUC18NP	pUC18 without *Pst*I in the MCS	Williams 1997
pUC4K	Ap^r, Km^r, source plasmid for Km^r cassette	Vierira, 1982
pKOK6.1	Ap^r Km^r; promoterless lacZ	Kokotec,1989
pBAC98	6.7-kbp *Nsi*I/*Kpn*I fragment cloned from ACN73 containing Km^r cassette and the whole of *areCB* and part of *areR* and *areA*	Jones, 1999
pADPW21	1.6-kbp *Bam*HI*/*Sac*I fragment containing *areC* in pUC18	This study
pADPW22	pADPW21 with Km^r cassette from pUC4K cloned into *Bgl*II site in *areC*	This study
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing	This study

	Km^r cassette and the whole of *areA* in pUC18	
pADPW36	1.2-kbp *Eco*RI* fragment containing *areA* in pUC18	This study
pADPW38	pADPW36 with lacZ-Km^r cassette from pKOK6.1 cloned into *Nsi*I site in *areA*	This study
pADPW45	6.2-kbp *Hind*III fragment cloned from ADPW56 containing *areR* and *areC::*Km^r in pUC18	This study
pADPW46	6.2-kbp *Hind*III fragment cloned from ADPW56 containing *areR* and *areC::*Km^r in pUC18NP	This study
pADPW47	pADPW46 with lacZ-Km^r cassette from pKOK6.1 cloned into *Pst*I of the Km^r cassette	This study
	* denotes restriction site added by PCR	

Fig. 5.6.1. Southern hybridisation analysis to demonstrate that the lacZ-kanamycin cassette has recombined into *areA* to create ADPW61 (*areA*:*lacZ*).

Fig. 5.6.1.a. Map of the DNA sequence upstream and downstream of *areA*, pADPW36 (1.2kbp), and pADPW38 (*areC*::Km^r 5.9kbp). ADPW61 was created by transformation of pADPW38 into ADP1. ADPW61 has a chromosomal rearrangement where the 4.7kbp lacZ-Km^r cassette is integrated into the genome at the *Nsi*I restriction site. Restriction sites marked with an asterisk are on the plasmid multi cloning site.

Fig. 5.6.1.b. The agarose gel blot was hybridised with an *areA* probe purified by band extraction from *Eco*RI digested pADPW36. Lane S contains 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *areA* are shown (kbp):

Lane 1: pADPW36 (*Eco*RI), 1.0.
Lane 2: pADPW38 (*Cla*I), 5.1.
Lane 3: ADP1 (*Cla*I), 2.8.
Lane 4: ADPW61 (*Cla*I), 2.3, 1.6 (undigested aproximately 4.5).

Fig. 5.6.1.a

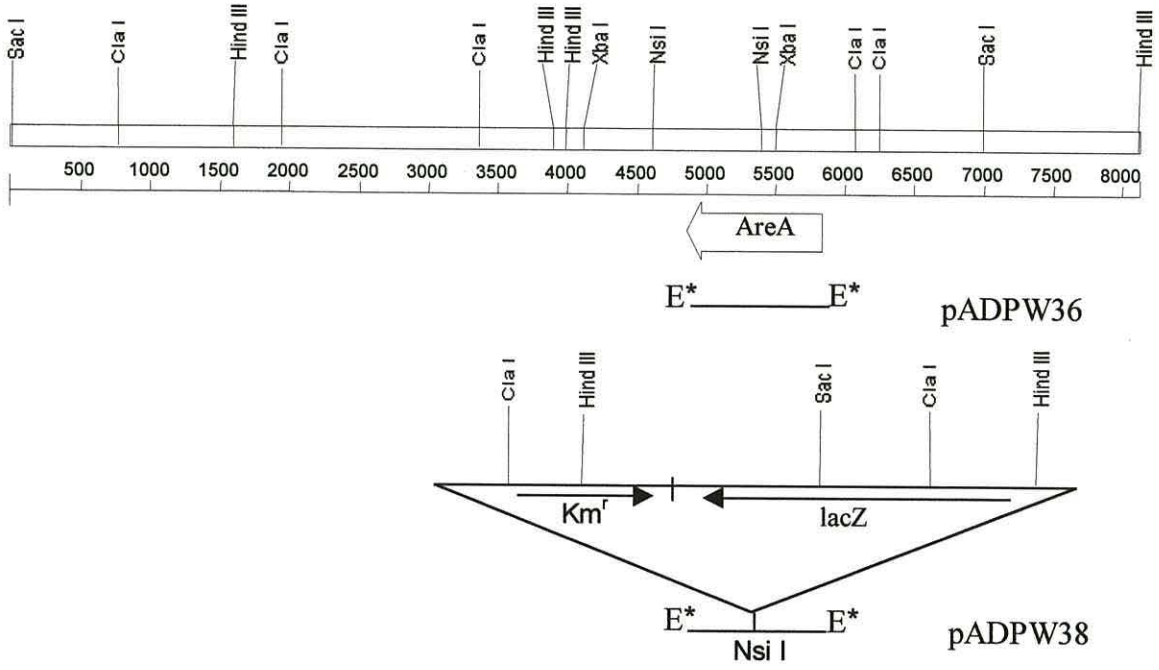


Fig. 5.6.1.b

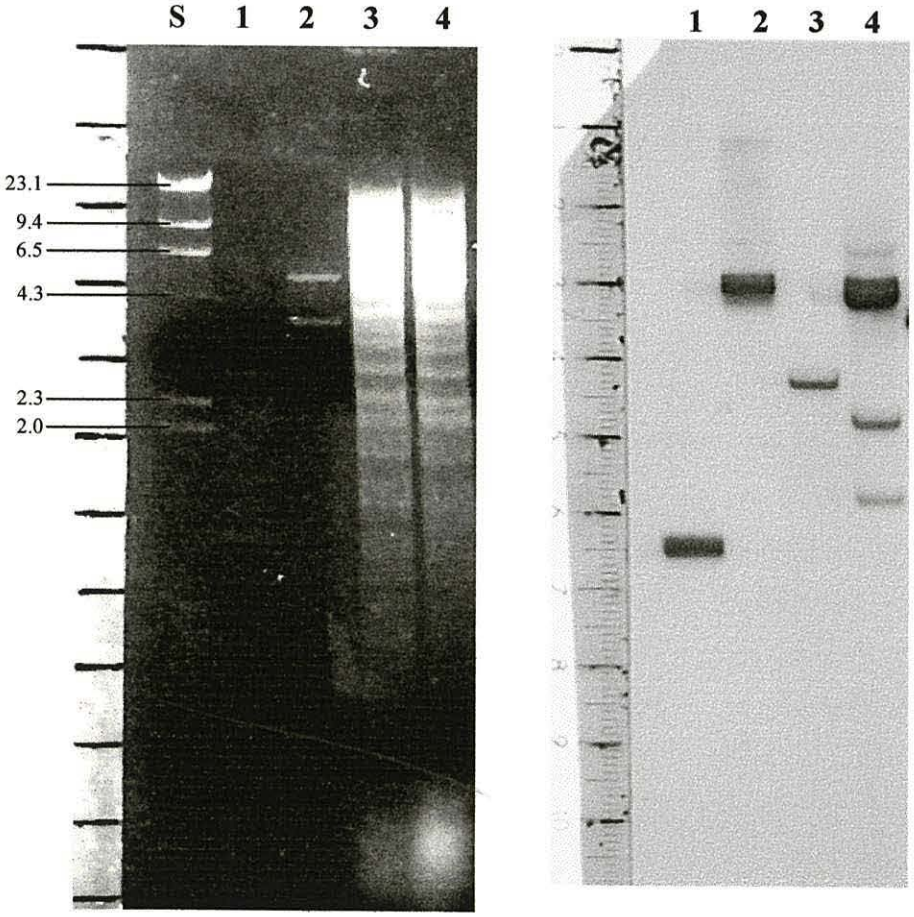


Fig. 5.6.2. Southern hybridisation analysis to demonstrate that the lacZ-kanamycin cassette has recombined into *areC* to create ADPW63.

Fig. 5.6.2.a. Map of the DNA sequence upstream and downstream of *areC*, pADPW45 (6.2kbp), and pADPW47 (*areC*::Km^r 10.9kbp). ADPW63 was created by transformation of pADPW47 into ADP1. ADPW63 has a chromosomal rearrangement where the 4.7kbp lacZ-Km^r cassette is integrated into the genome at the *Bgl*II restriction site within *areC*.

Fig. 5.6.2.b. The agarose gel blot was hybridised with an *areC* probe purified by band extraction from *Eco*RI/*Bam*HI digested pADPW21 (Fig.5.2.1.). Lanes M contain 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *areC* are shown (kbp):

Lane 1: pADPW45 (*Sac*I/*Kpn*I) 8.9. (band at 4.4. due to partial digest by *Kpn*I)
Lane 2: pADPW47 (*Sac*I/*Kpn*I) 3.4., 4.7. (band at 7.2. due to partial digest by *Kpn*I)
Lane 3: ADP1 (*Sac*I/*Kpn*I) 3.3.
Lane 4: ADPW61 (*Sac*I/*Kpn*I) 4.6., 3.4.

Lane 5: pADPW45 (*Cla*I) 4.8., 1.5.
Lane 6: pADPW47 *Cla*I) 4.6., 2.3.
Lane 7: ADP1 (*Cla*I) 2.5.
Lane 8: ADPW61 (*Cla*I) 2.2., 1.5.

Fig. 5.6.2.a.

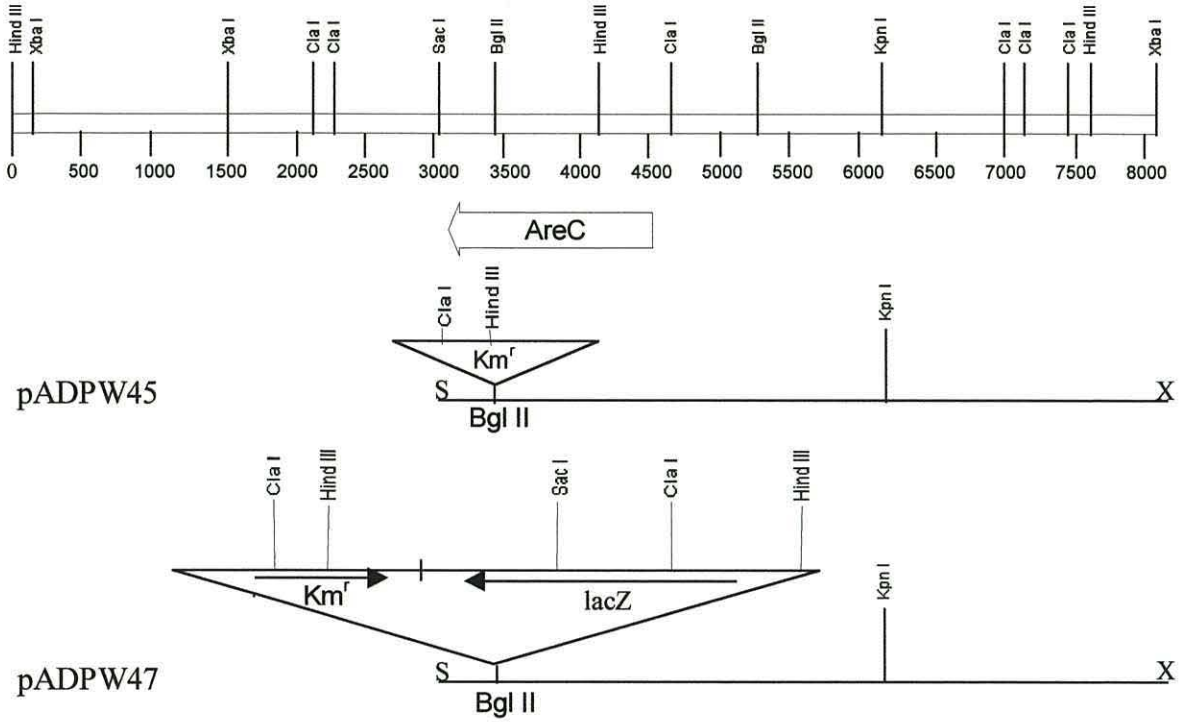
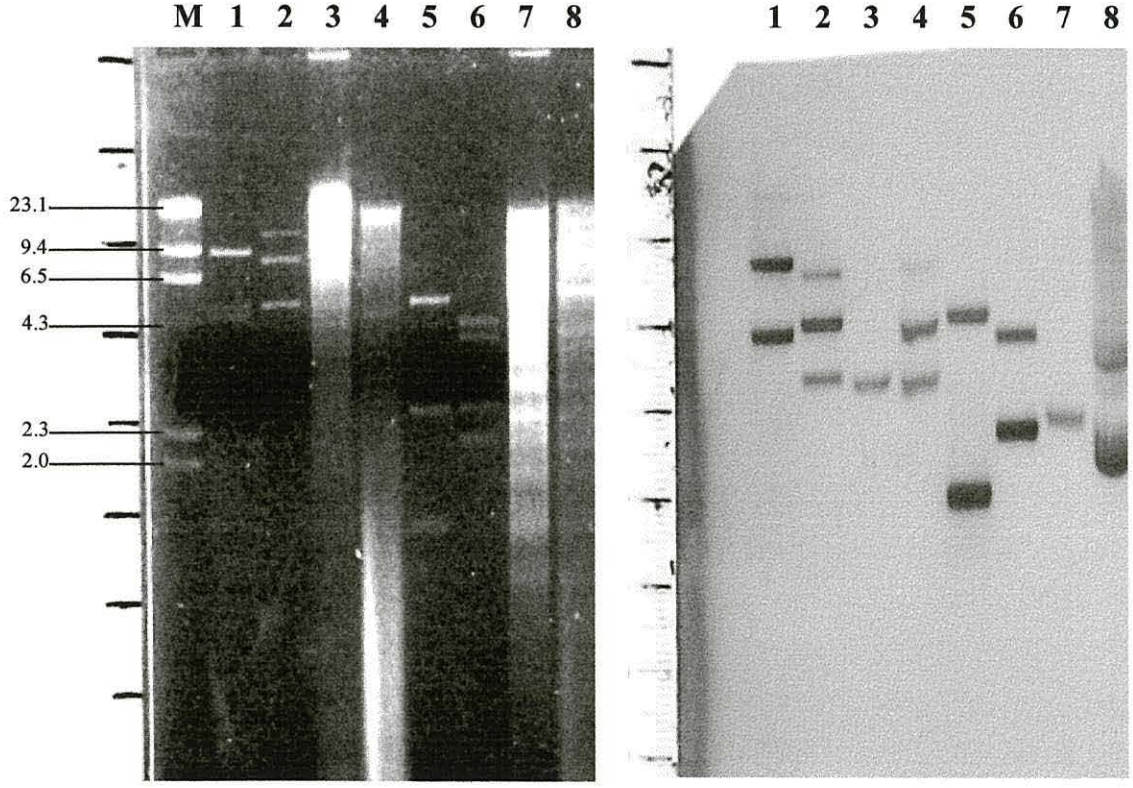


Fig. 5.6.2.b.



the cassette into the chromosomes of ADPW61 and ADPW63 was confirmed by Southern Blot analysis (Figs. 5.6.1; 5.6.2.).

It is also noted that the *lacZ*-Km^r cassette was also inserted into the sole *BglII* site of pADPW21 (Fig. 3.3.1) creating pADPW23. However, when pADPW23 was linearised and transformed into ADP1, the *lacZ*-Km^r cassette would not integrate into the chromosome for an undiagnosed reason (data not shown).

5.7. β-galactosidase activity of *lacZ* transcriptional fusions of *areA* and *areC* in ADP1.

5.7.1. Induction of *areA* and *areC* expression by *are* operon substrates in strains ADPW61 and ADPW63.

Acinetobacter ADPW61 and ADPW63 cultures were grown to late log phase on succinate in the presence of inducers, which was necessary due to the gene disruptions. β-Galactosidase activities were determined as described in Chapter 2.11. Benzyl acetate, benzyl alcohol, and benzaldehyde induced transcription of *areA* to about the same level (Table. 5.7.1) (Fig.5.7.1). Neither benzoate nor succinate induces *areA*. Also, benzyl acetate, benzyl alcohol, and benzaldehyde induced transcription of *areC*. Succinate does not induce *areC*, and benzoate induces *areC* about 2-fold of the succinate level (Table. 5.7.1) (Fig.5.7.2.). Moreover, *areC* is induced approximately 8 to 10 fold higher than is *areA* by the same substrates.

5.7.2. Inducer specificity of expression of *areA* and *areC* in strains ADPW61 and ADPW63.

The inducibility of the expression of the *areA-lacZ* and *areC-lacZ* chromosomal fusions in strains ADPW61 and ADPW63 with a variety of different esters was tested by assaying induced cells for β-galactosidase activity (Table 5.7.2.) (Fig. 5.7.3.). Benzyl acetate, 2-hydroxybenzyl acetate, and 4-hydroxybenzyl acetate induced transcription of *areA* to about the same level. The level of *are* gene induction decreased as the length of the aliphatic chain of benzyl esters increased. Benzyl acetate (2C) induced *areA* to about

Table 5.7.1. β-galactosidase activity expressed from *are*::*lacZ* chromosomal fusions. Induction of the *are* genes of *Acinetobacter* sp. ADP1 in the presence of *are* operon substrates was measured by the level of β-galactosidase activity as expressed in Miller Units (Miller, 1972). Values are averages of three independent trials conducted in duplicate. Standard deviations of the means are in parentheses. All chemicals were added to a final concentration of 1 mM.

Inducer	β-Galactosidase activity	
	ADPW61 (*areA*::LacZ)	ADPW63 (*areC*::LacZ)
Uninduced	44.0 (±21)	658 (±222)
Benzoate	136 (±45)	1548 (±365)
Benzaldehyde	1050 (±145)	9200 (±2200)
Benzyl alcohol	950 (±125)	8450 (±1900)
Benzyl acetate	1000 (±150)	8200 (±1600)

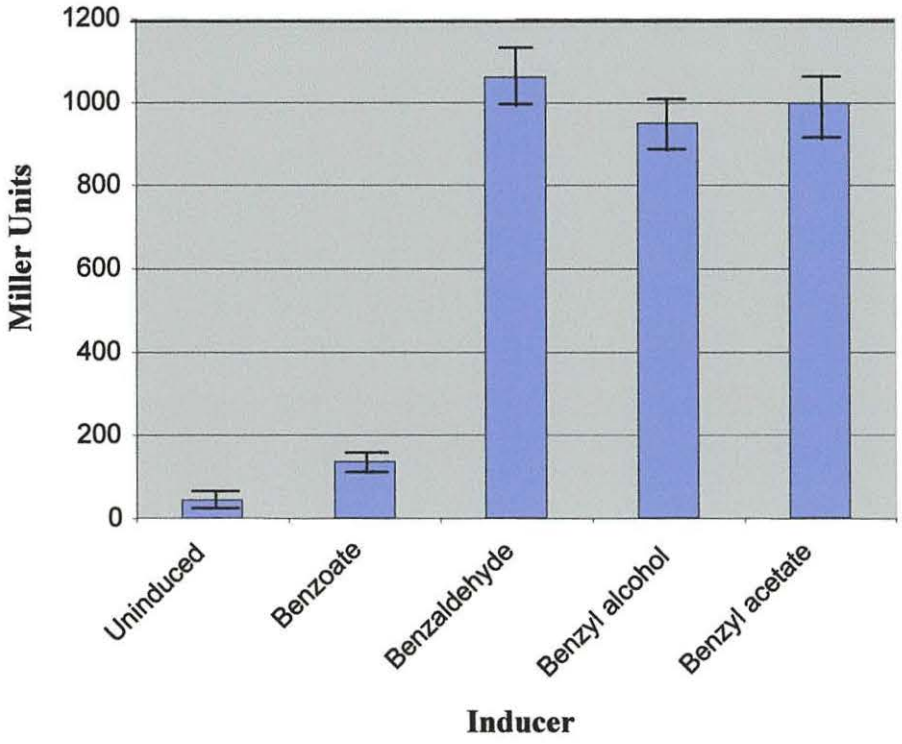


Fig 5.7.1. Expression of β-Galactosidase activity from a chromosomal *areA*::lacZ fusion in strain ADPW61. Strains were cultured in LB in the presence or absence of 1mM inducers. Activities are averages of three independent trials conducted in duplicate. Error bars represent standard deviations.

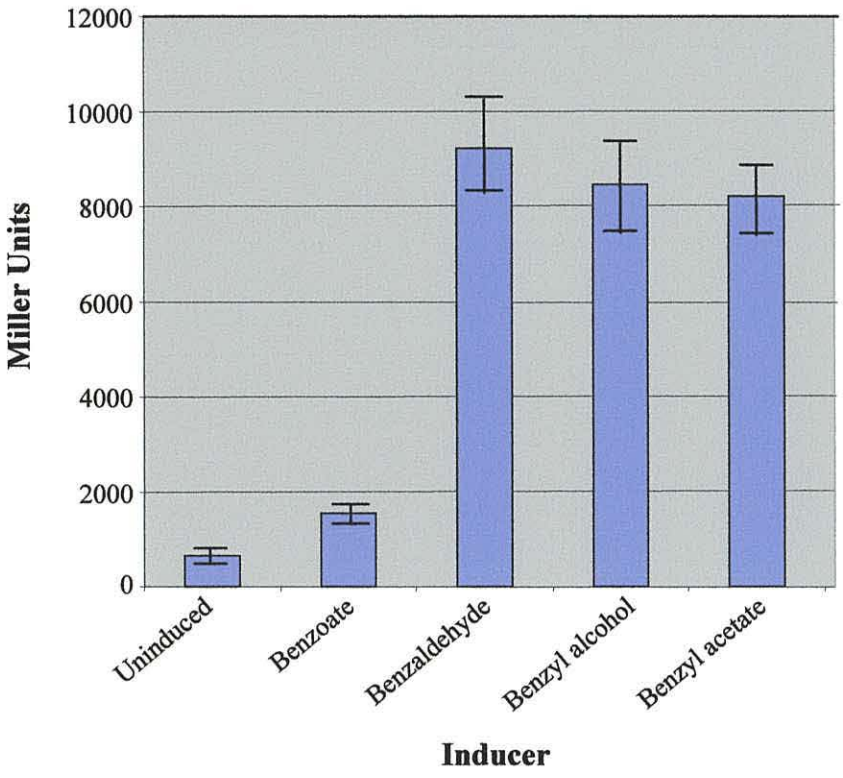


Fig 5.7.2. Expression of β-Galactosidase activity from a chromosomal *areC*::lacZ fusion in strain ADPW63. Strains were cultured in LB in the presence or absence of 1mM inducers. Activities are averages of three independent trials conducted in duplicate. Error bars represent standard deviations.

Table 5.7.2. β-galactosidase activity expressed from *are*::*lacZ* chromosomal fusions. Induction of the *are* genes of *Acinetobacter* sp. ADP1 in the presence of ester substrates. The level of β-galactosidase activity is expressed in Miller Units (Miller, 1972). Values are averages of three independent trials conducted in duplicate. Standard deviations of the means are in parentheses. All chemicals were added to a final concentration of 1.0 mM.

Inducer	β-Galactosidase Activity	
	ADPW61 (*areA*::LacZ)	ADPW63 (*areC*::LacZ)
Uninduced	33 (±15)	500 (±100)
Benzyl acetate	1190 (±130)	8320 (±636)
Benzyl propionate	728 (±65)	4150 (±701)
Benzyl butyrate	375(±50)	2100 (±350)
Benzyl benzoate	450(±20)	3800 (±500)
2-hydroxybenzyl acetate	1000 (±50)	8040 (±820)
4-hydroxybenzyl acetate	1100 (±200)	8400 (±1000)

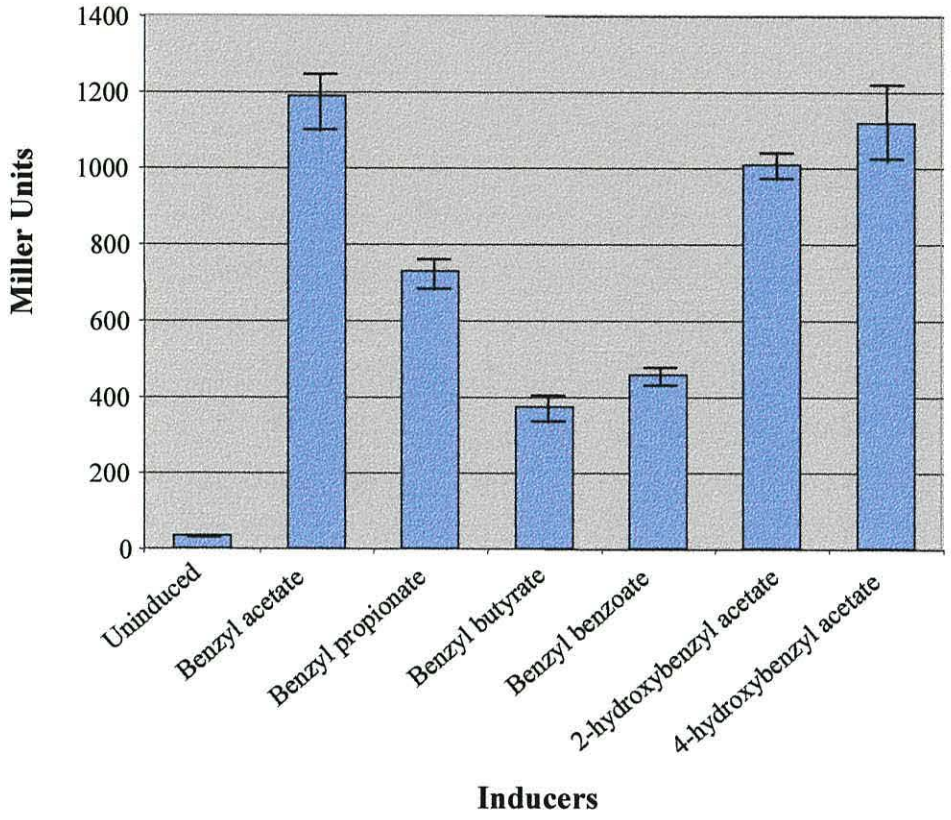


Fig. 5.7.3. Expression of β-Galactosidase activity from a chromosomal *areA*::lacZ fusion in strain ADPW61. Strains were cultured in LB in the presence or absence of 1mM ester substrates of *Acinetobacter* sp.ADP1. Activities are averages of three independent trials conducted in duplicate. Error bars represent standard deviations.

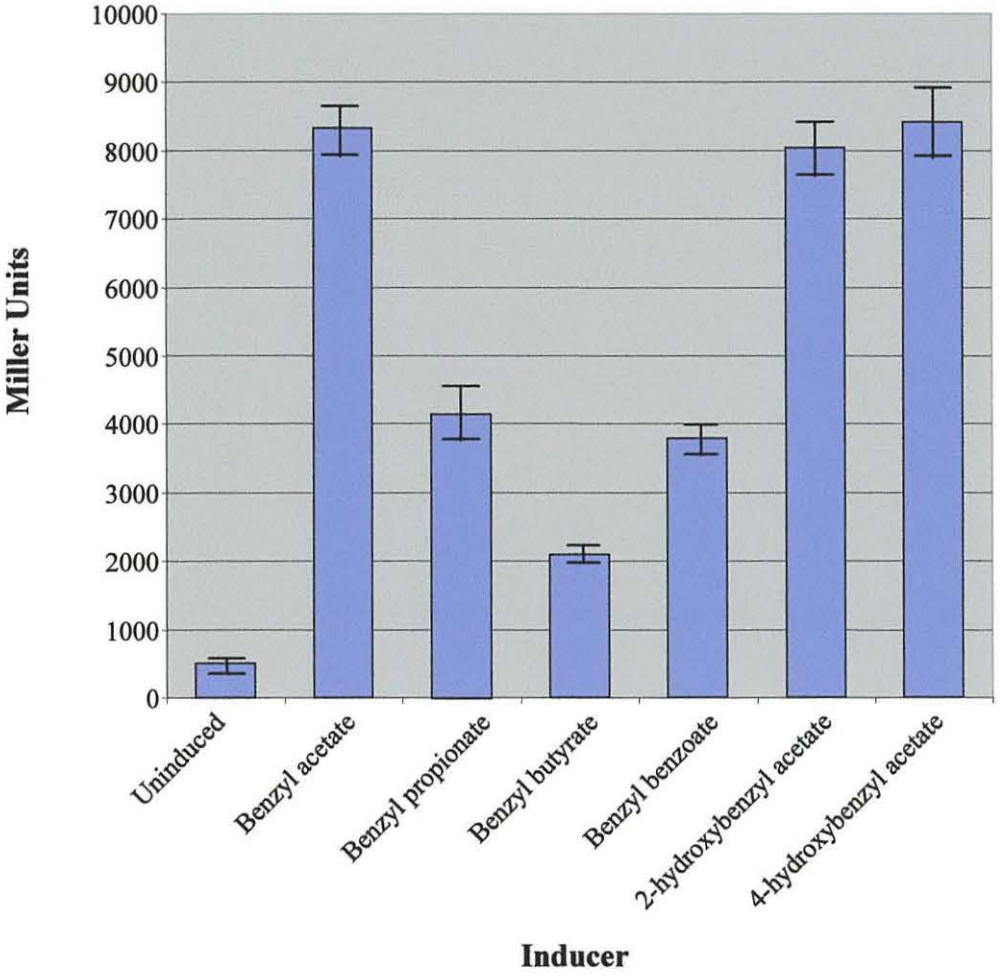


Fig 5.7.4. Expression of β-Galactosidase activity from a chromosomal *areC*::*lacZ* fusion in strain ADPW63. Strains were cultured in LB in the presence or absence of 1mM ester substrates of *Acinetobacter* sp.ADP1. Activities are averages of three independent trials conducted in duplicate. Error bars represent standard deviations.

twofold the induction level of benzyl butyrate (4C). Benzyl benzoate, which is not a growth substrate for ADP1, induces *areA* to half the activity compared with benzyl acetate.

5.8. Discussion.

Studies of the *areABC* genes by RT-PCR, proved that a single message is transcribed across the three genes when ADP1 is grown on benzyl alcohol or benzyl acetate. Also, enzyme assays showed that AreA activity is induced when ADP1 is grown on benzaldehyde or benzyl alcohol, and AreB and AreA activity are induced when ADP1 is grown on benzaldehyde. The coding regions of *areC* and *areB* are separated by 28 nucleotides, and the coding regions of *areB* and *areA* are separated by 84 nucleotides. The accumulated results suggest that the *areABC* genes are part of a single operon.

β-Galactosidase assays on ADPW61 (*areA*::*lacZ*-Km^r) and ADPW63 (*areC*::*lacZ*-Km^r) confirm that both *areC* and *areA* are induced by benzyl acetate, benzyl alcohol, or benzaldehyde. Moreover, the β-Galactosidase assays consistently show that *areC* is induced to a much higher level than *areA* by the same substrate. β-Galactosidase assay values were approximately ten times higher in ADPW63 than in ADPW61. This suggests that there might be differential expression of the operon where *areCB* are transcribed to at higher level than *areA*. The small inverted repeat sequence between *areB* and *areA* might be acting as a partial transcriptional terminator resulting in the general effect of differential transcription of the *areABC* genes. The evidence for this is based upon β-Galactosidase assays and sequence browsing. Further evidence must be gathered to substantiate the hypothesis.

CHAPTER 6

Functional Analysis of AreR, a σ^{54}-type Bacterial Transcription Enhancer

6.1. Introduction.

In Chapter 5, it was shown that *areCBA* is expressed as an operon, and is inducible by a range of benzyl compounds. This indicates the existence of a regulatory system controlling the expression of the operon. It was also shown that upstream of the transcriptional start of *areC* is a region with high homology to the –12,–24 type promoter sequence that is recognised by the σ^{54} sigma factor. In addition, an inverted repeat sequence characteristic of a bacterial enhancer sequence is located 76 bp upstream of the –12,–24 region. Analysis of other systems of this type showed that genes are expressed when a bacterial enhancer-binding protein binds co-operatively to the inverted repeats. This protein then contacts the σ^{54}–holoenzyme through DNA looping and catalyses the isomerisation of the closed complex between the σ^{54}–holoenzyme and the promoter to an open complex, in a process that requires ATP hydrolysis by the regulator protein.

Several types of enhancer-binding proteins (also called 'activators of σ^{54} – holoenzyme' or 'NtrC family of transcriptional regulators') have been identified in bacterial systems. The proteins have a characteristic modular structure containing three functional domains that fold and function independently of one another. In most cases, the amino-terminal domain is responsible for signal reception, the central domain functions in ATP hydrolysis and activation, and the carboxy-terminal domain is involved in DNA binding (Shingler, 1996). The gene adjacent to *areC*, designated *areR*, shows high sequence similarity to this family of transcriptional regulators. In this chapter, experiments are described to determine whether AreR controls the expression of *areCBA*.

6.2. Analysis of the *areR* Gene Sequence.

The AreR amino acid sequence was analysed by searches of the GenBank database using BLASTP (Altschul *et al.* 1990). The analysis revealed that the closest relatives to AreR are members of the bacterial enhancer binding protein family. The AreR amino acid sequence was then aligned with the two most homologous proteins, which were AcoR from *Clostridium magnum* (Kruger *et al.* 1994) and AcoR from *P.putida* (GenBank accession CAA72019), as well as the extensively studied DmpR gene sequence which regulates the phenol catabolic genes from a *Pseudomonas* sp (Shingler *et al.* 1992) using the ClustalX computer programme (Thompson *et al.* 1994) (Fig 6.2.1). The alignment identifies that the three domains characteristic of the bacterial enhancer binding protein family are present in AreR. The amino-terminal domain, or A-domain, is the least conserved, and the numbers of amino acids in the A-domains vary considerably. The AreR and the two AcoR proteins, from *P.putida and* from *C. magnum* A-domains have 270, 280, and 290 amino acid residues respectively, whereas the DmpR A-domain has 211 amino acids. The percentage similarities of the A-domains of AreR to those of *P.putida* AcoR, of *C. magnum* AcoR and of the DmpR A-domains are 37, 34 and 23% respectively. There is high homology between residues 159 and 174 of AreR where the sequence PVFNGQGKILGALDIT is conserved in AreR and both AcoR proteins. This region of AreR aligns with the region which includes residue 135 of DmpR which has been implicated as having a role in interactions between the A- and C- domains (Ng *et al.* 1996).

The central or C-domain is well conserved in the four aligned sequences. The lengths of the C-domains of each gene is very similar: AreR (233 amino acids), AcoR from *P.putida* (236), AcoR from *C. magnum* (251), and DmpR (238). The similarities of AreR C-domain to the *P.putida* AcoR, the *C. magnum* AcoR and DmpR are 61, 57 and 55% respectively. The C-domain can be subdivided into seven well-conserved sub-regions which are designated C1-C7 (Morret & Segovia, 1993). Each of the seven regions of the C-domain can be identified in AreR (Fig. 6.2.1.). The C-1 region of AreR contains the glycine-rich phosphate-binding loop common in ATP- and GTP- binding motifs [LIVMFY](3)-x-G-[DEQ]-[STE]-G-[STAV]-G-K-x(2)-[LIVMFY] (Bairoch *et al.* 1997). Subregion C-4 contains a motif that resembles the Walker-type b sequence found

Fig. 6.2.1. The alignment of AreR with other proteins from the NtrC family of transcriptional regulators. Alignment of the predicted amino acid sequences of AreR with the 2 most similar members of the NtrC family: AcoR from *Pseudomonas putida* (p.put) (Acc.CAA72019), and AcoR from *Clostridium magnum* (c.mag) (Acc.I40789). The alignment also includes the extensively studied protein DmpR from *Pseudomonas* sp. (Acc. CAA48174). The alignment was done using the ClustalX (PAM350 matrix) (Thompson *et al.* 1994) computer program. Residues that are conserved in all four sequences are highlighted in black, residues that are conserved in three of the four sequences are highlighted in dark shade grey, and residues that are conserved in two of the four sequences are highlighted in light grey. The ruler shown above the sequences corresponds to the amino acid sequence of AcoR (p.put). Gaps introduced to maximise the alignment are dashed. The regions corresponding to the different domains are indicated, and their names are displayed below the sequences. The borders of the A, C, and D domains are set according to those proposed for the DmpR protein (Shingler, 1996). The sections C1 to C7, shown below the sequences, correspond to the conserved regions within the C domain (Morett & Segovia, 1993).

Fig.6.2.1.

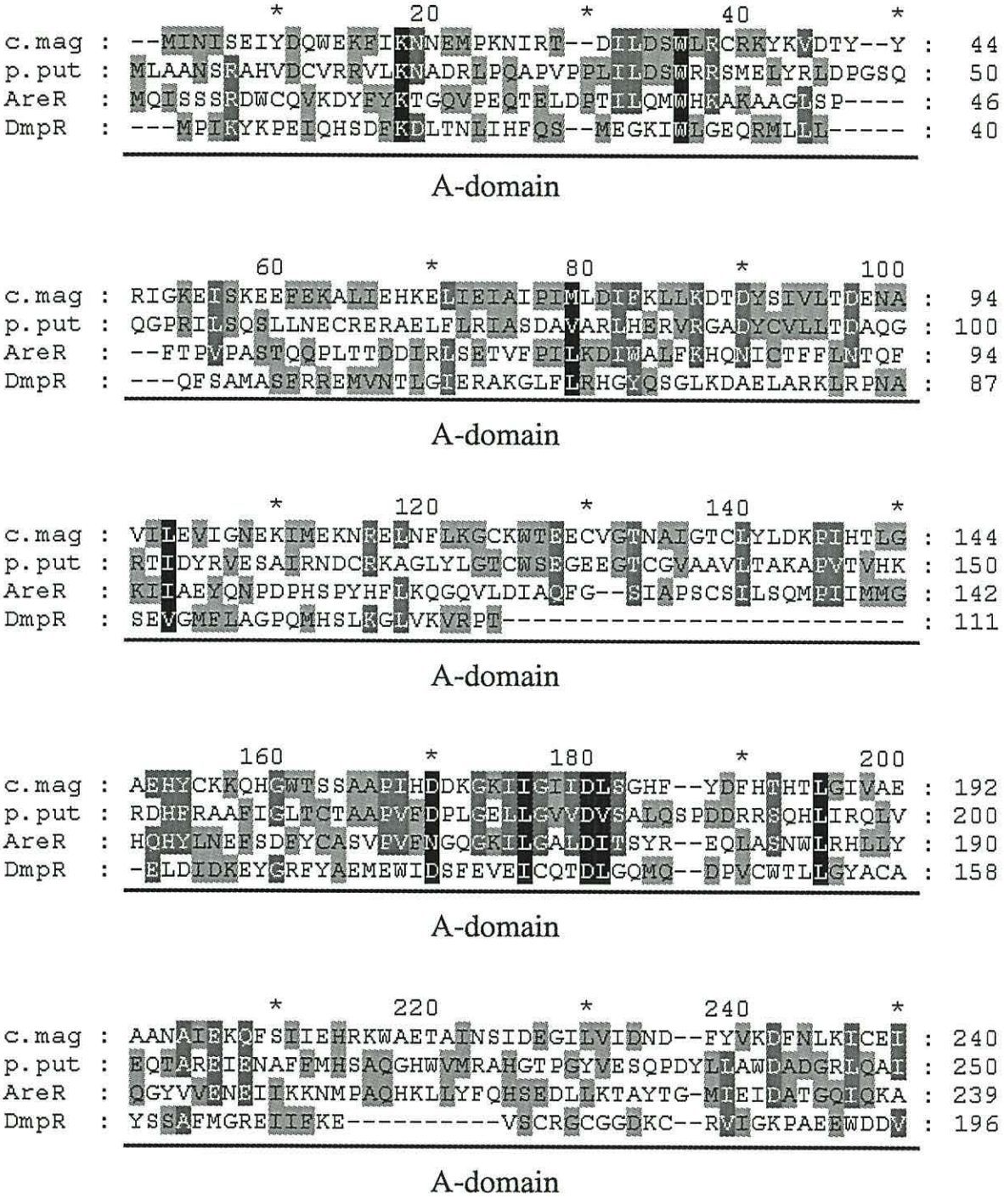

c.mag : --MINISEIYDQWEKEIKNNEMPKNIRT--DILDSWLRCRKYKVDTY--Y : 44
p.put : MLAANSRAHVDCVRRVLKNADRLPQAPVPPLILDSWRRSMELYRLDPGSQ : 50
AreR : MQISSSRDWCQVKDYFYKTGQVPEQTELDPTILQMWHKAKAAGLSP---- : 46
DmpR : ---MPIKYKPEIQHSDFKDLTNLIHFQS--MEGKIWLGEQRMLLL----- : 40
A-domain
c.mag : RIGKEISKEEFEKALIEHKELIEIAIPIMLDIFKLLKDTDYSIVLTDENA : 94
p.put : QGPRILSQSLLNECRERAELFLRIASDAVARLHERVRGADYCVLLTDAQG : 100
AreR : --FTPVPASTQQPLTTDDIRLSETVFPILKDIWALFKHQNICTFFLNTQF : 94
DmpR : ---QFSAMASFRREMVNTLGIERAKGLFLRHGYQSGLKDAELARKLRPNA : 87
A-domain
c.mag : VILEVIGNEKIMEKNRELNFLKGCKWTEECVGTNAIGTCLYLDKPIHTLG : 144
p.put : RTIDYRVESAIRNDCRKAGLYLGTCWSEGEEGTCGVAAVLTAKAPVTVHK : 150
AreR : KIIAEYQNPDPHSPYHFLKQGQVLDIAQFG--SIAPSCSILSQMPIIMMG : 142
DmpR : SEVGMELAGPQMHSLKGLVKVRPT------------------------- : 111
A-domain
c.mag : AEHYCKKQHGWTSSAAPIHDDKGKIIGIIDLSGHF--YDFHTHTLGIVAE : 192
p.put : RDHFRAAFIGLTCTAAPVFDPLGELLGVVDVSALQSPDDRRSQHLIRQLV : 200
AreR : HQHYLNEFSDEYCASVPVFNGQGKILGALDITSYR--EQLASNWLRHLLY : 190
DmpR : -ELDIDKEYGRFYAEMEWIDSFEVEICQTDLGQMQ--DPVCWTLLGYACA : 158
A-domain
c.mag : AANAIEKQFSIIEHRKWAETAINSIDEGILVIDND--FYVKDFNLKICEI : 240
p.put : EQTAREIENAFFMHSAQGHWVMRAHGTPGYVESQPDYLLAWDADGRLQAI : 250
AreR : QGYVVENEIIKKNMPAQHKLLYFQHSEDLLKTAYTG-MIEIDATGQIQKA : 239
DmpR : YSSAFMGREIIEKE----------VSCRGCGGDKC--RWIGKPAEEWDDV : 196
A-domain

Fig.6.2.1. Cont.

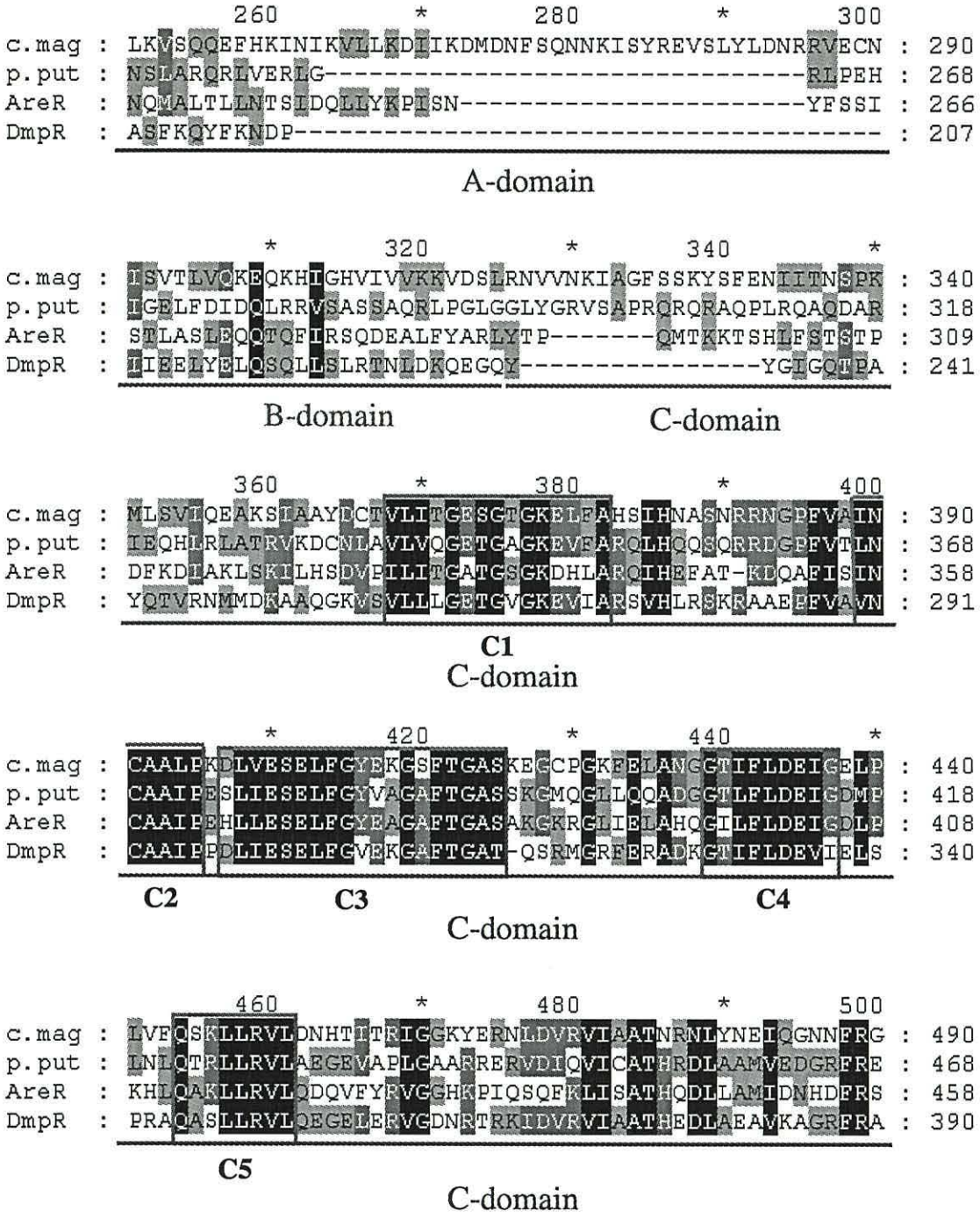
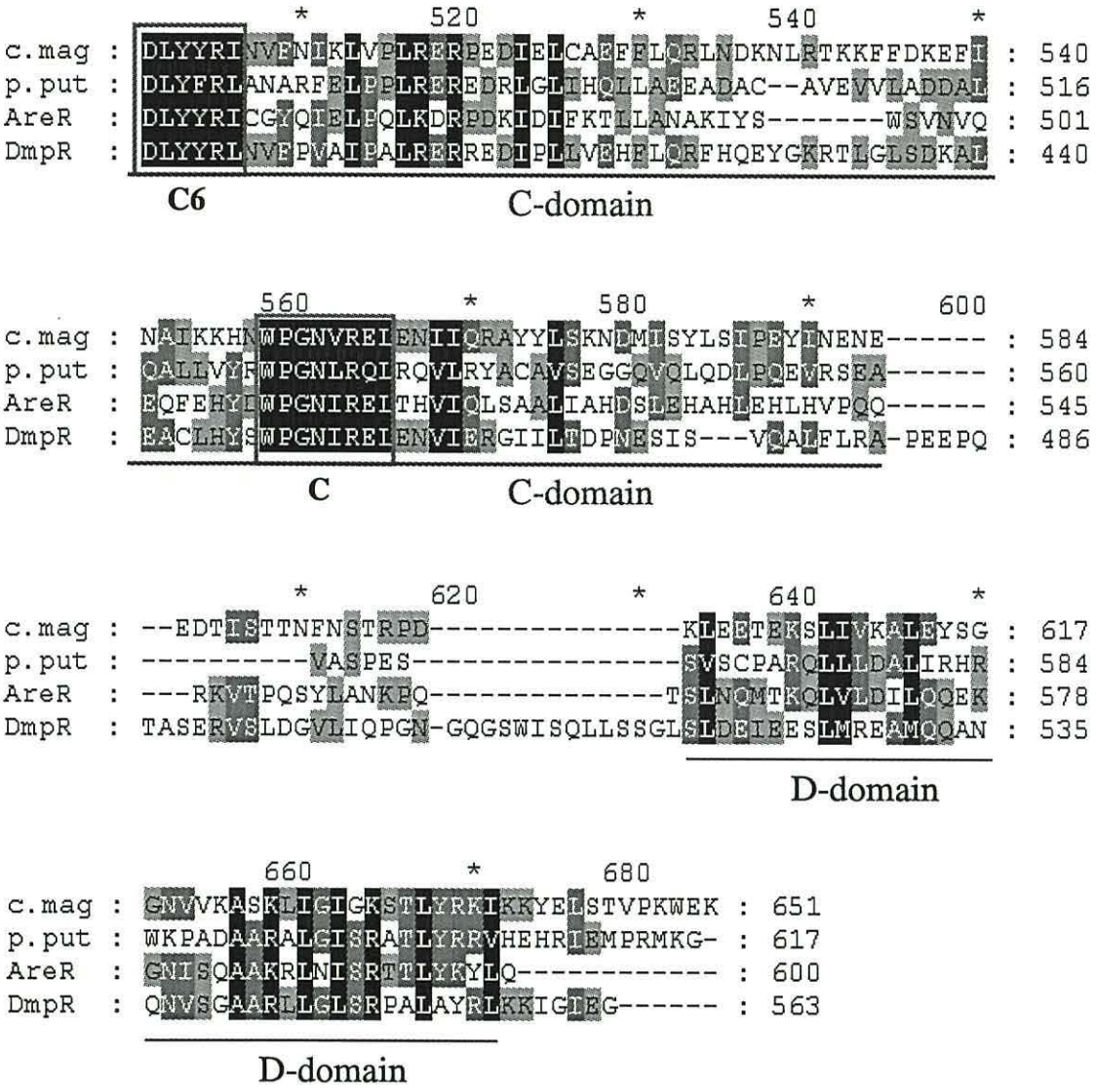

260 * 280 * 300
c.mag : LKVSQQEFHKINIKVLLKDIIKDMDNFSQNNKISYREVSLYLDNRRVECN : 290
p.put : NSLARQRLVERLG---------- RLPEH : 268
AreR : NQMALTLLNTSLDQLLYKPISN---------- YFSSI : 266
DmpR : ASFKQYFKNDP---------- : 207
A-domain
* 320 * 340 *
c.mag : ISVTLVQKEQKHIGHVIVVKKVDSLRNVVNKIAGFSSKYSFENIITNSPK : 340
p.put : IGELFDIDQLRRVSASSAQRLPGLGGLYGRVSAPRQRQRAQPLRQAQDAR : 318
AreR : STLASLEQQTQFIRSQDEALFYARLYTP------- QMTKKTSHLFSTSTP : 309
DmpR : IIEELYELQSQLLSLRTNLDKQEGQY---------- YGIGQTPA : 241
B-domain
C-domain
360 * 380 * 400
c.mag : MLSVIQEAKSIAAYDCTVLITGESGTGKELFAHSIHNASNRRNGPFVAIN : 390
p.put : IEQHLRLATRVKDCNLAVLVQGETGAGKEVFARQLHQQSQRRDGPFVTLN : 368
AreR : DFKDIAKLSKILHSDVPILITGATGSGKDHLARQIHEFAT-KDQAFISIN : 358
DmpR : YQTVRNMMDKAAQGKVSVLLLGETGVGKEVIARSVHLRSKRAAEPFVAVN : 291
C1
C-domain
* 420 * 440 *
c.mag : CAALPKDLVESELFGYEKGSFTGASKEGCPGKFELANGGTIFLDEIGELP : 440
p.put : CAAIPESLIESELFGYVAGAFTGASSKGMQGLLQQADGGTLFLDEIGDMP : 418
AreR : CAAIPEHLLESELFGYEAGAFTGASAKGKRGLIELAHQGILFLDEIGDLP : 408
DmpR : CAAIPPDLIESELFGVEKGAFTGAT-QSRMGRFEQADKGTIFLDEVIELS : 340
C2
C3
C4
C-domain
460 * 480 * 500
c.mag : LVFQSKLLRVLQNHTITRIGGKYERNLDVRVIAATNRNLYNEIQGNNFRG : 490
p.put : LNLQTRLLRVLAEGEVAPLGAARREPVDIQVICATHRDLAAMVEDGRFRE : 468
AreR : KHLQAKLLRVLQDQVFYRVGGHKPIQSQFKLISATHQDLLAMIDNHDFRS : 458
DmpR : PRAQASLLRVLQEGELERVGDNRTRKIDVRVIAATHEDLAEAVKAGRFRA : 390
C5
C-domain

Fig.6.2.1. Cont.


c.mag : DLYYRINVENIKLVPLRERPEDIELCAEFFLQRLNDKNLRTKKFFDKEFI : 540
p.put : DLYFRLANARFELPPLRERERDRLGLIHQLLAEEADAC--AVEVVLADDAL : 516
AreR : DLYYRICGYQIELPQLKDRPDKIDIFKTLLANAKIYS-------WSVNVQ : 501
DmpR : DLYYRLNVEPVAIPALRERREDIPLLVEHFLQRFHQEYGKRTLGLSDKAL : 440
C6
C-domain
c.mag : NAIKKHNWPGNVREIENIIQRAYYLSKNLMISYLSIPEYINENE------ : 584
p.put : QALLVYRWPGNLRQIRQVLRYACAVSEGGQVQLQDLPQEVRSEA------ : 560
AreR : EQFEHYIWPGNIREITHVIQLSAALIAHDSLEHAHLEHLHVPQQ------ : 545
DmpR : EACLHYSWPGNIREIENVIERGIILTDPNESIS---VQALFLRA-PEEPQ : 486
C
C-domain
c.mag : --EDTISTTNFNSTRPD----------KLEETEKSLIVKALEYSG : 617
p.put : ----------VASPES----------SVSCPARQLLLDALIRHR : 584
AreR : ---RKVTPQSYLANKPQ----------TSLNQMTKQLVLDILQQEK : 578
DmpR : TASERVSLDGVLIQPGN-GQGSWISQLLSSGLSLDEIEESLMREAMQQAN : 535
D-domain
c.mag : GNVVKASKLIGIGKSTLYRKIKKYELSTVPKWEK : 651
p.put : WKPADAARALGISRATLYRRVHEHRIEMPRMKG- : 617
AreR : GNISQAAKRLNISRTTLYKYLQ---------- : 600
DmpR : QNVSGAARLLGLSRPALAYRLKKIGIEG------ : 563
D-domain

in ATP- and GTP- binding proteins (Kelly & Hoover, 1999). The C-5 domain has the conserved QXLLRVL, which, on the basis of sequence comparisons with other ATP-binding proteins, has also been shown to be involved in ATP-binding and hydrolysis (Shingler, 1996). The C-3 region has the consensus amino acid sequence ESELFGXXXGAFTGA and is diagnostic for activators of σ^{54} –holoenzyme. The C-3 region is predicted to consist of two alpha helices separated by a turn. The C-3 region is thought to adjoin the σ^{54} –holoenzyme. AreR also contains the other characteristic motifs C-6 and C-7, but no functions have yet been assigned to these subregions.

The carboxy-terminal domain (the D-domain), of AreR also aligns well with the D-domains of the other three proteins (Fig. 6.2.1.). The D-domains of the proteins of this family are thought to constitute a helix-turn-helix motif involved in the binding of the protein to the DNA (Shingler, 1996).

6.3. Mutational analysis of *areR* function.

6.3.1. Insertional inactivation of *areR* in *Acinetobacter* sp. ADP1.

AreR has all of the characteristics of σ^{54}-type enhancer-binding proteins, but is *areR* the gene that encodes the activator protein that binds to the enhancer sequence upstream of *areC*? To prove this hypothesis, insertional inactivation of *areR* was performed, and the ability of the resulting mutant to activate the transcription of *areCBA* analysed. Firstly, pADPW45 was cloned from ADPW56 (*areC*::*Km*^{r}) (Table 6.3.1.) on a *Sac*I/*Xba*I fragment as described in Chapter 5.6.1 (Fig 6.3.1.). A *Hin*dIII subclone of pADPW45 was created and designated pADPW73 (Fig 6.3.1.). Then a Km^r cassette from pUC4K was digested with *Bam*HI and inserted into the sole *Bgl*II restriction site of pADPW73 that is located 1339 bases or 446 amino acids into the *areR* open reading frame, and lies in the C-domain of the protein, to create pADPW90 (Fig. 6.3.1.). The pADPW90 was linearised and transformed into ADP1 to create ADPW79 (*areR*::Km^r). The ADPW79 genomic DNA was analysed by Southern blotting, and the genomic insertion confirmed (Fig. 6.3.2).

Table 6.3.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter		
ADP1 (BD413)	Wild type	Juni, 1969
ADPW56	*areC*::Km^r; transformation of ADP1 with pADPW22	This study
ADPW79	*areR*::Km^r; transformation of ADP1 with pADPW90	This study
E. coli		
DH5α	F$^-$, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	Gibco BRL
pUC18	Ap^r, cloning vector	Yanisch-Perron, 1982
pADPW22	pADPW21 with Km^r cassette from pUC4K cloned into *Bgl*II site in *areC*	This study
pADPW45	6.2-kbp *HinId*III fragment cloned from ADPW56 containing *areR* and *areC*::Km^r in pUC18	This study
pADPW71	3.5-kbp *Hin*dIII fragment from pADPW45 in pUC18	This study
pADPW90	pADPW71 with Km^r cassette from pUC4K cloned into the *Bgl*II site in *areR*	This study
	* denotes restriction site added by PCR	

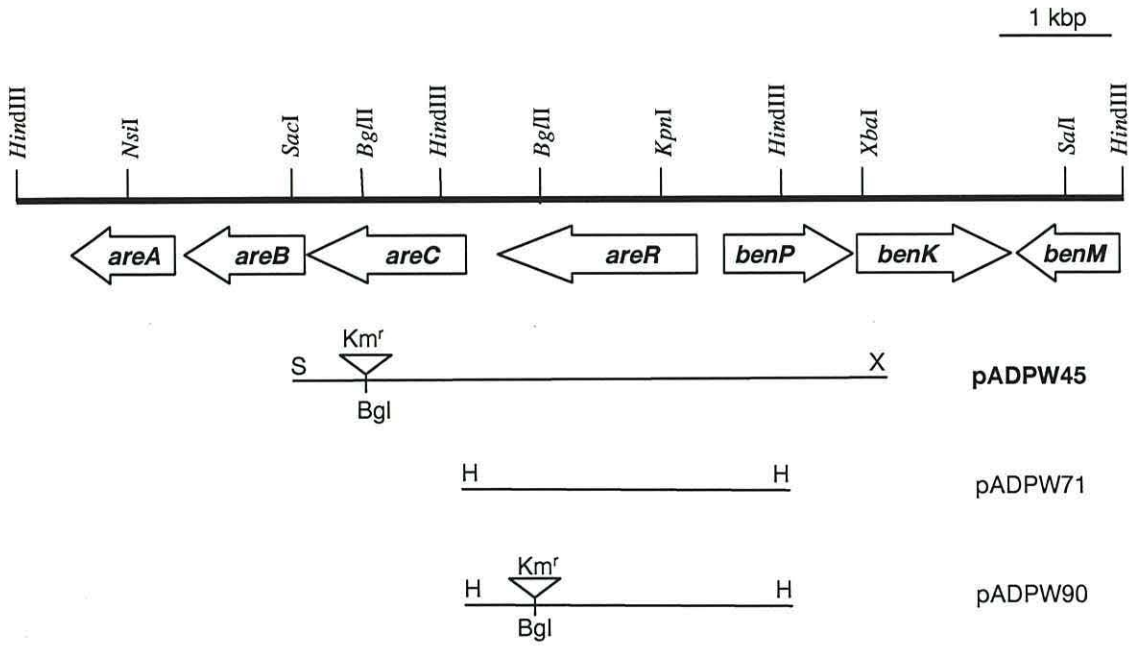


Fig. 6.3.1. Physical map of the *are* genes and their location relative to the *ben-cat* cluster. The inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 6.3.1. pADPW45 named in boldface type was cloned directly from genomic DNA. Other plasmids were produced by subcloning from plasmids containing genomic DNA. The Km^r cassette insertion is not to scale. The abbreviations for the restriction sites are: H, *Hin*dIII; S, *Sac*I; X, *Xba*I.

Fig. 6.3.2. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *areR* to create ADPW79 (*areR*::Km^r).

Fig. 6.3.2.a. Map of the DNA sequence upstream and downstream of *areR*, pADPW71 (3.5kbp), and pADPW90 (*areR*::Km^r) (4.6kbp). ADPW79 was created by transformation of pADPW90 into ADP1. ADPW79 has a chromosomal rearrangement where the 1.2kbp Km^r cassette is integrated into the genome at the *Bgl*II restriction site that is within *areR*.

Fig. 6.3.2.b. The agarose gel blot was hybridised with an *areR* probe purified by band extraction from *Hin*dIII digested pADPW71 (Fig.6.3.1.). Lane S contain 8*Hin*dIII size marker (Promega) and the fragment sizes are stated in kbp. The sizes of the bands produced following hybridisation to *areR* are shown (kbp):

Lane1: pADPW71 (*Hin*dIII) 3.5.
Lane2: pADPW90 (*Hin*dIII) 2.8, 1.9.
Lane3: pADPW71 (*Cla*I) 3.3, 2.4.
Lane4: pADPW90 (*Cla*I) 3.3, 2.8.

Lane5: ADP1 (*Hin*dIII) 3.5.
Lane6: ADPW79 (*Hin*dIII) 2.4, 1.8.

Fig. 6.3.2.a.

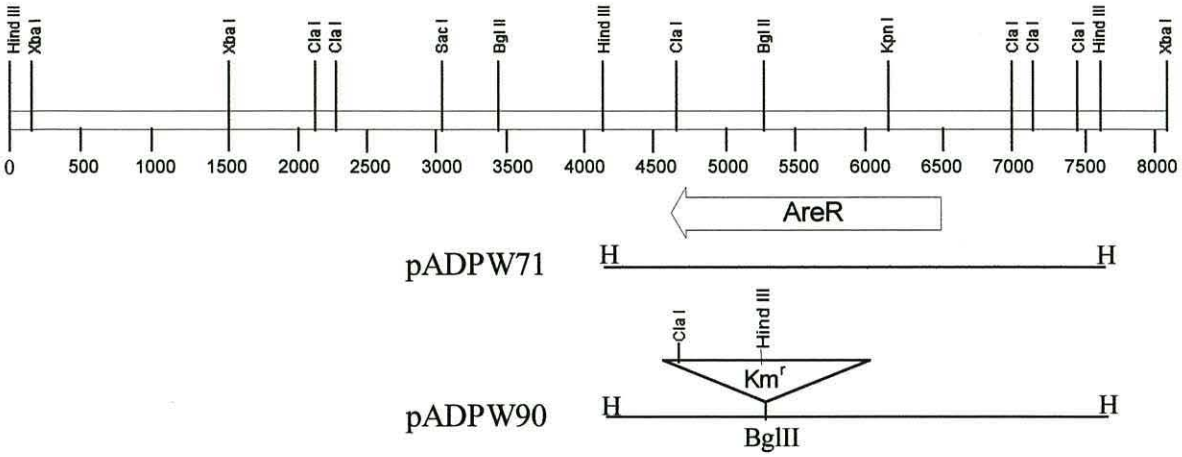
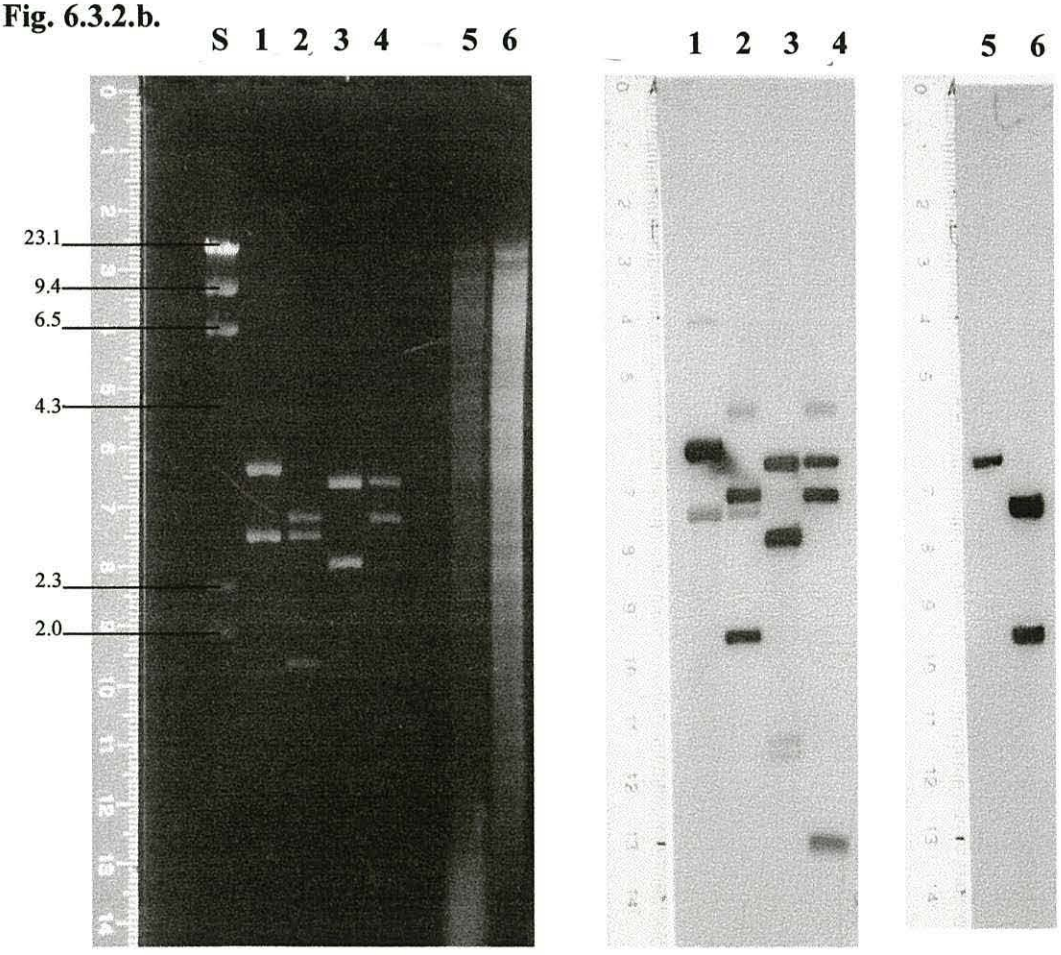


Fig. 6.3.2.b.



6.3.2. Analysis of ADPW79 (*areR*::*Km*^{r}) growth phenotype.

ADPW79 was patched on minimal media plates containing either benzyl acetate, benzyl alcohol, benzaldehyde, or succinate. Whereas wild type ADP1 grows on all four substrates overnight, growth of ADPW79 on benzyl acetate, benzyl alcohol, or benzaldehyde is considered challenged in that it takes two to three days for ADPW79 to grow to the same colony size. Thus, as with the *areCBA* gene disruptions, insertional inactivation of the gene does not result in the complete loss of ability of the strain to grow on benzyl acetate, benzyl alcohol or benzaldehyde. As with the *areCBA* gene disruptions, competition experiments were carried out in which ADPW79 competed against ADP1 for growth on substrates (Chapter 3.4.1). The competition experiment was done three times, and in each experiment, the ADPW79 could not compete with ADP1 when both were competing for growth on succinate as well as on benzyl acetate, benzyl alcohol and benzaldehyde. The wild type ADP1 had out-competed ADPW79 even when grown on succinate, and thus it must be concluded that the insertion in the regulatory gene *areR* caused a general weakness in the mutant.

6.3.3. Biochemical assays of ADPW79 against *are* operon substrates.

Cell free extracts of ADP1 cells grown on minimal media containing succinate, benzyl alcohol, or benzyl alcohol + succinate, and of ADPW79 grown on minimal media containing benzyl alcohol + succinate were prepared. The cultures were harvested at 4°C and washed with 0.1M phosphate buffer pH7.4 before being resuspended in 10ml of the same buffer. Cell free extracts were prepared as described in Chapter 2.9. Two independent preparations of the cell free extracts were assayed in triplicate at two enzyme concentrations, against benzyl acetate, benzyl alcohol, and benzaldehyde. The specific activity was determined for each assay (Table 6.3.2).

No increased AreC, AreB nor AreA activity could be measured in ADPW79 cell free extracts grown on succinate plus benzyl alcohol. AreA, AreB, and AreC activity was only detected in crude cell free extracts from ADP1 cells grown on benzyl alcohol, or benzyl alcohol plus succinate indicating that succinate does not repress the induction of the *are* operon. No AreB or AreC activity could be measured in succinate-grown cells.

Table 6.3.2. Activities of ADP1 and ADPW79 against *are* substrates. Specific activities of AreA (Benzyl esterase), AreB (Benzyl alcohol dehydrogenase), and AreC (Benzaldehyde dehydrogenase) against benzyl acetate, benzyl alcohol, and benzaldehyde respectively in crude extracts of ADP1 or ADPW79 cells grown on succinate, benzyl alcohol, or succinate plus benzyl alcohol.

		Specific activity (U/mg of protein) against benzyl substrates		
Strain	Growth Substrate	Benzaldehyde dehydrogenase	Benzyl alcohol dehydrogeanase	Benzyl esterase
ADP1	Succinate	<0.04	<0.04	0.34
ADP1	Benzyl alcohol	0.37	1.12	4.48
ADP1	Succinate + Benzyl alcohol	0.35	0.91	6.36
ADPW79	Succinate + Benzyl alcohol	<0.04	<0.04	0.14

[a]All reaction rates were based on the average of three measurements , none of which varied by >5%.

The AreA specific activity measured in succinate-grown cells was over 500 times lower than when grown on benzyl alcohol. Thus, insertional inactivation of the regulator gene *areR* results in the loss of ADP1's ability to induce AreC, AreB, or AreA activity, when grown on benzyl alcohol, which normally induces the expression of the genes.

6.4.Discussion.

6.4.1. Experimental evidence for the function of AreR.

In this chapter the AreR protein from *Acinetobacter* sp.ADP1 has been characterised. The nucleotide sequence of the *areR* gene was determined, encoding a predicted protein mass of 67kDa. The function of AreR as the regulator protein responsible for the control of the expression of AreABC was inferred. This was established from studies where AreR function was disrupted by the insertion of a Km cassette, creating the *areR*$^-$ mutant ADPW79. Induction experiments with ADPW79 provided evidence that AreR is responsible for activated expression from the AreC promoter. Whereas benzyl alcohol, or benzyl alcohol plus succinate, induces AreA, B, or C activity in ADP1, no AreA, B, or C activity could be detected in ADPW79 when grown on either medium.

6.4.2. Sequence analysis evidence for the function of AreR.

AreR has significant amino acid sequence homology to members of the NtrC family of bacterial transcriptional activators. The highest homology to other members of the NtrC family was, as expected, in the central C-domain. Homologies were also found in the A-domain of two proteins referred to as AcoR from *P.putida* and from *C. magnum.* There is little homology throughout the A-domains of these genes. However, neither AcoR proteins has been extensively characterised. Nevertheless, the location of the *acoR* gene from *C. magnum* adjacent to genes encoding the enzymes for acetoin (3-hydroxy-2-butanone) dehydrogenase in that bacterium suggests a function of 'aco operon expression regulatory protein' (Kruger *et al.* 1994). A further similarity between the *areR* and *acoR* is that both genes are transcribed in the same direction as the gene that they putatively regulate, with AreR regulating expression of *areC* (benzaldehyde dehydrogenase), and

AcoR regulating the expression of *acoA* (acetoin dehydrogenase). This is unusual since the vast majority of NtrC type transcriptional regulators are transcribed divergently to the gene that they regulate (Shingler, 1996).

There is little evidence for the mechanism by which AreR is activated by the effector. Little is known of the way AcoR if activated and it is difficult to draw conclusions by sequence alignment alone. Nevertheless, the A domain of AreR has no significant homology to the A domains of response regulators within the NtrC family and there does not appear to be a conserved Asp residue corresponding to the Asp54 by which the membrane-bound autokinase, termed NtrB, controls the level of phosphorylation of its cognate activator, NtrC (Klose *et al.* 1993). This leaves the direct effector activation method of control, which is the way the XylR (Delgado & Ramos, 1994) and DmpR (O'Neill *et al.* 1998) subclass of NtrC proteins are activated. Both XylR and DmpR are activated by direct effector activation by specific ligands that are the distinct aromatic substrates of the catabolic pathways they control. There is extensive homology between the A-domains of DmpR and XylR. When the A-domain of AreR is compared to those of DmpR and XylR, there is limited homology between the residues. No conclusions can be drawn from the sequence evidence. Further studies might provide an insight and may assign a mechanism by which AreR is activated by an effector molecule.

CHAPTER 7

Cloning and characterising of the DNA adjacent to the *are* genes in *Acinetobacter* sp. ADP1

7.1. Introduction.

In this chapter, the DNA downstream of the *are* genes is cloned and characterised. This will extend the genetic region comprising the *cat*, *ben*, and *are* genes described in Chapter 3 and Chapter 4 of this thesis, and will show whether there are further genes coding for enzymes involved in aromatic catabolism.

7.2. Cloning and characterisation of the DNA adjacent to the *are* genes in *Acinetobacter* sp. ADP1.

7.2.1. Cloning of the ADP1 DNA downstream of *areA*.

Using the chromosomal antibiotic-resistance cassette in *areB* of strain ADPW57 (Table 7.2.1), DNA adjacent to *areABC* was cloned as described in Chapter 3.1.3. A plasmid library in pUC18 in *E. coli* was made from the chromosomal DNA of ADPW57, from which plasmid pADPW32 with an 8.0 kbp *Sac*I insert (Fig. 7.2.1) was selected for by Km^rAp^r resistance. pADPW34 was constructed as a *Sac*I to *Xba*I subclone of pADPW32 (Fig. 7.2.1) so as to have sequence overlap with the previously sequenced pADPW33 (Fig. 3.2.3) (Jones *et al.* 1999). pADPW34 was used as the DNA sequencing template. The sequencing was done to publication quality by MWG-Biotech Ltd (Ebersberg, Germany). Sequence alignment confirmed its overlap with pADPW33 (Appendix 1).

7.2.2. Analysis of nucleotide sequence for protein products from the DNA adjacent to *areCBA*.

Analysis of the nucleotide sequence downstream of *areA* revealed the presence of four open reading frames (Fig. 7.2.1). Immediately downstream of *areA*, but separated by 195 bp and the presence of a putative inverted repeat which could serve as a termination loop

TABLE 7.2.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter ADP1 (BD413)	Wild type	Juni, 1969
ADPW57	*areB*::Km^r; transformation of ADP1 with pADPW26	This study
ADPW67	*salA*::Km^r; transformation of ADP1 with pADPW44	This study
ADPW70	*salE*::Km^r; transformation of ADP1 with pADPW76	This study
ADPW72	*salR*::Km^r; transformation of ADP1 with pADPW79	This study
ADPW78	*salD*::Km^r; transformation of ADP1 with pADPW86	This study
E. coli DH5α	F$^-$, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ^-, *thi-1*, *gyrA96*, *relA1*	This study
XL1-Blue MRF'	Δ (*mcrA*)*183*, Δ(*mcrCB-hsdSMR-mrr*)*173*, *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, *lac* [F' *proAB lacI^{q}Z* ΔM15 Tn*10* (Tet^r)]	Gibco BRL
BL21(DE3)pLysS	F$^-$, *omp*T, *hsd*S_B(r_B^-, m_B^-), *dcm*, *gal*, (DE3), pLysS, Cm^r	Stratagene
pUC18	Ap^r, cloning vector	Yanisch-Peron, 1985
pUC4K	Ap^r, Km^r, source plasmid for Km^r cassette	Vieira, 1982
pUI1637	Ap^r, Km^r, source plasmid for Km^r cassette	Eraso, 1994
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and the whole of *areA* in pUC18	Jones, 1999
pADPW33	3.6-kbp *Hin*dIII subclone of pADPW32 in pUC18	Jones, 1999
pADPW34	4.1-kbp *SacI*/*Xba*I subclone of pADPW32 in pUC18	This study
pADPW41	1.6-kbp *SacI*/*Hin*dIII subclone of pADPW34 in pUC18	This study

pADPW44	pADPW41 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salA*	This study
pADPW49	1.0-kbp *Eco*RI* fragment containing *salE* in pUC18	This study
pADPW70	*Nde*I*/*Eco*RI* fragment containing *salE* in pET5a	This study
pADPW76	pADPW49 with Km^r cassette from pUI1637 cloned into *ClaI*I site in *salE*	This study
pADPW78	1.0-kbp *Eco*RI*/*Hin*dIII* fragment containing part of *salR* in pUC18	This study
pADPW79	pADPW78 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salR*	This study
pADPW82	1.2-kbp *Eco*RI* fragment containing part of *orf1* in pUC18	This study
pADPW86	pADPW82 with Km^r cassette from pUC4K cloned into *Nsi*I site in *orf1* * denotes restriction site added by PCR	This study

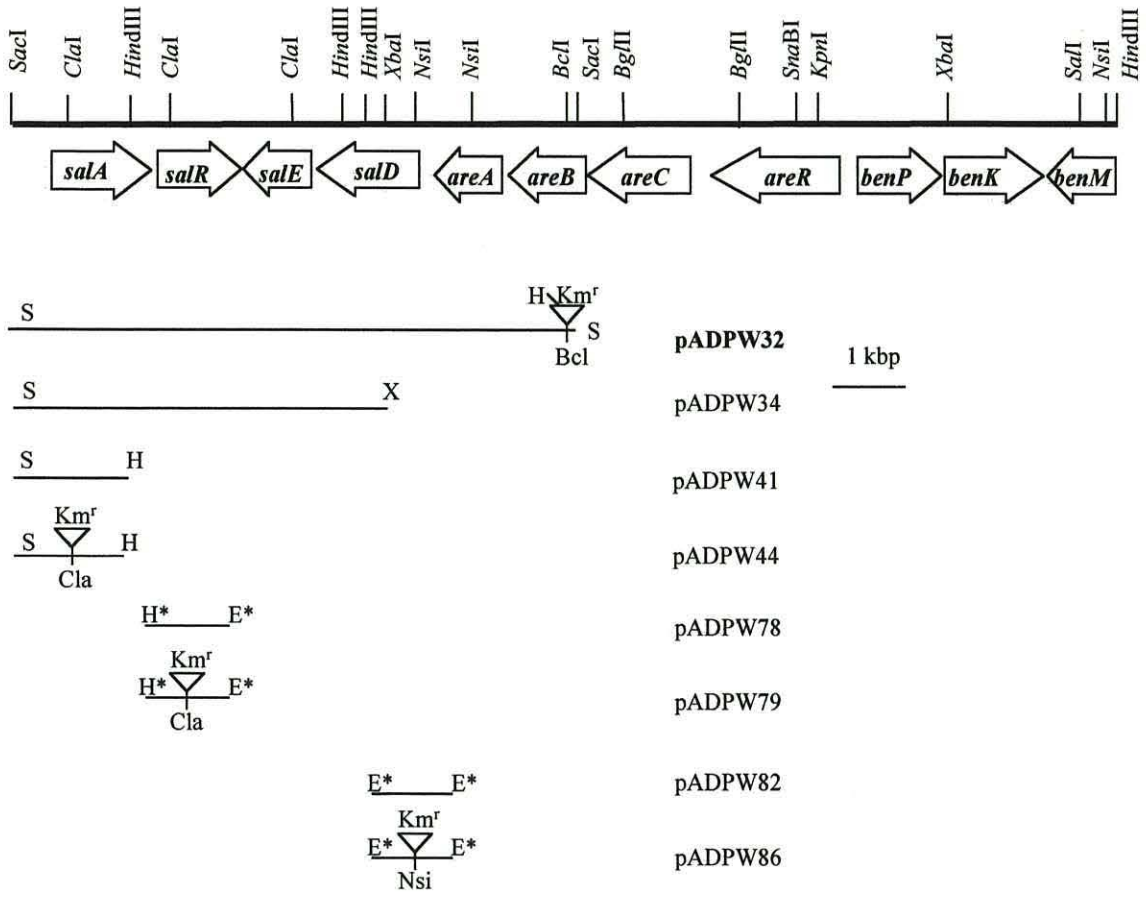


Fig. 7.2.1. Physical map of the *salA, salR, salE* and *salD* genes. The locations of the *salA, salR, salE* and *salD* genes relative to *areABC* at the left hand end (as drawn) of the supraoperonic *ben-cat* cluster is shown. The inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 7.2.1. pADPW32 named in boldface type was cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA (denoted by an asterisk) or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The Km^r cassette insertions are not to scale. The abbreviations for the restriction sites are: Bc, *Bcl*I; C, *ClaI;* E, *Eco*RI; H, *Hin*dIII; N, *Nsi*I; Nd, *Nde*I; S, *Sac*I; X, *XbaI.*

for *areA* expression (Chapter 5.2.1) is a gene denoted *salD*. SalD shows limited homology to putative proteins found in other operons of aromatic catabolism in different bacteria which include TodX (36% similarity) (Wang *et al.* 1995) and XylN (36% similarity) (Harayama S., Genbank accession no. D63341) both of which have been tentatively assigned functions as membrane proteins involved in the transport of hydrocarbons.

Downstream of *salD* is a gene we designated *salE* (Fig. 7.2.1), the product of which shows homology with serine esterases of the α/β-hydrolase family, but which has very limited similarity to the benzyl esterase (AreA) upstream of it. The characterisation of *salE* is the topic discussed in Chapter 8 of this thesis.

Transcribed convergently to *salE* is a gene designated *salR*, apparently encoding a regulatory protein of the LysR family. The closest matches to SalR are two NahR proteins involved in naphthalene catabolism, one from *Pseudomonas stutzeri* (Bosch *et al.* 1999) and one from the *P. putida* NAH7 plasmid (Schell & Sukordhaman, 1989) (Table 7.2.2). This latter protein has a dual role as positive regulator of the two functionally related operons, for the conversion of naphthalene to salicylate and for the further conversion of salicylate to central metabolites via catechol and the subsequent extradiol (*meta*) cleavage pathway (You *et al.* 1988; Schell & Poser, 1989; Schell & Sukordhaman, 1989).

The final gene in this cluster is *salA* (Fig. 7.2.1) which encodes a putative product, the closest relatives of which have both been named NahG (Table 7.2.2). These are salicylate hydroxylases, converting the salicylate produced from naphthalene to catechol and both are the first gene in the salicylate operons on the NAH7 plasmid (You *et al.* 1991) and in *P.stutzeri* (Bosch *et al.* 1999) respectively. However, ADP1 does not grow on napthalene as a sole carbon source. In ADP1 both *salA* and *salR* have a GTG start codon.

Table 7.2.2. *Acinetobacter* sp. ADP1 genes and gene products.

Gene designation	Putative function of gene product	Size of gene (bp)	%(A+T)	Size of gene product		Most similar gene products (% amino acid identity/similarity) (Acc. no.)
				Residues	kDa	
salD	putative membrane protein	1161	58	386	41.8	XylN (*P.putida* pWW0) (19/36%) (BAA09665) TodX (*P.putida*) (21/36%) (S53995)
salE	Esterase	720	54	239	27.0	Esterase (*M. pneumoniae*) (17/34%) (P75311)
salR	Regulatory protein	891	65	296	34.1	Regulatory Protein (*P. stutzeri*) (33/53%) (AAD02145) NahR (*P.putida* NAH7) (31/52%) (P10183)
salA	Salicylate Hydroxylase	1272	54	423	46.9	Salicylate hydroxylase (*P.putida*) (48/65%) (S51322) NahG (*P.putida* NAH7) (47/65%) (P23262)

7.3. Insertional inactivation of the *sal* genes.

7.3.1 Insertional inactivation of *salD*.

The chromosomal copy of *salD* was specifically disrupted by the insertion of a Km^r cassette into the gene. This work was undertaken by Weiske Pool under the guidance of Rheinallt M. Jones. The disruption of *salD* whilst on a plasmid was made after first creating pADPW82, which has a PCR-generated *Eco*RI insert in pUC18 (Fig 7.2.1). This 1.2 kbp fragment, with 0.6 kbp flanking either side of the *Nsi*I site in *salD*, was amplified from pADPW32 (Fig. 7.2.1). Primer sequences with the *Eco*RI sites underlined and the altered bases in bold, were as follows: salD (forward) 5'-AGGGGGAATTCTGGCAGCAATCACTG -3', and salD (reverse) 5'-GGGCTGGAATTCCCAAGTACTACCTAT -3'. The central *Nsi*I site in pADPW82 was used as the insertion site for a Km^r cassette from pUC4K (Vieira & Messing, 1982) to create pADPW86 (Fig 7.2.1.). This was done by excising the Km^r cassette from pUC4K with *Pst*I, which shares compatible overhanging ends with *Nsi*I. pADPW86 was linearised and the disrupted gene introduced into ADP1 by natural transformation to create ADPW78 (*salD*::Km^r) (Table 7.2.1.). ADPW78 (*salD*::Km^r) was a very unhealthy strain, grew slowly on minimal medium and lost viability when maintained on agar after 2 to 3 days. However within these limitations ADPW78 grew on salicylate. The unhealthy state of the mutant is probably related to the assigned function of SalD-related proteins as a membrane protein involved in the transport of hydrocarbons. Colonies appeared to be sticky as if they were lysed following 2 to 3 days on minimal media plus succinate or salicylate plates or on LB agar plates.

7.3.2. Insertional inactivation of *salR* and *salA*.

The chromosomal copies of *salA* and *salR* were specifically disrupted by the insertion of a Km^r cassette into each gene. As a first step, pUC18-derived plasmids were made that each carried part or whole of one of the genes disrupted by a Km^r cassette. For *salA*, a SacI/*Hin*dIII fragment of pADPW34 was cloned to create pADPW41 (Fig. 7.2.1.). The

Km^r cassette of pUI1637 (Eraso & Kaplan, 1994) was cloned into a unique *Cla*I site of pADPW41, creating pADPW44 (Fig. 7.2.1.).

Plasmid-borne *salR* disruption was achieved after first creating pADPW78 (Fig. 7.2.1), which has a PCR-generated *Eco*RI insert in pUC18. This 1.0 kbp fragment, internal to *salR* and having 0.5 kbp flanking both sides of the unique *Cla*I site, was amplified from pADPW34. Primer sequences, with the *Eco*RI sites underlined and the altered bases in bold, were as follows: (forward) 5'-TG**G**AATTCATGAACAGATCCGAAAAGAACG -3', and (reverse) 5'-CATGA**A**TTCCCTGAGTATGCCCGGTA -3'. The central *Cla*I site in the pADPW78 was used as the insertion site for the Km^r cassette from pUI1637 (Eraso & Kaplan, 1994) to create pADPW79 (Fig. 7.2.1).

The disrupted genes on pADPW44 (*salA*::Km^r) and pADPW79 (*salR*::Km^r) were linearised and transformed into ADP1, resulting in strains ADPW67 (*salA*::Km^r) and ADPW72 (*salR*::Km^r). The integration of the Km cassettes into the ADP1 genome was confirmed by Southern Blot analysis (Fig. 7.3.1) and (Fig. 7.3.2). Both ADPW67 (*salA*::Km^r) and ADPW72 (*salR*::Km^r) failed to grow on salicylate unlike ADP1 which grows vigorously on salicylate overnight.

7.3.3.Test for the formation of catechol by the catalytic activity of SalA.

Salicylate hydroxylase catalyses the conversion of salicylate to catechol. The formation of catechol produces a characteristic brown colour in the media surrounding the colony. *E.coli* DH5α containing plasmids pADPW34 or pADPW41 in (Fig. 7.2.1) were plated on LB ampicillin media containing 2.5mM salicylate and incubated at 37°C. Following two days incubation, a strong brown colour had developed in the media surrounding the strain containing pADPW34, but not in the media surrounding pADPW41. pADPW34 contains a 4.1 kbp *Sac*I/*Xba*I ADP1 DNA fragment containing all of the *salA* gene as well as the whole of the *salR* gene, whereas the strain containing pADPW41 contains a 1.8 kbp *Sac*I/*Hin*dIII ADP1 fragment containing an incomplete *salA* open reading frame. The truncated *salA* in pADPW41 was unable to convert salicylate to catechol.

Fig. 7.3.1. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *salA* to create ADPW67.

Fig. 7.3.1.a. Map of the DNA sequence upstream and downstream of *salA*, pADPW41 (1.6kbp), and pADPW44 (*salA*::Km^r 3.8kbp). ADPW67 was created by transformation of pADPW44 into ADP1. ADPW67 has a chromosomal rearrangement where the 2.2kbp Km^r cassette is integrated into the genome at the *Cla*I restriction site that is within *salA*. Restriction sites marked with an asterisk are on the plasmid multi cloning site. The *Hind*III site destroyed during cloning is marked [d].

Fig. 7.3.1.b. The agarose gel blot was hybridised with a *salA* probe purified by band extraction from *Sac*I /*Xba*I digested pADPW34 (Fig.7.2.1.). Lane S contain 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *salA* are shown (kbp):

Lane 1: pADPW41 (*Hin*dIII /*Eco*RI) 4.25. (*Hin*dIII in MCS destroyed during cloning)

Lane 2: pADPW44 (*Hin*dIII /*Xba*I) 4.3.

Lane 3: pADPW41 (*Cla*I) 4.25

Lane 4: pADPW44 (*Cla*I) 4.25

Lane 5: ADP1 (*Hin*dIII) 2.2, 2.0. (next *Hin*dIII site is 0.4kbp from the *Sac*I site)

Lane 6: ADPW67 (*Hin*dIII) 1.1, 0.8.

Lane 7: ADP1(*Xba*I/*Sac*I) 4.1.

Lane 8: ADPW67 (*XbaI*/*SacI*) 5.4, 0.8.

Fig. 7.3.1.a.

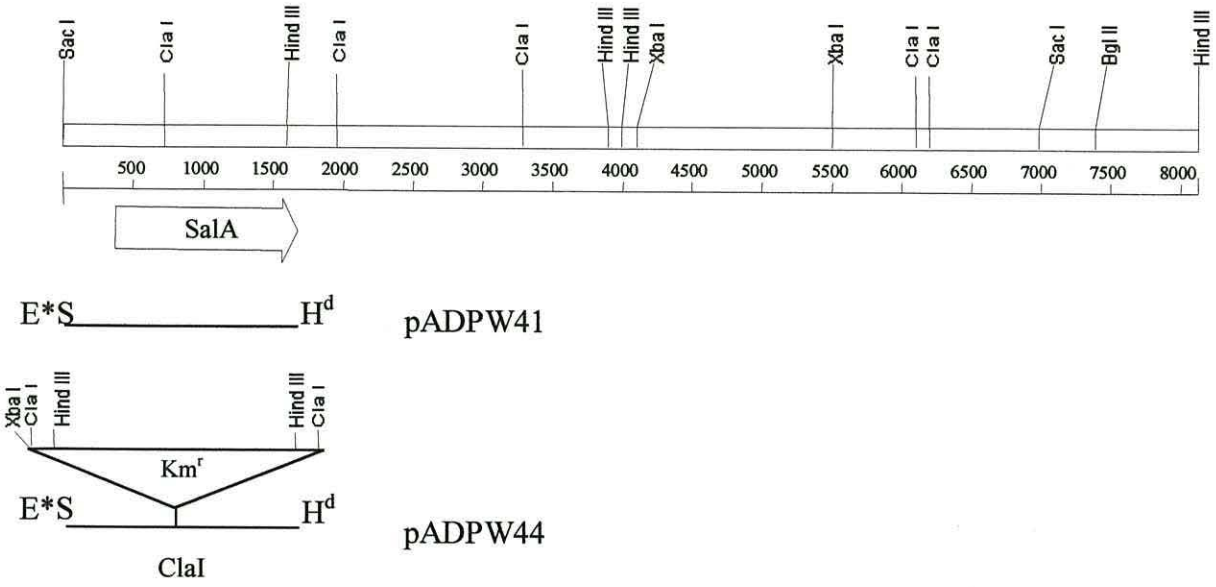


Fig. 7.3.1.b.

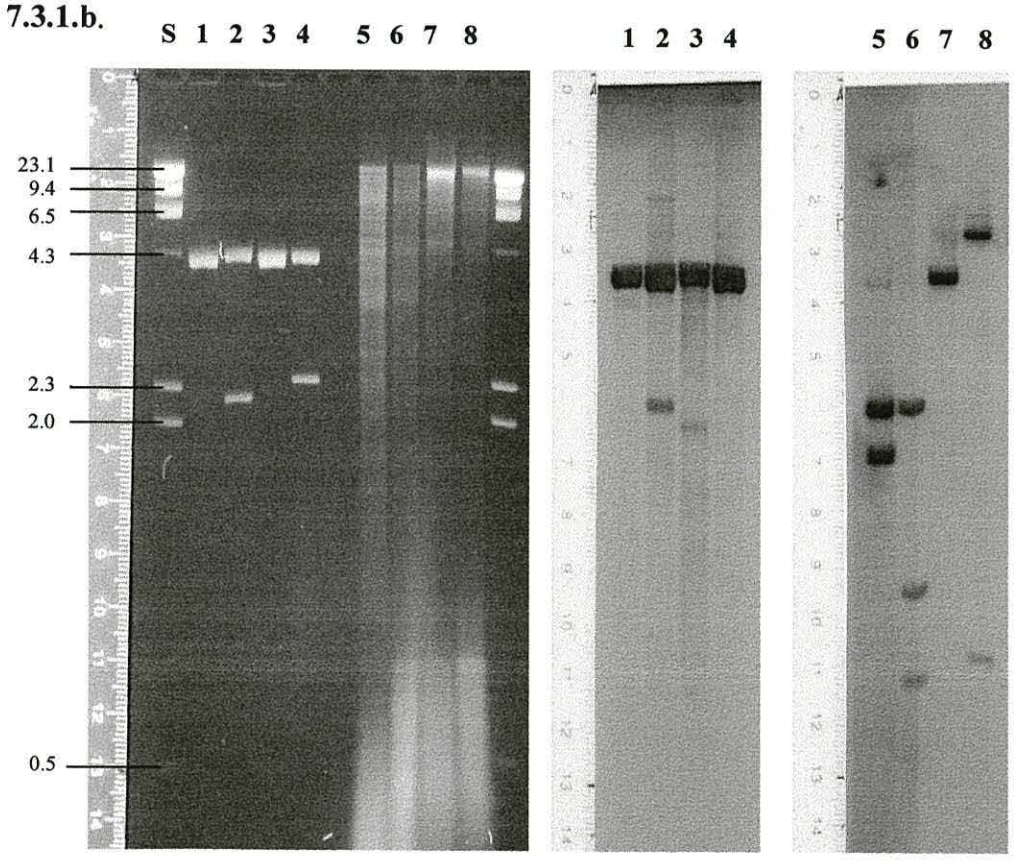


Fig. 7.3.2. Southern hybridisation analysis to demonstrate that the kanamycin cassette has recombined into *salR* to create ADPW72 (*salR*::Km^r).

Fig. 7.3.2.a. Map of the DNA sequence upstream and downstream of *salR*, pADPW78 (1.0kbp), and pADPW79 (*salR*::Km^r 3.2kbp). ADPW72 was created by transformation of pADPW79 into ADP1. ADPW72 has a chromosomal rearrangement where the 2.2kbp Km^r cassette is integrated into the genome at the *Cla*I restriction site that is within *salR*. Restriction sites marked with an asterisk are on the plasmid multi cloning site.

Fig. 7.3.2.b. The agarose gel blot was hybridised with a *salR* probe purified by band extraction from *Eco*RI/*Hin*dIII digested pADPW78. Lane S contains 8*Hin*dIII size marker (Promega) and the fragment sizes stated in kbp. The sizes of the bands produced following hybridisation to *salR* are shown (kbp):

Lane 1: pADPW78 (*Eco*RI/*Hin*dIII) 0.9.
Lane 2: pADPW79 (*Eco*RI/*Cla*I) 0.55, 0.45.
Lane 3: pADPW78 (*Eco*RI) 1.0.
Lane 4: pADPW79 (*Eco*RI) 2.8, 0.5.

Lane 5: ADP1 (*Hin*dIII) 2.2.
Lane 6: ADPW72 (*Hin*dIII) 1.9, 0.4 (faint).
Lane 7: ADP1 (*Xba*I/*Sac*I) 4.1.
Lane 8: ADPW72 (*XbaI*/*SacI*) 4.1, 2.2.

Fig. 7.3.2.a.

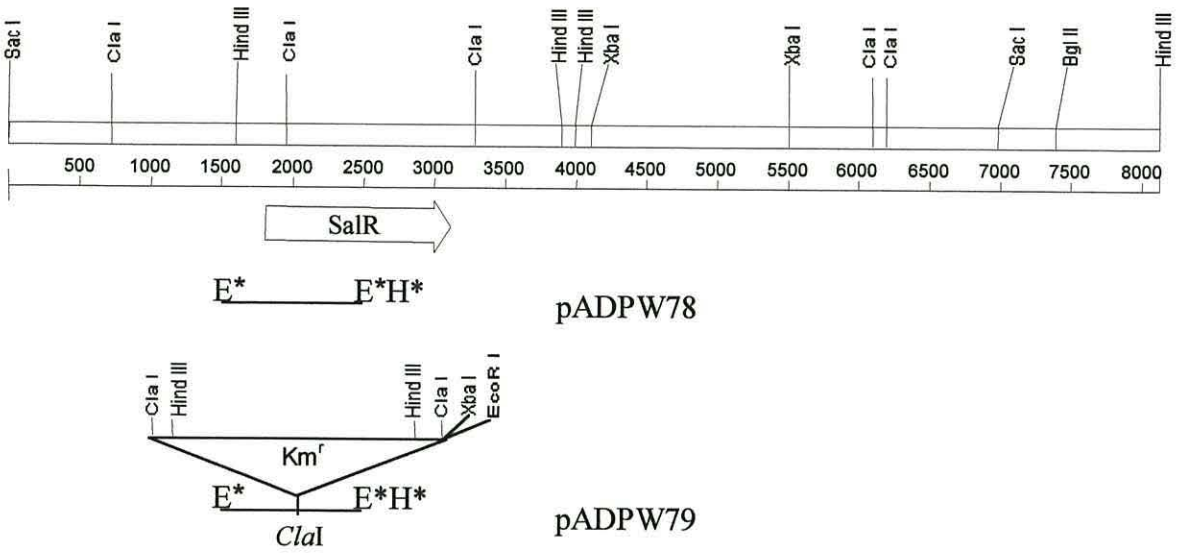
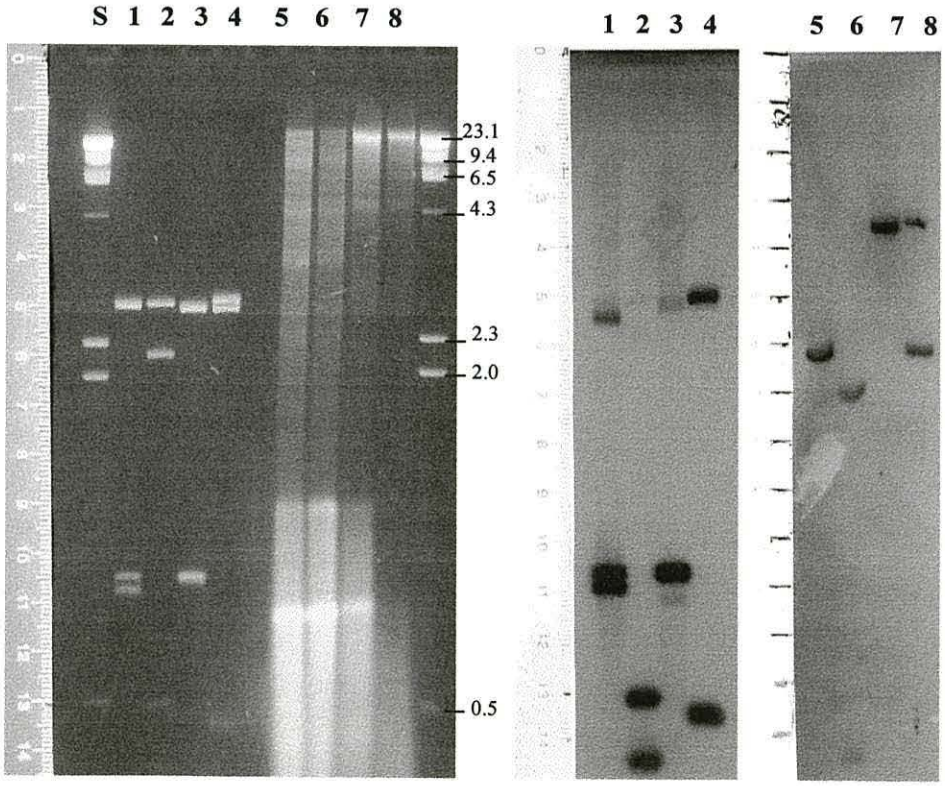


Fig. 7.3.2.b.



7.4. Biochemical Assays determining Salicylate hydroxylase activity.

7.4.1. Verification of the quality of the SalA biochemical assays.

Biochemical assays to determine salicylate hydroxylase activity were done as described in Chapter 2.10.2. Each rate determination was determined in triplicate at two enzyme concentrations, with two independently prepared cell free extracts. Preliminary rate determinations were done to check that the rate of NADH disappearance was linear in respect to enzyme concentrations. A 1mM salicylate solution was prepared by dissolving salicylate in distilled H_2O, while the 1mM substituted salicylate solutions were prepared by first dissolving the substituted salicylic acids in methanol to 1M concentration, and then adding a calculated volume of the methanol to distilled H_2O to a concentration of 1mM. A precipitate appeared in these solutions, which was dissolved by heating at 65°C for 1 hour. Determining the total NADH utilised at assay completion was used to check the substrate concentration in the final assay mixture. For example, if the salicylate concentration in the assay was 100μM then a total change of 0.622 absorbance units at 340nm would be expected since the Molar Extinction Coefficient of NADH is 6220 mol^{-1} cm^{-1}. All substrates assayed were found to be soluble at 1mM. All assay rates were taken as the initial assay velocity over a linear curve.

Spectral scans of all assay substrates and products were performed to check that none of the salicylates absorbed light at 340nm, the wavelength at which NADH disappearance is followed. Spectral scan overlays of the salicylate hydroxylase assay (Fig. 7.4.1) showed a decrease in salicylate concentration (peak at 290nm) and the formation of the catechol and *cis*, *cis*-muconate product (peak at 260nm) when a cell free extract from salicylate-grown ADP1 was added to the biochemical assay.

7.4.2. Stability of SalA enzyme activity.

Biochemical assays were replicated within 48 hours. The SalA activity in cell free extract diminished by <2% following storage at -20°C compared with extract assayed immediately following harvesting.

Fig. 7.4.1.a. Spectral overlay of SalA mediated salicylate disappearance. The disappearance of salicylate catalysed by a cell free extract from *Acinetobacter* sp. ADP1 grown in the presence of 2 mM salicylate. The sample cell contained 50μM salicylate, 50 mM Tris pH 7.5, and 100μM NADH, and the cell free extract, while the reference cell had the same content save salicylate. Spectra were recorded before the addition of cell free extract, and at 0.5, 1.0, 1.5, and 2.0 min following the addition of cell free extract. The disappearance of salicylate (**A**) is followed at 294nm. The formation of a peak at 257nm (**B**) corresponds to the absorbance peak of *cis,cis* muconate. The disappearance of NADH is followed at 340mn (**C**).

Fig. 7.4.1.b. The spectra of 50μM salicylate in 50 mM Tris pH 7.5, which has an absorbance peak at 294nm.

Fig. 7.4.1.c. The spectra of 50μM *cis,cis* muconate in 50 mM Tris pH 7.5, which has an absorbance peak at 259nm.

Fig.7.4.1.a

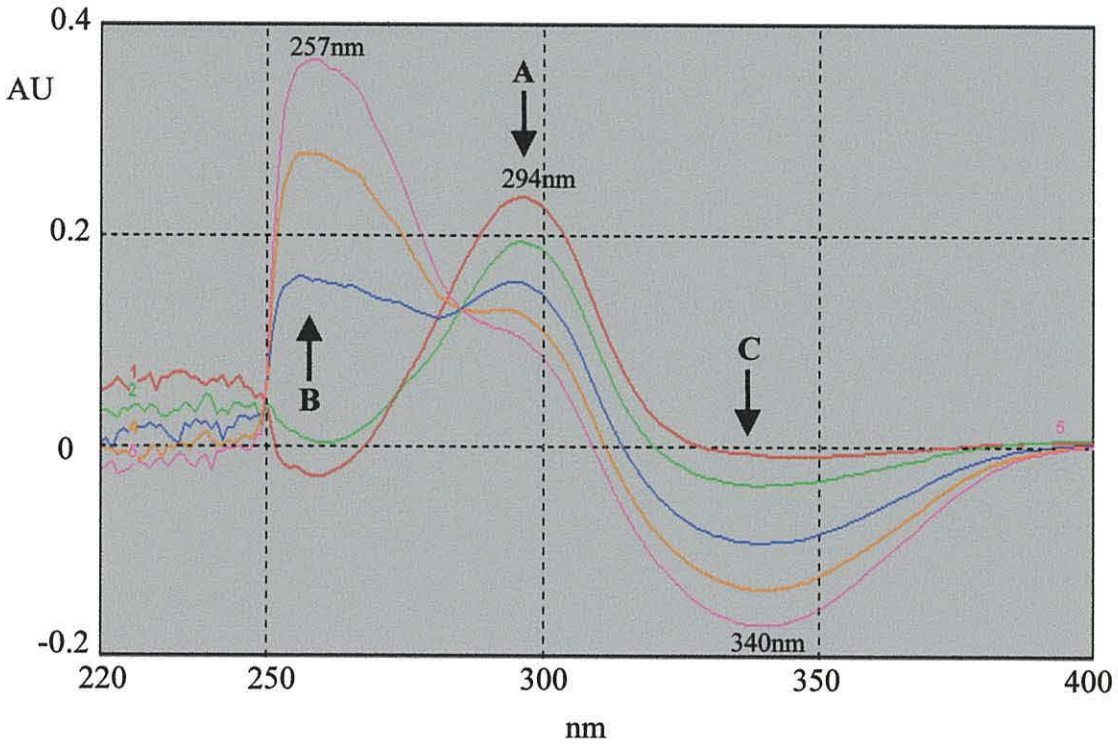


Fig.7.4.1.b

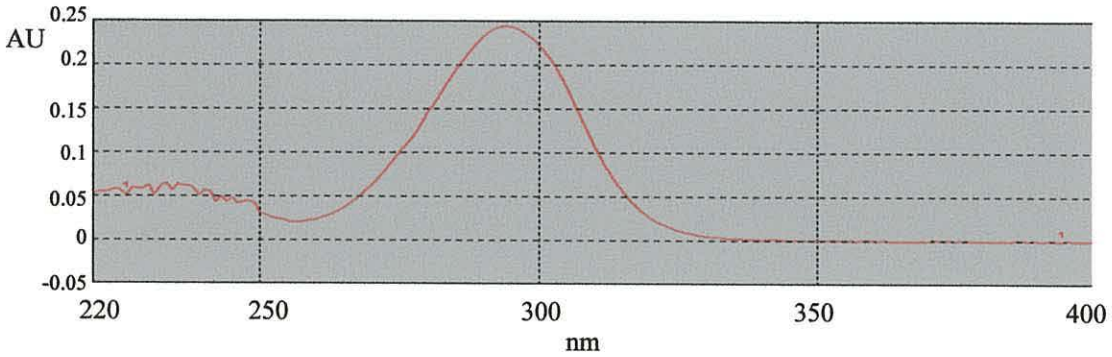
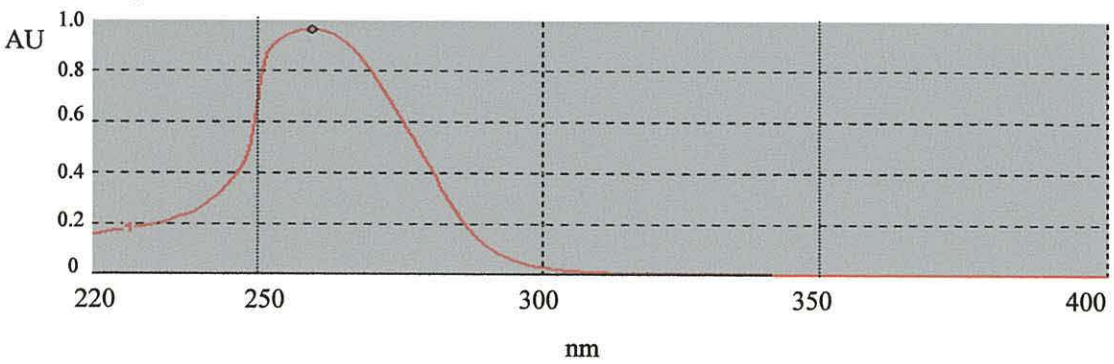


Fig.7.4.1.c



7.4.3. Expression of salicylate hydroxylase in *Acinetobacter* sp. ADP1.

Using the standard NADH-linked assay procedure described in Chapter 2.10.2, salicylate hydroxylase activity was detected in a number of strains (Table 7.4.1). In ADP1, activity was not detectable in succinate-grown cells but was induced by growth on both salicylate and ethyl salicylate and, when grown on succinate in the presence of these, the activity was significantly induced although at a lower level than in the absence of succinate. In both ADPW67 (*salA*::Km^r) and ADPW72 (*salR*::Km^r) no activity was detected when grown on succinate plus salicylate or, for the latter, succinate plus ethyl salicylate. However in ADPW70 (*salE*::Km^r) (Jones *et al.* 2000) there was high induction when grown on salicylate and low, but significant, induction when grown on succinate plus ethyl salicylate which it is unable to metabolise.

7.4.4. Expression of salicylate hydroxylase in *E.coli* (pADPW34).

Using the standard NADH-linked assay procedure described in Chapter 2.10.2, salicylate hydroxylase activity was detected in *E.coli* DH5α pADPW34, which carries *salA, salR* and *salE* (Table 7.4.1.). Moreover, the activity was induced only when 2 mM salicylate was added to the LB medium. pADPW34 contains 440 bps upstream of *salA* that probably contain promoter elements regulating the expression of *salA*. It seems that the regulatory system for expression of SalA is present on pADPW34, which is induced by salicylate.

7.4.5. Specificity of salicylate hydroxylase from *E.coli* (pADPW34) and from *Acinetobacter* sp. ADP1.

The relative activity expressed from cell free extracts of *Acinetobacter* sp. ADP1 and *E.coli* DH5α (pADPW34) against a range of substituted salicylates was measured (Table 7.4.2). The specific activity was 2-fold greater in salicylate-induced *E.coli* (pADPW34) than in salicylate-grown ADP1. The relative activity expressed from the cloned gene against a range of substituted salicylates was compared with that expressed from the wild-type ADP1 and found to be identical within experimental error, having a broad

Table 7.4.1. Specific activity of salicylate hydroxylase in crude extracts of cells grown on different carbon sources. Growth of *Acinetobacter* strains was in minimal medium containing aromatic substrates at 2mM and/or succinate at 10mM as indicated.

Strain	Growth substrate	Sp.Activ U/mg protein[a]
ADP1	succinate	<0.04
ADP1	salicylate	0.70
ADP1	ethyl salicylate	0.30
ADP1	succinate + salicylate	0.55
ADP1	succinate + Ethyl salicylate	0.11
ADPW67	succinate	<0.04
ADPW67	succinate + salicylate	<0.04
ADPW70	succinate	<0.04
ADPW70	salicylate	0.50
ADPW70	succinate + ethyl salicylate	0.09
ADPW72	succinate	<0.04
ADPW72	succinate + salicylate	<0.04
ADPW72	succinate + ethyl salicylate	<0.04
E.coli (pADPW34)	LB + salicylate	1.37
E.coli (pADPW34)	LB	<0.04

[a]Reaction rates were based on the average of three measurements, none of which varied by >5%.

Table 7.4.2. Relative activities of SalA salicylate hydroxylases in crude extracts of cells.

	Relative activity[a] in cell extracts of:	
Assay substarte	ADP1	DH5∀ (pADPW34)[c]
Salicylate	100	100
3-Methyl salicylate	>0.1	>0.1
4-Methyl salicylate	56	54
5-Methyl salicylate	35	33
4-Chlorosalicylate	45	40
5-Chlorosalicylate	27	22

[a] Activity relative to salicylate which is set at 100. All reaction rates were based on the average of three measurments, from two independently prepared cell free extracts, none of which varied by >5%.

[b] ADP1 was grown on minimal medium containing 2 mM salicylate: the specific activity was 0.71 U/mg protein.

[c] *E.coli* DH5∀ (pADPW34) was grown on LB containing 2 mM salicylate: the specific activity was 1.37 U/mg protein.

[d] No salicylate hydroxylase activity was detected in extracts of *E.coli* DH5∀ (pADPW34) when grown on LB alone, nor in *E.coli* DH5∀ extracts when grown on LB containing 2 mM salicylate.

substrate specificity except against the only available 3-substituted salicylate (Table 7.4.2).

7.5 Discussion.

7.5.1. SalA is responsible for the conversion of salicylate to catechol in ADP1.

Salicylate hydroxylase, SalA, is responsible for the catabolism of salicylate to catechol (Fig.7.5.1). This is the route by which salicylate is fed into the β-ketoadipate pathway. Salicylate is a growth substrate for *Acinetobacter* sp. ADP1 and a functional *salA* gene product is necessary for growth. Also, 2-hydroxybenzyl (salicyl) acetate is a growth substrate for *Acinetobacter* sp. ADP1. By the sequential actions of AreA, AreB, and AreC, 2-hydroxybenzyl acetate is catabolised to 2-hydroxy benzoate (salicylate) and SalA provides the route by which 2-hydroxybenzyl acetate feeds into the β-ketoadipate pathway. In this study SalA has been expressed from a pUC18 clone carrying *salERA*. The SalA salicylate hydroxylase activity shows the same relative substrate preferences as the specificity found in wild type ADP1 grown on salicylate itself.

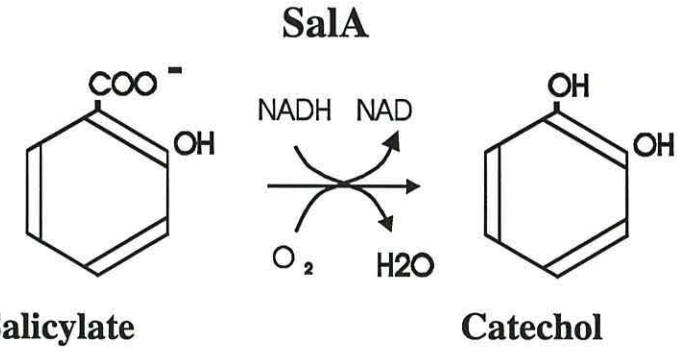


Fig. 7.5.1. SalA catalysed conversion of salicylate to catechol in *Acinetobacter* sp. ADP1.

An insertional knock-out of *salA* shows the phenotype expected from the proposed pathway (Fig.7.5.1). The further catabolism of the aromatic moiety is into the *ben-cat* branch of the β-ketoadipate pathway at the level of catechol. The confirmation of this

will be discussed in Chapter 9, where an ADP1 with a mutation in *catA* (for catechol 1,2-dioxygenase) is tested for the ability to grow on salicylate.

The SalR protein is a member of the LysR family of regulator proteins. The signature motif contains the DNA-binding helix-turn-helix at residues 17-47 (NISKAAEILNLSQPSVTYNLNRLRKHLNNPL). The close similarity of both it and SalA to the two *Pseudomonas* isofunctional proteins NahR and NahG from the *P.putida* NAH7 plasmid (Schell & Poser; 1989, You *et al.* 1988) and *P.stutzeri* (Bosch *et al.* 1999) imply the past intergeneric exchange of genes by horizontal transfer.

The conservation of the amino acid sequence between the genes has been maintained with a high amino acid % similarity. However in spite of the retention of amino acid sequence, the DNA composition has equilibrated to become more similar to that of the general host in terms of AT/GC ratio. Whereas the *Pseudomonas* genes have a G+C ratio of between 60 to 65%, the *salA* and *salR* genes have a composition more characteristic of the A+T-rich *Acinetobacter* genome (Table 7.2.2). This has also been reported by Ornston and his colleagues with other genes of the β-ketoadipate pathway, the gene products of which in *Pseudomonas* and *Acinetobacter* share considerable homology but the nucleotide compositions are clearly characteristic of the different chromosomal environment of each host (Harayama *et al.* 1991).

7.5.2. Regulation of *sal* genes.

SalA activity is induced during growth of ADP1 on salicylate (Table 7.4.1), hence the expression of *salA* must be regulated. SalR has high sequence homology to other salicylate hydroxylase LysR type proteins that regulate SalA expression (Table 7.4.1). What evidence is there to prove that SalR is the protein that regulates expression of SalA in *Acinetobacter* sp. ADP1? There are three findings that support the hypothesis:

(a) Insertion of a Km^r cassette into *salR* in strain ADPW72 (*salR*::Km) stops induction of salicylate hydroxylase activity by salicylate.

(b) Salicylate hydroxylase is also salicylate-inducible in *E.coli* (pADPW34) which carries only *salARE*.

(c) Close protein homologies exist between *salR*, *salA* and the *Pseudomonas* homologues *nahR* and *nahG*.

A major difference between the *sal* and *nah* genes is that the *P.putida nahR* is transcribed divergently from its adjacent catabolic gene *nahG*. However *salA* and *salR* are cotranscribed as a single regulatory unit as shown by the RT-PCR results (Chapter 8.4). This implies that SalR thus controls its own expression.

7.5.3. Analysis of the *salA* promoter region.

There is homology between the DNA upstream of *salA* in *Acinetobacter* sp. ADP1 with the regions of the NAH7 plasmid from *P.putida* identified as being the site at which NahR binds to the DNA (Schell & Poser, 1989). The sequence upstream of *nahAa* (napthalene dioxygenase from NAH7) which binds to NahR is TCA-N_6-TGA, and the sequence upstream of *nahG* (salicylate hydroxylase from NAH7) which also binds to NahR is the sequence TCA-N_3-TGATGA (Schell & Poser, 1989) (GenBank accession no. M11863). Upstream of *nahG* (salicylate hydroxylase) in *P.stutzeri* is the sequence TCA-N_3-TGAATA. Upstream of *salA* in *Acinetobacter* sp. ADP1 is the sequence TCA-N_3-TGATGT. Alignment of these sequences shows that the sequence TCA-N_3-TGA is conserved (Fig. 7.5.2.). This implies that the conserved residues are involved in regulator binding. Just downstream of the TCA-N_3-TGA there is also a -35 sequence TAGGCA that shares homology with the paradigm -35 sequence TTGACA (deHaseth *et. al.* 1998) (Fig.7.5.3.). Further downstream there is a –10 sequence TGATATAAT which shares perfect homology to the consensus extended –10 sequence TGNTATAAT (deHaseth *et. al.*1998) (Fig.7.5.3.). The presence of a conserved -10, -35 region also adds weight to the hypothesis that the conserved TCA-N_3-TGATGT is the regulator binding site. Nevertheless, conclusions cannot be drawn by sequence gazing. To prove the hypothesis, studies must be undertaken that measure the expression of SalA when the hypothesised regulator binding site has been mutated.

Fig. 7.5.2. Alignment of the DNA sequence upstream of *salA* in *Acinetobacter* sp. ADP1 (ADP1). The alignment against the *salA* upstream region was done with the corresponding upstream region of: *nahG* from *P.stutzeri* (P.stutz.); *nahG* from *P.putida* (NAH7); and orf7 from the vanillate cluster of genes in *Acinetobacter* sp. ADP1 (vanorf7). The regulator binding site for *nahR* upstream of *nahG* from *P.putida* (NAH7) is underlined (Schell & Poser, 1990). Highlighted bases are identical in all four sequences, and shadowed sequences are identical in three or less of the sequences.

ATGTTTTAAGATCATAATGATGGATATGAAAACTTAAACCTATACCGGATAGGCAATTAA-

Putative Regulator binding **-35**

GAAAGAAATATGATATAATTGTTTTAAGATATT - 100bps - CTCAAAGGAAATGAGTCGTG

ext. -10 **Shine-Delgano** → **Start SalA**

Fig. 7.5.3. Putative promoter region upstream of *salA* in *Acinetobacter* sp.ADP1.

CHAPTER 8

SalE is responsible for the catabolism of salicylate esters in *Acinetobacter* sp. ADP1

8.1 Introduction.

The initial laboratory work on characterising the *salE* gene product was done as part of a third year honours project by Vassillis Pagmantidis [BSc (Hons) 1999, University of Wales, Bangor] under the working supervision of Rheinallt M. Jones. V. Pragmatidis created ADPW70 (*salE*::Km^r) with an insertion within *salE*, and over-expressed recombinant DNA containing *salE* in pADPW70 (Table 8.1.1). Growth tests of ADPW70 determined that this mutant could not utilise ethyl salicylate as a sole carbon source while ethyl salicylate supports vigorous growth of ADP1 (Jones *et al.* 2000). Also, V. Pagmantidis determined the specificity of over-expressed SalE and AreA against 4-nitrophenyl esters of acyl chain length ranging from 2 carbon to 16 carbon, and the K_m and V_{max} of over-expressed SalE against 4-nitrophenyl acetate and butyrate (Jones *et al.* 2000).

The remainder of the investigation into the function of SalE is the subject of this chapter. Wild type *Acinetobacter* sp. ADP1 cell free extracts, and over-expressed SalE cell free extracts were assayed against ethyl salicylate in an NAD^+- linked assay. Also RT-PCR was carried out across the intergenic region between *salA* and *salR*, and across the intergenic region between *salD* and *salE* genes to determine whether they are inducible and transcribed on a single mRNA as implied by the gene organisation.

8.2. The *salE* gene function.

8.2.1. Sequence analysis of *salE* by database similarity searches.

The *salE* gene was first cloned and sequenced (Chapter 7.2.1) on pADPW34 (Fig. 8.1.1). *salE* is located downstream of *salD*, and its product shows homology with serine esterases of the α/β-hyrolase family (Ollis *et al.* 1992). SalE has very limited similarity to

TABLE 8.1.1. Bacterial strains and plasmids

Strains/plasmids	Genotype/phenotype	Reference/ source
Acinetobacter ADP1 (BD413)	Wild type	(Juni, 1969)
ADPW67	*salA*::Km^r; transformation of ADP1 with pADPW44	This study
ADPW70	*salE*::Km^r; transformation of ADP1 with pADPW76	This study
ADPW72	*salR*::Km^r; transformation of ADP1 with pADPW79	This study
E. coli DH5α	F$^-$, ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsd*R17 (r_K^-, m_K^+), *phoA*, *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	This study
BL21(DE3)pLysS	F$^-$, *omp*T, *hsd*S_B(r_B^-, m_B^-), *dcm*, *gal*, (DE3), pLysS, Cm^r	Stratagene
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and the whole of *areA* in pUC18	This study
pADPW34	4.1-kbp *SacI*/*Xba*I subclone of pADPW32 in pUC18	This study
pADPW49	1.0-kbp *Eco*RI* fragment containing *salE* in pUC18	This study
pADPW70	*Nde*I*/*Eco*RI* fragment containing *salE* in pET5a	This study
pADPW76	pADPW49 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salE*	This study
	* denotes restriction site added by PCR	

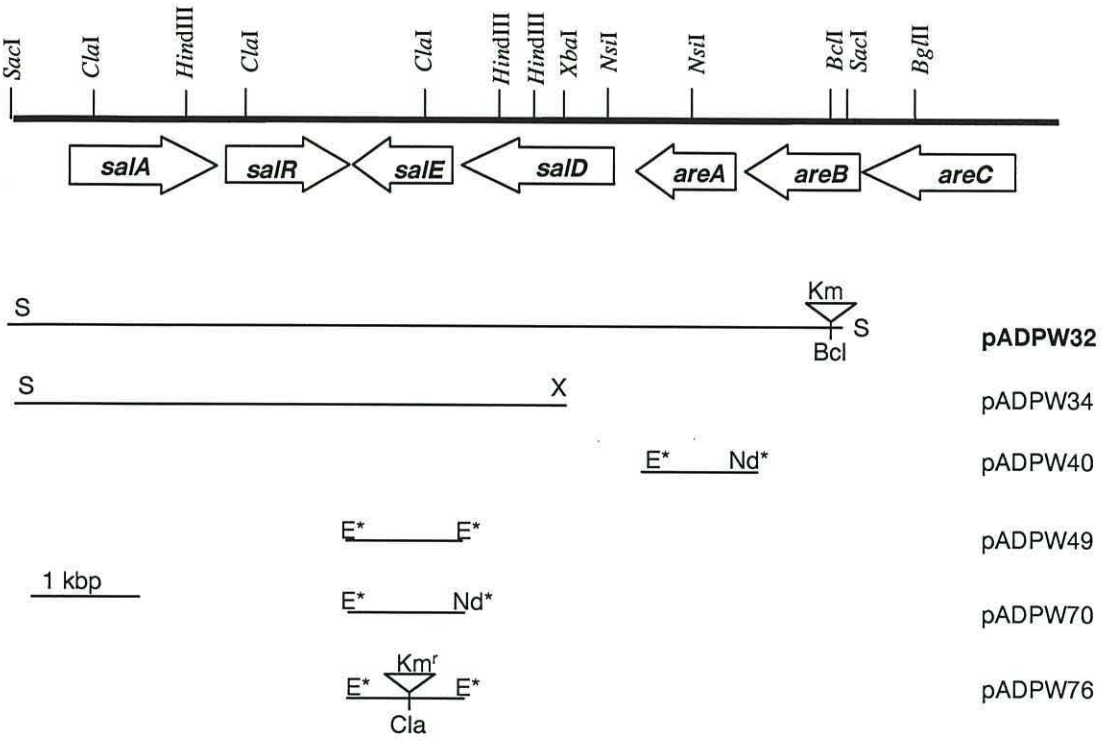


Fig. 8.1.1 Physical map of the *salA, salR, salE* and *salD* genes and their location relative to *areABC* at the left hand end (as drawn) of the supraoperonic *ben-cat* cluster. The inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 8.1.1. pADPW32 named in boldface type was cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA (denoted by an asterisk) or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The Km^r cassette insertions are not to scale. The abbreviations for the restriction sites are: E, *Eco*RI; H, *Hin*dIII; Nd, *Nde*I; S, *Sac*I; X, *Xba*I.

the benzyl esterase AreA upstream of it. Examination of the deduced amino acid sequence of SalE in PROSITE (Bairoch *et al.*1997) shows that from residues 68-77 (IVLLGHSYGG) it has the signature characteristic of serine lipases [LIV]-x-[LIVFY]-[LIVMST]-G-[HYWV]-S-x-G-[GSTAC] in which the serine is the active site nucleophile. The primary sequence of SalE does not align closely with other reported *Acinetobacter* esterases from *A. lwoffii* RAG-1 (Alon & Gutnick, 1993) nor the two esterases from strain ADP1, a carboxylesterase (Kok *et al.* 1993) and the adjacent benzyl esterase AreA (Jones *et al.* 1999), which are of different lengths and have their putative active site serines further from the N-terminus.

8.2.2. Insertional Inactivation of *SalE.*

The insertional inactivation of *salE* was done by V. Pagmantidis (Jones *et al.* 2000). A PCR generated-fragment carrying *salE* was used to create pADPW49, and then the insertion of a Km^r resistant cassette from pUI1637 was inserted (Eraso & Kaplan, 1994) into the sole *Cla*I site in pADPW49 to create pADPW76 (Fig. 8.1.1). pADPW76 was linearised and transformed into ADP1 to create ADPW70 (*salE*::Km^r).

Because SalE showed protein homologies with other esterases, ADPW70 and ADP1 were tested for their abilities to grow on a range of esters containing aromatic components both as the alcohol and as the acid moiety and also a number of exclusively aliphatic esters (Jones *et al.* 2000). These include *n*-propyl acetate, benzyl acetate, *n*-butyl propionate, ethyl propionate, benzyl propionate, ethyl valerate, ethyl butyrate, benzyl butyrate, ethyl caproate, ethyl benzoate, *n*-propyl benzoate, *n*-butyl benzoate, ethyl benzoylacetate and six were substrates for neither ethyl acetate, *n*-butyl acetate, *n*-butyl butyrate, benzyl benzoate, *n*-propyl cinnamate, benzyl salicylate. Of the range of ester compounds, only ethyl salicylate and methyl salicylate supported growth of ADP1 but did not support growth of ADPW70 suggesting that SalE is a hydrolase specific for these esters. Neither ADPW67 (*salA*:Km), ADPW72(*salR*:Km) nor ADPW78(*salD*:Km) (Table 8.1.1) were able to grow on the two alkyl salicylates.

8.2.3. The over-expression of the *salE* gene product in *E.coli.*

salE was over-expressed by V. Pragmantidis (Jones *et al.* 2000) by subcloning the *Nde*I/*Eco*RI fragment from pADPW49 into the pET5a high expression vector to create pADPW70 (Fig. 8.1.1.). SalE was induced from pADPW70 with IPTG, and the resulting cell free extract was run on a SDS-PAGE electrophoresis gel (Fig 8.2.1). A strong protein band with the expected molecular weight of 27 kDa corresponds to predicted weight of SalE.

8.3. Analysis of the function of SalE by a biochemical NAD^+ linked assay.

8.3.1. Verification of the quality of the Ethyl Salicylate Linked Assay.

An assay was set up to test the activity of SalE against ethyl salicylate. This was done by linking the products of ethyl salicylate cleavage (salicylate and ethanol) with yeast alcohol dehydrogenase (ADH) (Chapter 2.10.3.). The assay components were 50 mM Tris-HCl Buffer, 100μM ethyl salicylate, 10 units yeast ADH, 2mM NAD^+, and the reaction was started by addition of SalE. Ethyl salicylate was redistilled under reduced pressure to remove small amounts of contaminating ethanol. Rates were confirmed as <0.001 Δabs/min prior to the addition of SalE into the assay mixture. The spectra of ethyl salicylate and salicylate were determined to check that neither absorbed at 340nm.

The solubility and concentration of ethyl salicylate was determined from the total NADH produced at assay completion. A 100μM concentration of ethyl salicylate produced a net change of 0.622 absorbance units as would be expected. This also verifies the stoichiometry of one molecule of NADH produced per one molecule of ethyl salicylate cleaved. All assay rates were taken as the initial assay velocity whilst the trace was still linear.

In addition, spectral scan overlays of the progress of the linked assay were carried out (Fig 8.3.1). The scans showed a disappearance in ethyl salicylate concentration (peak at 302nm), the formation of the salicylate product (peak at 294nm), and the linked formation of NADH (peak at 340nm), when cell extracts from IPTG-induced *E.coli* (pADPW70) or from ethyl salicylate-grown *Acinetobacter* sp. ADP1 were added to the

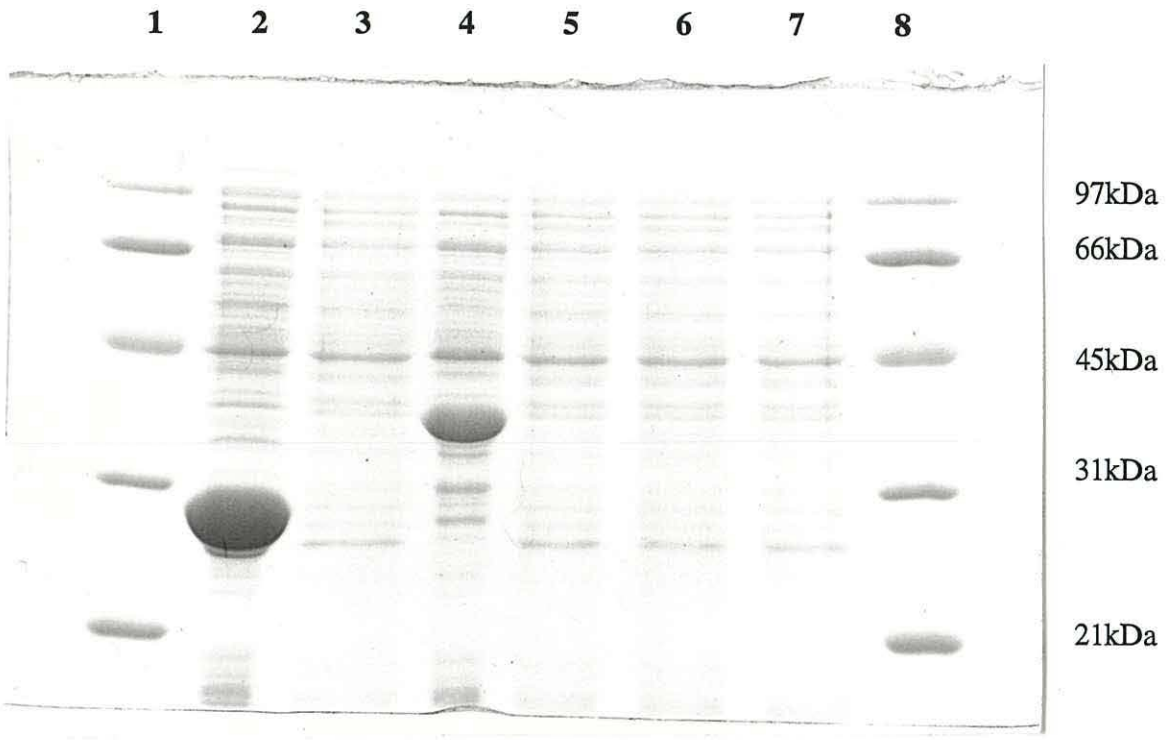


Fig. 8.2.1. SDS-PAGE of over-expressed *SalE and AreA* proteins. The lanes contain lysates from *E.coli* BL21(DE3)pLysS carrying the following plasmids, induced with IPTG (I) or uninduced (U). Lanes: 1, 8, molecular weight standards; 2, pADPW70 (I) ; 3, pADPW70 (U) ; 4, pADPW40 (I); 5, pADPW40 (U); 6, pET5a ; 7, no plasmid. The molecular weight standards are (in kDa): A, 97.4; B, 66.2; C, 45.0; 31.0; E, 21.5. The estimated M_r values for the overexpressed bands are SalE (lane 2) 27 kDa, AreA (lane 4) 37 kDa. Each lane contains approximately 50μg of protein.

Fig. 8.3.1.a. Spectral overlay of the Ethyl salicylate/NAD^+ linked assay.

The disappearance of ethyl salicylate is catalysed by a cell free extract from IPTG induced *E.coli* (pADPW70). The sample cell contained 100μM ethyl salicylate, 100 mM pH 8.0 phosphate buffer, and 2 mM NAD^+, 10U NAD^+, and the cell free extract, while the reference cell had the same content save ethyl salicylate. Spectra were recorded before the addition of cell free extract, and at 0.5, 1.0, 1.5, and 2.0 min following the addition of cell free extract. The disappearance of ethyl salicylate (**A**) is followed at 302nm. The formation of a peak at 294nm (**B**) corresponds to the absorbance peak of salicylate. The formation of NADH is followed at 340nm (**C**).

Fig. 8.3.1.b. The spectra of 100μM ethyl salicylate in 50 mM Tris pH 7.5, which has an absorbance peak at 302nm.

Fig. 8.3.1.c. The spectra of 100μM salicylate in 50 mM Tris pH 7.5, which has an absorbance peak at 294nm.

Fig. 8.3.1.a.

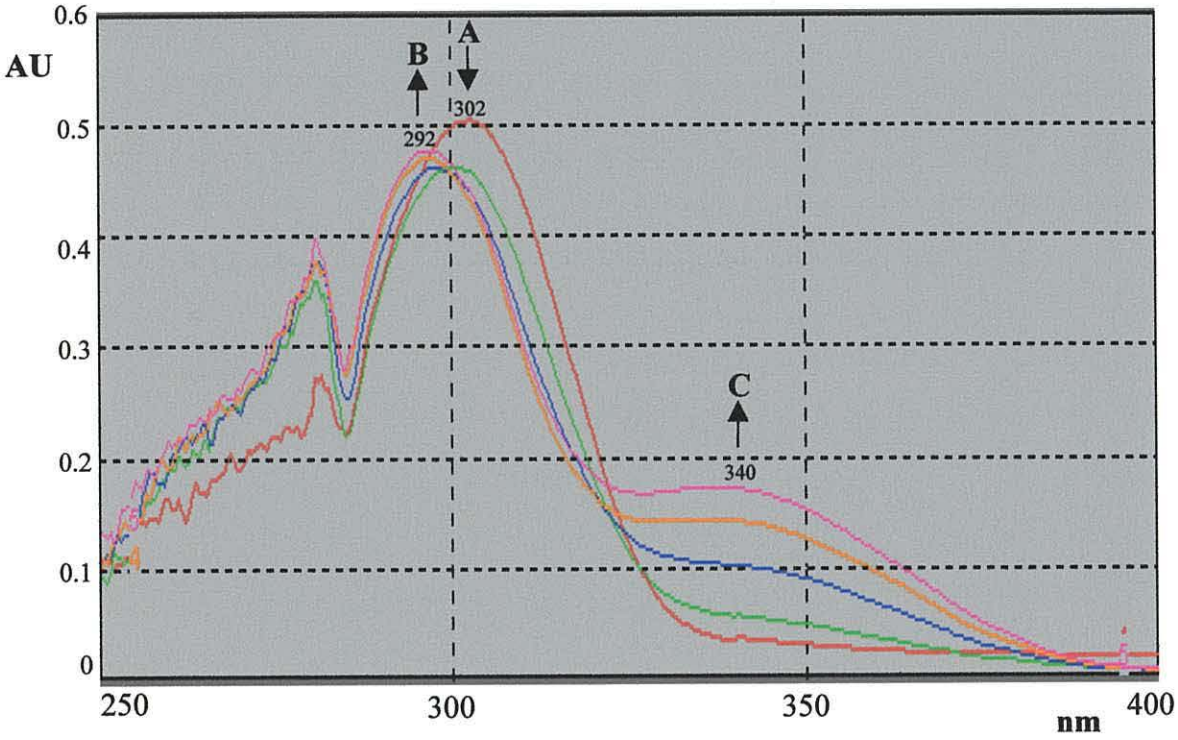


Fig. 8.3.1.b.

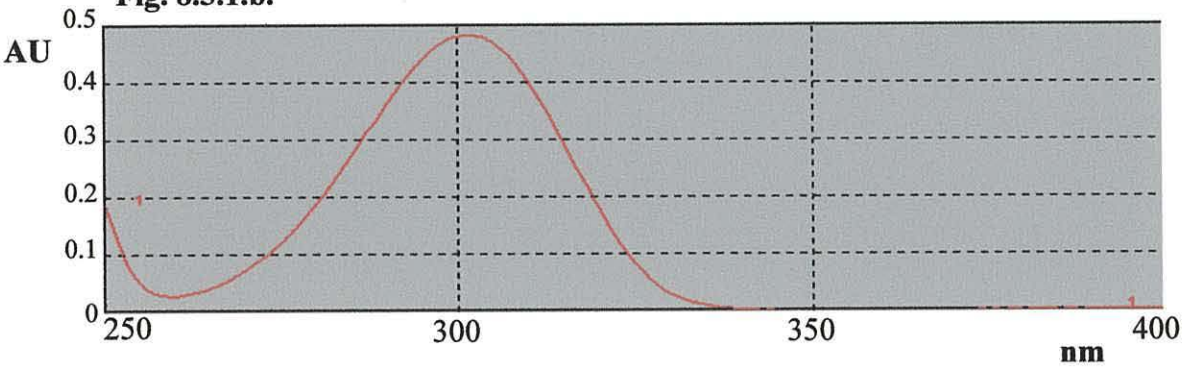
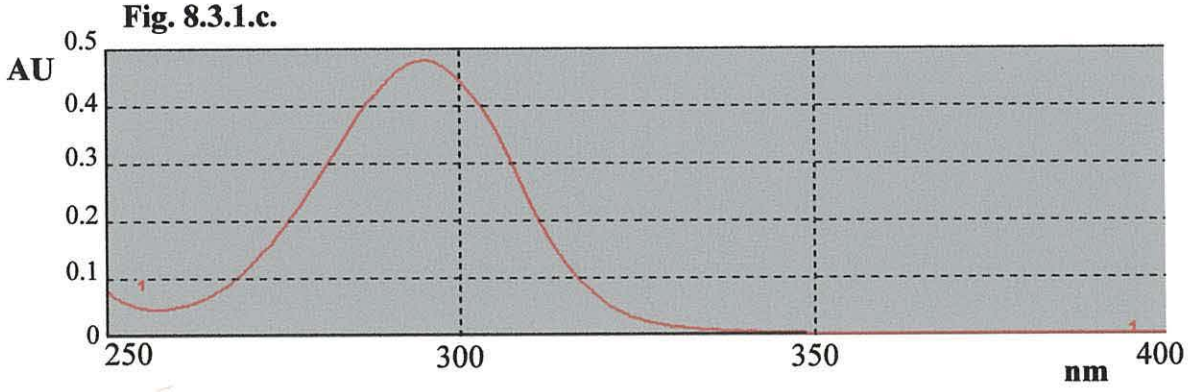


Fig. 8.3.1.c.



assay (Fig. 8.3.1.). No diminution in ethyl salicylate concentration occurred when cell extracts from uninduced *E.coli* (pADPW70), from salicylate-grown ADP1, or from ADPW70 (*salE*::Km^r) grown on succinate in the presence of ethyl salicylate were added to the biochemical assay.

8.3.2. Biochemical assays of SalE from IPTG-induced *E.coli* (pADPW70) against ethyl salicylate.

pADPW70 (Table 8.1.1) (Fig. 8.1.1) was transformed into *E.coli* BL21 (DE3) pLysS and expressed by growth in 500 ml of LB plus ampicillin to an O.D. of 0.6 at 600 nm, and then inducing for 4 h by the addition of 0.4 mM IPTG. Cell free extracts were prepared as described in Chapter 2.9. SDS-PAGE of extracts of the induced *E.coli* hosts showed high levels of expression of SalE with the predicted size (Fig. 8.2.1.).

The NAD$^+$-linked assay was carried out against ethyl salicylate using various cell free extracts. The determination of the initial rate velocity at 340nm showed that, after IPTG induction, the *E.coli* (pADPW70) contained high hydrolytic activity against ethyl salicylate (Table 8.3.1). No hydrolytic activity against ethyl salicylate was detected in the uninduced *E.coli* (pADPW70) cell free extract. The specific activity of IPTG-induced *E.coli* (pADPW70) was 15 times higher than that measured in ethyl salicylate-grown cell free extracts of *Acinetobacter* sp. ADP1 (Table 8.3.1).

8.3.3. Biochemical assays of SalE from *Acinetobacter* sp. ADP1.

Using the standard alcohol dehydrogenase /NAD$^+$-linked assay procedure (Chapter 2.10.3.), esterase activity against ethyl salicylate was detected in a number of strains (Table 8.3.1). In ADP1, SalE activity was not detectable in succinate-grown cells nor in salicylate-grown cells, but was induced by growth on ethyl salicylate or when grown on succinate plus ethyl salicylate. When ADP1 was grown on succinate plus ethyl salicylate, the SalE activity was significantly induced although to a six-fold lower specific activity than in the absence of succinate. No SalE activity was detected in ADPW70 (*salE*::Km^r) cells when grown on succinate, salicylate, nor succinate plus ethyl salicylate.

Table 8.3.1. Specific activities of SalE salicylate esterase in crude extracts of cells grown on different media.

Strain	Growth substrates	Specific Activity[a] (U/mg of protein)
ADP1	succinate	<0.04
ADP1	salicylate	<0.04
ADP1	ethyl salicylate	3.44
ADP1	succinate + ethyl salicylate	0.67
ADPW70	succinate	<0.04
ADPW70	salicylate	<0.04
ADPW70	succinate + ethyl salicylate	<0.04
ADPW72	succinate	<0.04
ADPW72	succinate + salicylate	<0.04
ADPW72	succinate + ethyl salicylate	0.45
E.coli (pADPW70)	LB + IPTG	52.49
E.coli(pADPW70)	LB	<0.04

[a] All reaction rates were based on the average of three measurements, none of which varied by >5%.

No SalE activity was detected when ADPW72 (*salR*::Km^r) was grown on succinate nor when grown on succinate plus salicylate. However SalE activity was detected in succinate plus ethyl salicylate-grown ADPW72 (*salR*::Km^r), although at a seven fold lower specific activity.

8.4. RT-PCR analysis of ADP1 ethyl salicylate induced transcripts.

In order to show whether the *salA* and *salR* genes, and *salD* and *salE* genes are inducible and transcribed on a single mRNA as implied by the gene organisation, reverse transcriptase (RT)-PCR analysis was carried out across the intergenic regions between the two sets of genes.

8.4.1 Purification of RNA from *Acinetobacter* sp. ADP1.

Total RNA was isolated from *Acinetobacter* sp. ADP1 cells were grown on minimal medium containing either succinate, salicylate or ethyl salicylate, and total RNA was prepared as described in Chapter 2.8.

8.4.2. RT-PCR analysis of the intergenic region between *salA* and *salR*.

RT-PCR amplifications were carried out with the Access RT-PCR kit (Promega), across the intergenic region between *salA* and *salR* using primer pairs (*salAR* forward) 5'-CCATGGACACGTGCGGTAGAC-3'; (*salAR* reverse) 5'-TTTTTGGTGCATGTGCTCGTAAGT-3' (Fig. 8.4.1.). RT-PCRs were carried out as described in Chapter 2.8.2. Negative control reactions to eliminate the possibility of amplifying residual genomic DNA were performed in the same way, except that the reverse transcriptase was omitted from the reaction mixtures.

The expected RT-PCR size of the *salA-salR* amplicon is 988 bp. The PCR products obtained, together with restriction digests chosen to confirm the presence of the restriction sites, were analysed by agarose gel electrophoresis (Fig. 8.4.1.). The products obtained across *salAR* from the total RNA of cells grown on both ethyl salicylate and

salicylate were of the expected sizes. The presence of restriction sites in the expected positions within the fragments also confirmed that the amplified fragment was obtained from the *salAR* region. No products were obtained from total RNA of succinate-grown cells or from reaction mixtures from which the reverse transcriptase had been omitted.

8.4.3. RT-PCR analysis of the intergenic region between *salE* and *salD*.

RT-PCR analysis of *salE* and *salD* was done with the Access RT-PCR kit (Promega) on total RNA isolated from succinate-, salicylate-, and ethyl salicylate-grown cells by amplifications across the intergenic regions between *salE* and *salD* using primer pairs (*salED* forward) 5'-AGATTTAGGTATTCAGCAATTCAGGGCAAAAGGTG-3'; and (*salED* reverse) 5'-AAGGCTCAGGCGTAAGCATCTTGTAAGTTTCCTC-3' (Fig. 8.4.1.b). PCR was done as described in Chapter 2.8.2. Negative control reactions to eliminate the possibility of amplifying residual genomic DNA were performed in the same way, except that the reverse transcriptase was omitted from the reaction mixtures.

The PCR products obtained, together with restriction digests chosen to confirm the presence of expected restriction sites, were analysed by agarose gel electrophoresis (Fig. 8.4.1.a.). An 1195 bp amplification product was obtained only from the total RNA of cells grown on ethyl salicylate. The presence of a restriction site in the amplified fragments confirmed that the product was amplified from the *salDE* region. No products were obtained from total RNA of succinate-grown, or salicylate-grown cells, or from reaction mixtures from which the reverse transcriptase had been omitted.

8.5. Discussion.

8.5.1. The catabolic role of the *sal* gene products.

The salicylate esterase, SalE, and salicylate hydroxylase, SalA catalyse the sequential catabolism of alkyl salicylates via salicylate to catechol (Fig 8.5.1). This is another route into the β-ketoadipate pathway for ester substrates. In this study, SalE has been expressed from its cloned gene in the expression vector pET5a where it is expressed to very high

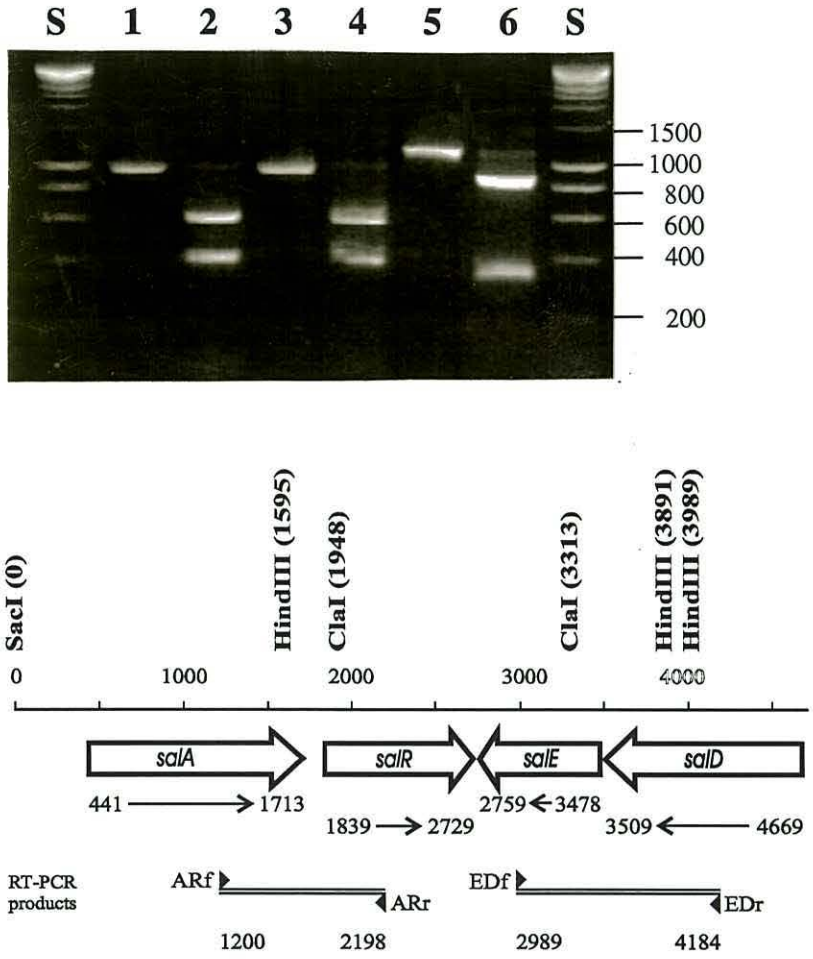


Fig. 8.4.1. Agarose gel electrophoresis of RT-PCR products amplified from ADP1 grown on ethyl salicylate and salicylate. The sizes of molecular size markers (in bp) in lanes S (HyperLadder I, from Bioline, London, UK) are indicated by arrows (1,500 1,000, 800, 600, 400 and 200 bp). Lanes: 1, *salAR*, salicylate-grown cells (expected size, 988 bp); 2, *salAR*, salicylate-grown cells digested with *Hin*dIII (582 and 406 bp); 3, *salAR*, ethyl salicylate-grown cells (expected size, 988 bp); 4, *salAR*, salicylate-grown cells digested with *Hin*dIII (582 and 406 bp); 5, *salED*, ethyl salicylate-grown cells (expected size, 1195 bp); 6, *salED*, ethyl-salicylate grown cells digested with *Cla*I (860 and 335 bp). No detectable products were obtained in control reactions, with each pair of primers, from which reverse transcriptase had been omitted or in reactions carried out on succinate-grown cells (data not shown).

specific activity. Characterisation of SalE from *Acinetobacter* sp. ADP1 also revealed its inducible ability to convert ethyl salicylate to salicylate and ethanol.

In addition, an insertional knock-out of *salE* (ADPW70) and of *salA* (ADPW67) showed the phenotype expected from the proposed pathway (Fig.8.5.1) where neither mutant could grow on ethyl salicylate as a sole carbon source. The catechol formed by SalA activity is probably further catabolised by the *cat* genes of the *ben-cat* branch of the β-ketoadipate pathway. This will be confirmed in Chapter 9 where *catA* ADP1 mutants will be tested for growth on salicylate and ethyl salicylate.

The location of the *sal* genes is almost directly adjacent to the *areCBA* operon (Chapter 3) (Jones *et al.* 1999), the role of which is to channel benzyl alkanoates into the β-ketoadipate pathway. The two sets of genes thus appear complementary in that the *are* genes are responsible for the catabolism of esters in which the alcohol moiety is aromatic whereas the *sal* genes encode the catabolism of esters in which the acid moiety is aromatic.

8.5.2. Induction of SalE and SalD is not regulated by SalR.

It is known that NahR regulates the lower (salicylate) and upper (napthalene) pathway in *P.putida* NAH7 plasmid (Schell, 1993). Since there is no other obvious regulator in the immediate vicinity, this leads to speculation whether SalR regulates both *salAR* and *salDE*. However, data suggest that this is not the case. RT-PCR analysis showed that *salD* and *salE* are cotranscribed and induced by ethyl salicylate, but are not induced by salicylate. The induction results show that growth on ethyl salicylate is necessary for induction of SalE activity and that salicylate does not act as an inducer (Table 8.3.1). Further, in ADPW72 (*salR*::Km^r) *salE* remains inducible by ethyl salicylate, and analysis of the ADPW72 growth media revealed an accumulation of salicylate. No obvious binding sites (TCA-N_3-TGA or TCA-N_6-TGA) for SalR were identified upstream of *salD*. Hence, there is no evidence that SalR regulates induction of SalE. There must be an additional regulator gene on the ADP1 chromosome involved in the induction of *salDE*, but not located in the immediate vicinity.

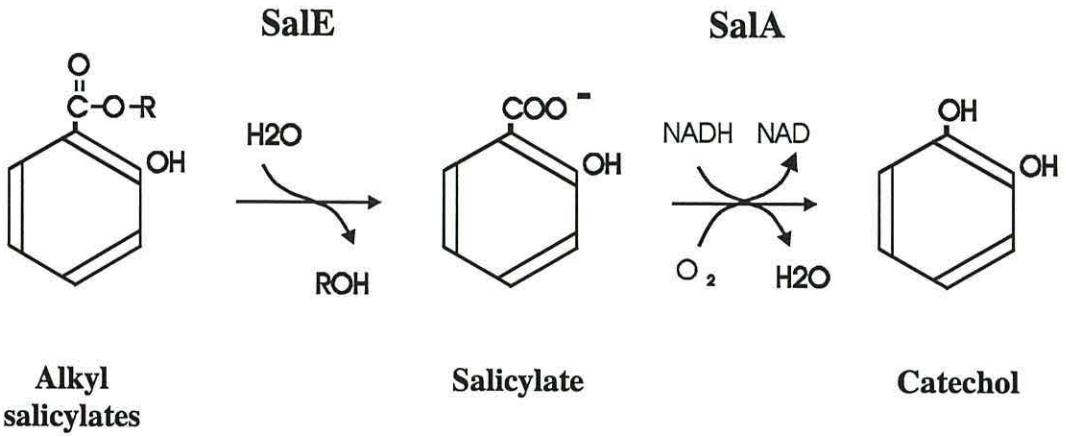


Fig. 8.5.1. Proposed pathway for the catabolism of alkyl salicylates in *Acinetobacter* sp. ADP1.

CHAPTER 9

Phenotypic analysis of ADP1 mutants with disruptions in *cat*, *ben*, *are* or *sal* genes

9.1. Introduction.

In this chapter, it will be shown that the aromatic moieties of aryl esters released by the hydrolytic activity of AreA, and of salicylate esters released by the hydrolytic activity of SalE are both channelled into the β-ketoadipate pathway. In addition, it will be shown that the substrate at which the aromatic moieties enter the β-ketoadipate pathway is dependent on the presence or absence of a hydroxyl side group on the aromatic ring.

The strategy taken was to test the growth phenotypes of three mutants disrupted in genes coding for a step in the β-ketoadipate pathway. The three mutants are (Table 9.1.1): ADP6 (*pcaG*$^-$) which does not grow on 4-hydroxybenzoate or any substrate which feeds into the protocatechuate branch of the pathway; ADPW1 (*catA*$^-$) which has a functionless catechol 1,2-dioxygenase that blocks the utilisation of benzoate or catechol, and accumulates catechol as a brown colouration from any substrate which feeds into the catechol branch; and ADPW38 (*ben*$^-$) with an uncharacterised lesion in *benABC* which is unable to convert benzoate to catechol (Table 9.1.1).

Further, the ability of double mutants to grow on the same set of substrates was examined. These were created using the kanamycin resistant cassette-insertions in *salE* (pADPW76) and *areA* (pADPW37) (Table 9.1.1.). The plasmids were individually introduced into ADP6, ADPW1 and ADPW38 by natural transformation and the double mutants selected for by the acquisition of kanamycin resistance (Table 9.1.1.).

Table 9.1.1. Bacterial strains and plasmids.

Strains/plasmids	Genotype/phenotype	Reference/ Source
Acinetobacter ADP1 (BD413)	Wild type	Juni, 1969
ADPW1	*catA*$^-$	Williams & Shaw, 1997
ADPW38	*ben*$^-$	Williams & Shaw, 1997
ADP6	*pcaG*$^{--}$	Gerischer & Ornston, 1995
ADPW56	*areC*::Km^r; transformation of ADP1 with pADPW22	This study
ADPW58	*areA*::Km^r; transformation of ADP1 with pADPW37	This study
ADPW67	*salA*::Km^r; transformation of ADP1 with pADPW44	This study
ADPW68	*areB*::Km^r; transformation of ADP1 with pADPW75	This study
ADPW70	*salE*::Km^r; transformation of ADP1 with pADPW76	This study
ADPW72	*salR*::Km^r; transformation of ADP1 with pADPW79	This study
ADPW76	*areA*::Km^r, *pcaG*$^-$; transformation of ADP6 with pADPW37	This study
ADPW77	*areA*::Km^r, *catA*$^-$; transformation of ADPW1 with pADPW37	This study
ADPW80	*areA*::Km^r, *benD*$^-$; transformation of ADPW38 with pADPW37	This study
ADPW86	*salE*::Km^r, *catA*$^-$; transformation of ADPW1 with pADPW76	This study
ADPW87	*salE*::Km^r, *benD*$^-$; transformation of ADPW38 with pADPW76	This study
ADPW88	*salE*::Km^r, *pcaG*$^-$; transformation of ADP6 with pADPW76	This study

9.2. Growth test analysis of mutants to confirm the position at which metabolites of aryl esters enter the β-ketoadipate pathway.

9.2.1. Phenotypes of single *are* mutants.

The chromosomal copies of *areA*, *areB*, and *areC* were disrupted by the insertion of a kanamycin resistant cassette, creating *Acinetobacter* strains ADPW58, ADPW68 and ADPW56, respectively (Table 9.1.1) (Chapter 3.3.) (Jones *et al.*1999). These mutants were tested for plate growth on benzyl acetate, 2-hydroxybenzyl acetate and 4-hydroxybenzyl acetate, benzyl alcohol, 2-hydroxybenzyl alcohol and 4-hydroxybenzyl alcohol and benzoate, 2-hydroxybenzoate (salicylate) and 4-hydroxybenzoate, as well as succinate as control (Table. 9.2.1).

With ADPW56 (*areC::*Km) and ADPW68 (*areB::*Km) the rates of growth on benzyl acetate and on the benzyl alcohol substrates were severely reduced, but not completely eliminated, and patches took from 2-3 days to achieve significant size compared with overnight growth for ADP1. ADPW56 and ADPW68 appeared to grow normally on benzoate. ADPW58 (*areA::*Km) grew normally on all substrates except for the benzyl acetates, where again growth was severely reduced and patches took 2 days to achieve comparable growth to the overnight growth of ADP1. Growth score for ADPW58 on benzyl acetate substrates is thus given a +/- 'poor growth' score since the growth of the mutants appeared quicker than growth of the mutants on benzyl alcohol substrates, which is adjudged as no growth. This is consistent with the linked biochemical assays to determine AreA activity in ADP1 (Chapter 6.3.3), where a considerable background esterase activity was detected in cell free extract from ADP1 cells grown on succinate. The results are consistent with disruption of a pathway involving sequential action of AreA, AreB and AreC, where benzyl acetate, and the 2-hydroxy- or 4-hydroxy-benzyl acetates are converted to their benzoate derivatives.

Table 9.2.1. Growth phenotypes of *Acinetobacter* strains with single mutations in one of the *are* genes.

	Growth of strain[a] (genotype)			
Growth Substrate	ADP1	ADPW58	ADPW68	ADPW56
	w.t	*areA*⁻	*areB*⁻	*areC*⁻
Benzoate	+	+	+	+
2-OH Benzoate	+	+	+	+
4OH Benzoate	+	+	+	+
Benzyl alcohol	+	+	-	-
2OH-Benzyl alcohol	+	+	-	-
4OH-Benzyl alcohol	+	+	-	-
Benzyl acetate	+	+/-	-	-
2OH-Benzyl acetate	+	+/-	-	-
4OH-Benzyl acetate	+	+/-	-	-
Succinate	+	+	+	+

[a] Growth assessed on plates: + Good growth; - no growth; +/- challenged growth

9.2.2. Phenotypes of single β-ketoadipate pathway mutants, and of double *are*/β-ketoadipate pathway mutants.

The Km cassette insertion in *areA* (pADPW37) was individually introduced into ADP6, ADPW1 and ADPW38 by natural transformation to produce three double mutants designated ADPW76, ADPW77, and ADPW80 respectively (Table. 9.1.1). The chromosomal insertion in *areA* for each mutant was analysed by Southern blot analysis (Data not shown). All mutants were tested for growth on the same carbon substrates as the single *are* mutants (Table 9.2.2.).

ADPW1 (*catA*$^-$) failed to grow on all the tested unsubstituted and 2-hydroxy-substituted aromatic substrates, accumulating catechol in the medium in all cases, indicating that catechol is an intermediate in the catabolism of unsubstituted and 2-hydroxylated aromatic substrates. ADPW38 (*ben*$^-$) failed to grow only on the unsubstituted aromatic substrates, and ADP6 (*pcaK*$^-$) failed to grow on only the 4-hydroxy substrates. The growth phenotypes are consistent with what is already known about the β-ketoadipate pathway. Benzoate is channelled to catechol by the enzymatic action of the *benABC* gene product, 2-hydroxybenzoate is channelled to catechol by the action of salicylate hydroxylase (Chapter 7.3), and 4-hydroxybenzoate is channelled to protocatechuate by the action of the *pob* gene products.

The phenotypes of the *areA*/β-ketoadipate double mutants grown on the substituted and unsubstituted benzoates and benzyl alcohols were identical to the phenotypes of the parent single β-ketoadipate pathway mutants (Table 9.2.2.). This is consistent with a mutation in *areA* not affecting growth on benzyl alcohol or any substrate below benzyl alcohol in the pathway.

However, ADPW77 (*catA*$^-$ *areA*$^-$) failed to grow on benzyl acetate or 2-hydroxybenzyl acetate and there was no accumulation of brown colour following 3 days incubation, unlike ADP1, showing that AreA was involved in the conversion of these substrates to catechol. Slight growth of ADPW77 appeared on 4-hydroxybenzyl acetate following three days incubation.

Table 9.2.2. Growth phenotypes of *Acinetobacter* strains with single or double mutations.

	Growth of strain[a] (genotype)				
Growth Substrate	ADPW38 *ben*⁻	ADPW1 *catA*⁻	ADP6 *pcaG*⁻	ADPW80 *ben*⁻/*areA*	ADPW77 *catA*⁻/*areA*	ADPW76 *pcaG*⁻/*areA*
Benzoate	-	-B	+	-	-B	+
2-OH Benzoate	+	-B	+	+	-B	+
4-OH Benzoate	+	+	-	+	+	-
Benzyl alcohol	-	-B	+	-	-B	+
2OH- Benzyl alcohol	+	-B	+	+	-B	+
4OH- Benzyl alcohol	+	+	-	+	+	-
Benzyl acetate	+/-	-B	+	+/-	-	-
20H- Benzyl acetate	+	-B	+	+/-	-	+/-
4OH- Benzyl acetate	+	+	-	+/-	+/-	+/-
Succinate	+	+	+	+	+	+

[a] Growth assessed on plates: + Good growth; - no growth; +/- challenged growth -B no growth with accumulation of brown/black colouration (catechol)

ADPW80 (*ben*⁻ *areA*⁻) failed to grow on benzyl acetate, but slight growth following two days incubation appeared when it was patched on both 4-hydroxy- and 2-hydroxy- benzyl acetates. ADPW76 (*pcaG*⁻ *areA*⁻) failed to grow on only 4-hydroxybenzyl acetate, while slight growth appeared when patched on the other two ester substrates following two days incubation.

The growth phenotypes of the double mutants were consistent with the proposed AreABC pathway where alkyl benzoates are channelled into the β-ketoadipate pathway. The unsubstituted benzoate nucleus from the benzyl acetate is fed into the benzoate branch to catechol via the action of the *ben* gene products, and growth is inhibited when there is a mutation in either the *ben* genes, in *catA* or in *areA*. The 2-hydroxy benzyl esters are fed into the β-ketoadipate pathway at the level of catechol and growth is inhibited only when there is a mutation in *areA* or in *catA*. In addition, the 4-hydroxybenzyl acetate is fed into the β-ketoadipate pathway via 4-hydroxybenzoate, and growth is inhibited only when *areA* and/or *pcaG* are disrupted.

9.3. Growth test analysis of mutants to confirm the position at which the product of salicylate ester catabolism enters the β-ketoadipate pathway.

9.3.1. Phenotypes of single *sal* mutants.

The chromosomal copies of *salA*, *salR*, and *salE* were disrupted by the insertion of a Km resistant cassette, creating *Acinetobacter* strains ADPW67 (*salA*::Km), ADPW72 (*salR*::Km) and ADPW70 (*salE*::Km), respectively (Table 9.1.1) (Chapter 7) (Jones *et al.* 2000). These mutants with a single mutation in a *sal* gene were tested for growth on ethyl salicylate, benzoate, 2-hydroxybenzoate, 4-hydroxybenzoate, and on succinate (Table 9.3.1). ADPW67 (*salA*⁻), ADPW72 (*salR*⁻) and ADPW70 (*salE*⁻) all failed to grow on both salicylate or ethyl salicylate unlike ADP1 which grows vigorously on both substrates overnight. Growth of these three mutant strains on benzoate or 4-hydroxybenzoate was unaffected.

Table 9.3.1. Growth phenotypes of *Acinetobacter* strains with single mutations in one of the *are* genes.

	Growth of strain[a] (genotype)			
Growth substrate	ADP1	ADPW67	ADPW72	ADPW70
	w.t	*salA*$^-$	*salR*$^-$	*salE*$^-$
Benzoate	+	+	+	+
2-OH Benzoate	+	-	-	-
4OH Benzoate	+	+	+	+
Ethyl 2-OH Benzoate	+	-	-	-
Succinate	+	+	+	+

[a] Growth assessed on plates: + Good growth; - no growth;

9.3.2. Phenotypes of single β-ketoadipate pathway mutants, and of double *sal*/β-ketoadipate pathway mutants.

In order to demonstrate that the aromatic moiety arising from salicylate and ethyl salicylate is channelled down the β-ketoadipate pathway, the phenotypes of mutants blocked both in the *sal* genes and in three different points in the β-ketoadipate pathway were tested (Table 9.3.2.). Double mutants were created by the transformation of pADPW76 (*salE*::Km) into ADPW1, ADP6 and ADPW38 respectively to create ADPW86 (*catA*$^-$ *salE*$^-$) ADPW87 (*ben*$^-$ *salE*$^-$), and ADPW88 (*catA*$^-$ *salE*$^-$) (Table 9.1.1.). Each double mutant strain was analysed by Southern hybridisation and the chromosomal insertion confirmed (data not shown).

Of the single β-ketoadipate pathway mutants only ADPW1 (*catA*$^-$) failed to grow on salicylate or ethyl salicylate, accumulating catechol in the medium from both substrates. The growth phenotypes of the these mutants were consistent with the proposed pathway (Fig.8.5.1.) where the salicylate nucleus is fed into the benzoate branch of the β-ketoadipate pathway at the level of catechol.

The phenotypes of the double mutants when grown on 4-hydroxy, 2-hydroxy, and unsubstituted benzoates are identical to their parent single mutants. However, none of the double mutants are able to grow on ethyl salicylate. Moreover, no accumulation of catechol was observed when ADPW86 (*catA*$^-$ *salE*$^-$) was grown on ethyl salicylate indicating that disruption of *salE*$^-$ renders the mutant unable to hydrolyse ethyl salicylate. The ability of *salE*$^-$ mutants to grow on ethyl benzoate and ethyl 4-hydroxybenzoate was not tested. Studies should be undertaken to determine the growth phenotype of this mutant against these two substrates.

9.4. Discussion.

9.4.1. Pathway feeding aryl esters and salicylate esters into the β-ketoadipate pathway.

Insertional knock-outs of *are* and *sal* genes, and the *areA*$^-$/β-ketoadipate pathway and the *salE*$^-$/β-ketoadipate pathway double mutants, show the phenotype expected from their

Table 9.3.2. Growth phenotypes of *Acinetobacter* strains with single or double mutations.

	Growth of strain[a] (genotype)					
Growth substrate	ADPW38 *ben*$^-$	ADPW1 *catA*$^-$	ADP6 *pcaG*$^-$	ADPW88 *ben*$^-$/*salE*$^-$	ADPW86 *catA*$^-$/*salE*$^-$	ADPW87 *pcaG*$^-$/*salE*$^-$
Benzoate	-	-B	+	-	-B	+
2-hydroxy benzoate	+	-B	+	+	-B	+
4-hydroxy benzoate	+	+	-	+	+	-
Ethyl 2-hydroxy benzoate	+	-B	+	-	-	-
Succinate	+	+	+	+	+	+

[a] Growth assessed on plates: + Good growth; - no growth; -B no growth with accumulation of brown/black colouration (catechol)

respective proposed pathways. When a particular step in the pathway was blocked by insertion or mutation, the resulting mutant was unable to grow on any substrate that precedes it in the pathway. When the step of the pathway that converts catechol to *cis, cis,* muconate was blocked, then any compound that has an unblocked route into catechol, accumulated catechol, which was seen as brown colouration on the agarose plate, and was evidence for the production of catechol.

Acinetobacter sp. ADP1 can utilise unsubstituted or hydroxy-substituted aryl esters or salicylate esters as sole carbon sources. The hydrolysis of these esters forms benzoate or 2-hydroxybenzoate or 4-hydroxybenzoate which are then fed into the β-ketoadipate pathway, and eventually converted into central metabolism intermediates (Fig. 9.4.1.).

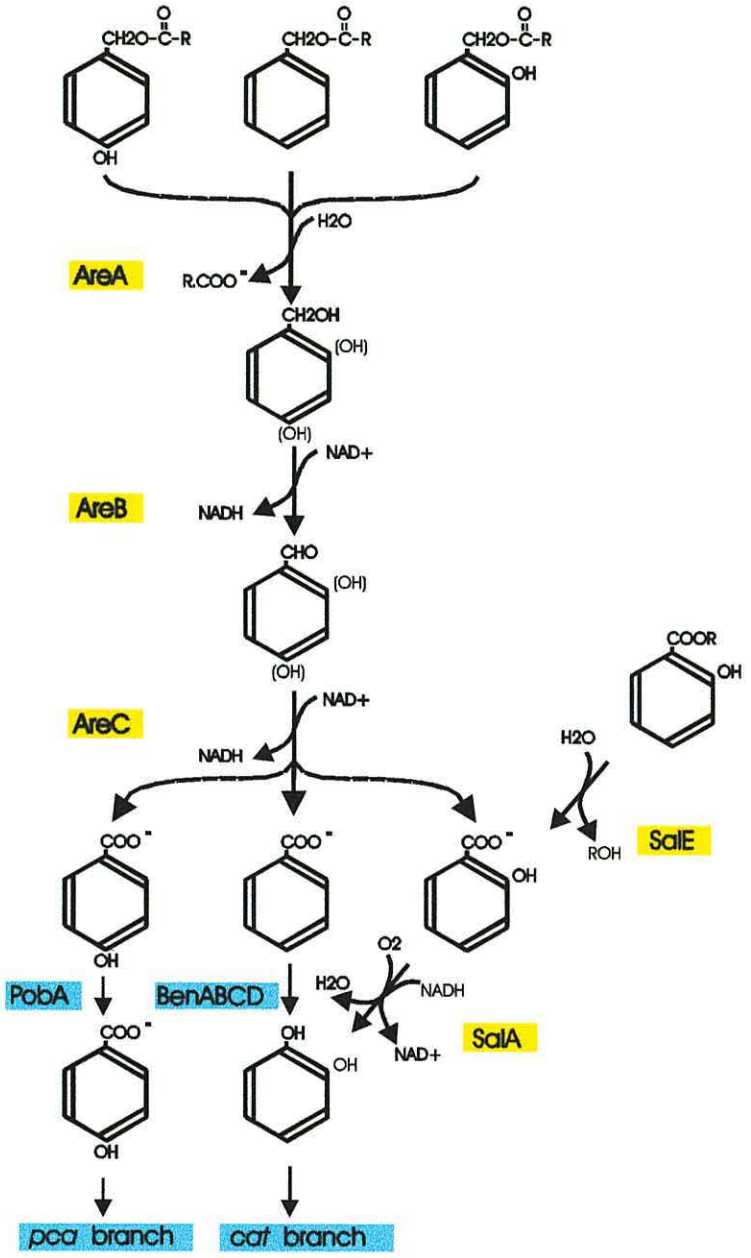


Fig. 9.4.1. The catabolism of aromatic esters through the β-ketoadipate pathway.

The previously described genes encoding the enzymes of the β-ketoadipate pathway are highlighted in blue (Harwood & Parales, 1996). The genes cloned and characterised in this thesis are highlighted in yellow (Jones *et al.*1999; Jones *et al.* 2000).

CHAPTER 10

Discussion

1.10.1. The catabolic role of the *are* genes.

The *areABC* genes of ADP1 that were characterised in these studies encode catabolic enzymes for the sequential conversions of the esters of benzyl alcohol to benzyl alcohol, benzaldehyde and benzoic acid. The *areABC* gene products are thus responsible for the funnelling of metabolites into the catechol branch of the β-ketoadipate pathway, which is encoded by the chromosomally adjacent *ben* and *cat* genes.

Insertional inactivation of each *are* gene on the *Acinetobacter* chromosome did not completely eliminate ADP1's ability to utilise the *are* operon substrates as carbon growth substrates. An explanation for this is that it is possible that similar enzymes within the cell can partially complement the defective genes. Nevertheless, the severe impairment of the catabolism of these substrates by the *are* mutants was clearly demonstrated by the ability of the wild-type strain to out-compete each mutant during growth on benzyl acetate.

Each of the AreABC gene products was over-expressed individually and their substrate profiles were determined. Enzyme assays confirmed the role of the AreABC proteins. In addition the assays also showed that they are capable of the conversion of esters of 2-hydroxybenzyl alcohol and 4-hydroxybenzyl alcohol to the corresponding acids. The 2-hydroxybenzyl (salicyl) alcohol esters are channelled to salicylate and the β-ketoadipate pathway via the adjacent *salA* gene. The 4-hydroxybenzyl alcohol is catabolised to 4-hydroxybenzoate, which funnels into the protocatechuate branch of the β-ketoadipate pathway that is encoded by genes 280kbp from the *are* genes on the ADP1 chromosome (Gralton *et al.* 1997). In addition, it was shown that growth of ADP1 on 4-hydroxybenzyl alcohol induces both benzaldehyde and benzyl alcohol dehydrogenase activities with the same substrate preference as when grown upon benzyl alcohol or benzyl acetate. This indicates that the same genes are expressed and are functional during growth on each substrate. Also, the chromosomal disruption of *areA, B,* or *C* severely reduces the ability of the mutant to grow on 2-hydroxybenzyl

acetate or 4-hydroxybenzyl acetate, which are both growth substrates for ADP1. This indicates that it is likely that these two substrates are also catabolised by the *are* genes.

The AreA protein may be able to degrade a much wider range of aryl esters than those that were tested in this study. This is probable because of the wide range of aromatic ester compounds in the environment, and the wide range of natural aromatic acids which ADP1 can utilise. Most commercially available compounds hypothesised as AreA substrates were assayed. However, most aromatic esters, which are a product of plant degradation, are not currently commercially available, and thus could not be assayed against AreA in this study.

The AreA protein is non-specific as far as the acid moiety of the aromatic ester is concerned and is able to hydrolyse 4-nitrophenyl esters of alkanoic acids with chain lengths of up to 16 carbons (Jones *et al.* 2000). Also, it is possible that esterase activity on a novel substrate might produce an aromatic alcohol that is not further metabolised. In the natural environment, the Are enzymes may therefore be responsible for supporting growth on a wide range of substrates.

However factors other than just substrate specificity also contribute to determining which carbon sources ADP1 can utilise as growth substrates. In these studies, benzyl benzoate was shown to act as a substrate for AreA to produce two products, benzoic acid and benzyl alcohol, both of which are growth substrates of ADP1: it would therefore appear to be an ideal growth substrate. However, tests also showed that benzyl benzoate is not a growth substrate of ADP1. This is possibly because of toxicity or blocks in uptake or regulation.

1.10.2. Genetic organisation and induction of the *areCBA* operon.

Sequence analysis found that the intergenic regions between *areC* and *B* (27 bp) and between *areB* and *A* (83 bp) are relatively short, suggesting coexpression of *areCBA* as an operon. Analysis of the *areCBA* genes showed that they are all cotranscribed on the same mRNA. Three lines of experimental analysis confirmed this:

1) RT-PCR analysis with primers whose products span the intergenic regions between *areC* and *areB*, and between *areB* and *areA*.

2) Increased AreC, AreB, and AreA activities were measured from ADP1 cells grown on benzaldehyde, benzyl alcohol, or benzyl acetate, but not from ADP1 cells grown on succinate. This showed that the substrates further down the *are* pathway could induce the enzymes at the beginning of the pathway. Benzyl alcohol dehydrogenase (AreB), and benzyl esterase (AreA) activity was found when the cell was induced by benzaldehyde; and benzyl esterase (AreA) activity was found when the ADP1 was induced by benzyl alcohol. This suggests a common regulatory mechanism that recognises all three substrates and causes co-expression of all three enzymes, rather than a series of sequential inductions.

3) Increased β-galactosidase activity was measured from the *are-lacZ* chromosomal fusions in ADPW61 (*areA*::*lacZ*), and in ADPW63 (*areC*::*lacZ*) when induced by benzaldehyde, benzyl alcohol, or by benzyl acetate. This shows that benzaldehyde, the substrate of the first gene in the operon, induces *areA*, which is the last gene in the operon.

Further studies using the *lacZ* reporter gene fused with the *are* genes showed increased β-galactosidase activity in ADPW61 and ADPW63 when induced by a range of ester substrates. The inducibility of both genes was highest with benzyl acetate, and with 2- or 4-hydroxybenzyl acetate, which all induced the *are* genes to similar levels of activity. The level of induction decreased as the length of the aliphatic moiety of the benzyl ester increased, with benzyl butyrate being the poorest inducer tested. Increased β-galactosidase activity was also measured from benzyl benzoate-induced cells. It is noteworthy that the level of induction by benzyl benzoate is higher than that from benzyl butyrate, but lower than that from benzyl propionate. These results exactly mirror the data from the study of the kinetic parameters for AreA against benzyl esters, where the Michaelis constant (Km) of AreA against benzyl benzoate was higher than that against benzyl butyrate, but lower than that against benzyl propionate. These results also show that regulation is not the reason why benzyl benzoate is not a growth substrate of ADP1. In addition, for the cell to achieve induction, the benzyl benzoate must have been transported across the

membrane. The growth tests to determine whether ADP1 could grow on benzyl esters were carried out by plating an ADP1 culture on a minimal medium plate, and then aliquoting the volatile esters in small caps on the lid of the plate. While benzyl propionate and benzyl butyrte readily support growth of ADP1, it is possible that the concentration of benzyl benzoate proved toxic to ADP1 in those conditions.

1.10.3. The *areCBA* genes are probably regulated by AreR.

The position of the transcriptional start of the *areCBA* operon was determined by Primer Extension analysis and was found to be located 98 bp from the *areC* start codon. Upstream of the transcriptional start is a σ^{54}-dependent -12, -24, promoter sequence (T**GG**CAC-N_5-CT**GC**) that is a nearly exact match to the consensus sequence T**GG**CAC-N_5-TT**GC** (Shingler, 1996). This indicated that a member of the σ^{54}-dependent NtrC transcriptional regulator family controls the expression of AreCBA. This hypothesis was further strengthened by the identification of an inverted repeat sequence upstream of the –24 -12, region, whose arms contain direct repeat sequences. This type of sequence is characteristic of an upstream activator site (UAS) where the NtrC type transcriptional regulator binds to the DNA (Shingler, 1996).

Upstream of the inverted repeat is the gene *areR*, which has all the amino acid sequence characteristics of a σ^{54}-dependent NtrC type transcriptional regulator. An insertional knockout disruption in the C-domain of *areR* resulted in a mutant that was unable to induce AreC, AreB, or AreA activity when grown in the presence of benzyl alcohol. These results indicate that AreR is the regulator involved in AreABC expression. These results clearly need to be backed up by data from Nuclease Protection assays and from Gel Mobility Shift assays showing that AreR indeed binds to the *areC* promoter DNA.

The mode by which the AreR's A-domain may be activated is uncertain. The two proteins, XylR and DmpR, which are members of the NtrC transcriptional regulator family, both induce aromatic catabolic operons where the pathway substrate binds to the A-domain of the protein to relieve suppression of ATPase activity (Delgado & Ramos, 1994; O'Neill *et al.* 1998). The DmpR and XylR A-domains share 73% amino acid identity, but neither has significant similarities to the A-domain

of AreR. Whereas the A-domains of the regulators that share similar effector substrates to AreR do not have sequence conservation, the genes which are arranged in a similar operonic structure as the *are* genes, do have a similar sized, and a reasonably conserved A-domain. The gene that was found to have the highest sequence homology to the A-domain of AreR was AcoR from *Clostridium magnum* (Kruger *et al.* 1994). AcoR regulates the genes encoding the enzymes for acetoin dehydrogenase in *C. magnum*. Whereas the *dmpR* and *xylR* are divergently transcribed from the genes that they regulate, both *areR* and *acoR* are transcribed in the same direction, but AreR and AcoR appear to be induced by structurally different substrates, i.e. alkyl esters and acetoin (3-hydroxy-2-butanone) respectively.

1.10.4. Comparison of *areBC* with the corresponding genes from *A. calcoaceticus* NCIB 8250.

The *xylBC* genes of *A. calcoaceticus* NCIB8250 are highly homologous to *areBC* of ADP1. The two *A. calcoaceticus* genes are in the same order as in ADP1, although no *areA* homologue has been identified downstream of these *xyl* genes. This was determined by sequence analysis of the published NCIB8250 *xylCB* sequence. There are 200 bp downstream of *xylB* which show negligible homology to the DNA downstream of *areB*. Furthermore, Southern Blot analysis was performed on a plasmid containing 1-2 kbp downstream of *xylBC* (a gift from Gillooly & Fewson, University of Glasgow, UK) with an *areA* probe. The *areA* probe did not hybridise to the plasmids which suggests that in the NCIB8250 strain, the *xylCB* genes may be involved in catabolism of only benzyl alcohol and not of benzyl esters.

1.10.5. The catabolic role of the *sal* genes.

The two enzymes salicylate esterase (SalE) and salicylate hydroxylase (SalA), characterised in this study, encode the sequential catabolism of ethyl salicylate to salicylate and of salicylate to catechol. These two genes form another route by which aromatic ester substrates can funnel into the catechol branch of the β-ketoadipate pathway. The pathway was deduced by expressing both enzymes from cloned genes, SalE from the expression vector pET5a where it is expressed to very high specific

activity, and SalA from a pUC18 clone carrying *salERA*. The cloned SalA salicylate hydroxylase activity showed the same relative substrate preferences as the activity found in wild type ADP1 grown on salicylate itself. Ethyl salicylate esterase (SalE) was only assayed against ethyl salicylate.

Insertional knock-outs of both *salE* and *salA* show the phenotype expected from the proposed pathway. In addition, the further catabolism of the aromatic moiety into the *ben-cat* branch of the β-ketoadipate pathway at the level of catechol was also confirmed by the fact that only mutation in *catA* (for catechol 1,2-dioxygenase) and not in the *ben* genes or in the *pca* genes destroyed the ability to utilise either ethyl salicylate or salicylate as growth substrates.

1.10.6. Regulation of *sal* genes.

An unsuccessful attempt was made to determine the position of the start of transcription of *salA* by Primer Extension probably because there was an insufficient quantity of DNA in the assay. Comparable total RNA quantities to those successfully used when determining the *are* operon transcriptional start proved insufficient. Nevertheless, sequence comparisons proved successful in locating a conserved σ^{70}-type promoter element upstream of *salA*. A recognised conserved extended -10 sequence (TGATATAAT) was located at a position 130 bps upstream to *salA*, accompanied by a region that could serve as a -35 (AGGCAA). Although the -35 region is not homologous to the paradigm sequence (TTGACA), the highly conserved extended -10 sequence has been shown to compensate for unconserved -35 regions (Dombroski, 1997; Sabelnikov *et al.* 1995). In addition to sequence analysis, biochemical assays determined that ADP1 has salicylate-inducible hydroxylase activity. The accumulated data firmly suggests that the transcription of *salA* gene is regulated by a σ^{70}-type regulator protein.

RT-PCR studies showed that *salA* is co-transcribed with *salR*. Analysis of the SalR amino acid sequence shows it is clearly a member of the LysR family of regulator proteins. SalR's closest relatives as determined by sequence alignments are two proteins designated NahR, which both regulate salicylate hydroxylases in *Pseudomonas*. A knockout mutant of *salR* (ADPW72) was shown to be unable to

grow on salicylate as a sole carbon source, and also stops induction of salicylate hydroxylase activity by salicylate. Also, salicylate hydroxylase is also salicylate-inducible in *E.coli* containing pADPW34 that carries only *salARE*. The accumulated data indicates that SalR is the protein that regulates expression of SalA. In addition, because *salA* and *salR* are co-transcribed, then it is implied that SalR thus controls its own expression.

Sequence analysis for a Regulator Binding Sequence (RBS) upstream of the putative –10,-35 promoter of *salA* revealed sequence similarities to regions that have been identified (Schell, 1993) as being the RBS at which NahR interacts with the DNA. Schell identified the sequence TCA-N_3-TGATGA upstream of the *nahG* (salicylate hydroxylase) in NAH7 of *Pseudomonas* as being the NahR binding site. DNA sequence alignments of the ADP1 and *Pseudomonas* DNA over this region showed that ADP1 has the sequence TCA-N_3-TGATGG. This conserved sequence is also present upstream of the *nahG* of *Pseudomonas stutzeri*, and upstream of orf7 in the unstable vanillate cluster of genes in ADP1, which has relatively high homology to *salA*. The promoter region upstream of *salA* needs to be analysed by Nuclease Protection assays and by Gel Mobility Shift assays to show that SalR binds to the hypothesised RBS.

A major difference between the *sal* and *nah* genes is that, whereas *nahR* is transcribed divergently from its adjacent catabolic gene *nahG*, *salA* and *salR* are co-transcribed as a single regulatory unit as shown by the RT-PCR results.

1.10.7. Ester hydrolases in *Acinetobacter* sp. ADP1.

The characterisation of the two cytoplasmic esterases, AreA and SalE, adds to the list of characterised ester hydrolases in ADP1. There are now five well characterised ester hydrolases (*estA, estB, lipA, areA,* and *salE*), and at least three other genes which show similarities to ester hydrolases (*estC, estD,* and *ppaA*).

An insertional inactivation of *salE* completely destroyed the ability of the mutant to grow on ethyl salicylate, while an insertional inactivation of *areA* did not completely destroy the ability of the mutant to grow on benzyl acetate. As stated above, out of the limited ADP1 DNA sequence cloned, several esterases have been

identified. There is no clear understanding of the actual number of ester hydrolases that are actually present in ADP1. It is possible that other esterases within the cell can partially complement the defective *areA* gene. These considerations also lead to the question as to why a *salE* disruption did not produce a mutant with partial reduction in the cell's ability to grow on ethyl salicylate. The reason why ADP1's array of ester hydrolases did not compensate for SalE activity must be that SalE has a much tighter functional specificity than AreA. It has already been determined that AreA can hydrolyse 4-nitrophenyl esters with an aliphatic moiety of up to 16 carbon atoms long, while SalE could not hydrolyse 4-nitrophenyl esters with an aliphatic moiety of longer than 6 carbons (Jones *et al.* 2000). The specificity of SalE has yet to be studied with respect to benzoate esters other than ethyl salicylate. It would be a useful study to investigate the ability of SalE to catabolise ethyl benzoate and ethyl 4-hydroxybenzoate such that a clear understanding of the metabolic capabilities of SalE can be appreciated.

1.10.8. The contributions of AreA and SalE to ADP1 metabolism.

The AreA and SalE proteins both function in channelling aromatic esters into the β-ketoadipate pathway. Each enzyme is specific to the type of aromatic ester which it hydrolyses. AreA has been shown to be specific to aromatic esters where the aromatic moiety of the molecule is the alcohol, and the aliphatic moiety is the acid. On the other hand, SalE has been shown to be specific for aromatic esters where the aromatic moiety is the acid and the aliphatic moiety is the alcohol. These two genes contribute to funnelling aryl esters and alkyl salicylate into central metabolites, which ultimately contributes to the metabolic diversity of ADP1. This metabolic diversity enables ADP1 to grow on substrates that are likely to be found as natural products either directly of plant origin or as microbial breakdown products of plant compounds.

1.10.9. Gene organisation of *areA*, *salD* and *salE*.

The *areA* and *salE* genes are within 1.5 kbp distance of each other on the ADP1 chromosome and both are transcribed in the same direction (fig. 7.2.1.). The two genes are separated by *salD*, an open reading frame with no designated function, but

which has similarities to genes coding for proteins that have been tentatively assigned functions as membrane proteins involved in the transport of hydrocarbons. These include TodX (Wang *et al.* 1995), XylN (Harayama, S. Genbank accession no. D63341.) and TbuX (Kahng *et al.* 2000). These studies have shown that *areABC* genes are transcribed on the same mRNA transcript, and that the *salDE* genes are transcribed on the same mRNA transcript.

It is hypothesised that the inverted repeat at the end of *areA* might be a transcriptional terminator, but as yet there is no conclusive evidence for this role. Further studies must be carried out using Northern Blotting techniques to determine the actual size of the *areCBA* transcript. Reverse Transcriptase PCR (RT-PCR) analysis must also be done using primers whose products span the *areA-salD* intergenic region. A negative RT-PCR result would indicate that the inverted repeat at the end of *areA* indeed functions as a terminator loop. If the RT-PCR result is negative, then studies could be done to determine the transcriptional start of the *salDE* mRNA transcript. Sequence gazing at the *salD* upstream region has met with limited success. Conserved -10 promoter consensus sequences have been located at 1) 120 bps upstream of the *salD* start codon which has the sequence TATAAA; and 2) 240 bps upstream of the *salD* start codon with the sequence TATAAT. Both are similar to the recognised conserved -10 sequence TATAAT (deHaseth *et al.*1998). However, no recognised –35 sequence could be found adjacent to the –10 promoter sequence.

If the conserved –10 sequence at 240 bps upstream of the *salA* start codon proves to be the promoter region serving *salDE* transcription, then this has implications regarding the regulation of expression of both *areA* and *salE*. This is because the intergenic region between *areA* and *salD* is only 195 bps long. It has been shown by RT-PCR analysis that the expression of *salDE* is inducible (see chapter 8.4.). Elements regulating *salDE* expression therefore would have to bind to the promoter that must be located more than 240 bps upstream of the *salA* start codon, which is internal to the *areA* open reading frame. Any binding of proteins to the *areA* open reading frame would disrupt efficient transcription of *areA*.

Since AreA and SalE are specific to the type of aromatic ester that they hydrolyse, it would appear inefficient for the cell to express both enzymes when only in the presence of one type of aromatic ester substrate. Further analysis must be done

to determine whether there is SalE activity in benzyl acetate grown ADP1 cells, and whether there is AreA activity in ethyl salicylate grown cells. In addition, studies could be done to determine the substrate preference of ADP1 by analysing the growth media when ADP1 is grown in the presence of benzyl acetate and ethyl salicylate concurrently.

1.10.10. A possible role for SalD.

SalD shows homology to proteins which include TodX (Wang *et al.* 1995), and XylN (Harayama, Genbank accession no. D63341) and which are putatively thought to function as membrane proteins involved in the transport of hydrocarbons. The result which indicated that SalD has a membrane related function was seen when ADPW78 (*salD*::Km^r) was found to be a very unhealthy strain, growing slowly on minimal medium and losing viability when maintained on agar after 2 to 3 days. This phenotype is characteristic of the knockout of a gene with a membrane-related function. In addition, when a Km^r-*lacZ* cassette from pKOK6.1 was inserted into the *salD* gene, the mutant was very unhealthy and lost viability quickly. No β-galactosidase activity could be measured from this fusion either as induced by ethyl salicylate or from background succinate induced fusions (Weiske Pool, unpublished results).

As well as having a membrane function, analysis of the amino acid sequence of SalD suggests that it may also have hydrolytic activity. SalD contains two stretches of sequence that resemble the deduced amino acid sequence at the serine active site of hydrolytic enzymes according to PROSITE (Bairoch *et al.*1997), which show the signature sequence of [LIV]-x-[LIVFY]-[LIVMST]-G-[HYWV]-S-x-G-[GSTAC], where residues at a point can be any of the ones stated in the parentheses. SalD contains the sequence AYGV**GASIGS** at residues 195 to 204, and the sequence LNGT**GNSTGT** at residues 360 to 369. Although they are not exactly homologous to the whole signature sequence, they do contain the invariable consensus sequence Gly-X_1-Ser-X_2-Gly (highlighted in boldface) (Brady *et al.* 1990). Studies should be undertaken to determine whether SalD has membrane bound ester hydrolase activity. This activity would possibly be against aromatic esters with longer aliphatic moieties

since it has already been determined that SalE can only hydrolyse p-nitrophenyl esters of up to six carbon atoms in length.

1.10.11. Expanding the perimeter of the *ben-cat* supraoperonic cluster.

The cloning of the 10kbp fragment which includes *areRABC* and *salDERA* presented in this thesis adds to the characterisation of genes involved in the catabolism of aromatic compounds in *Acinetobacter* sp. ADP1. The cloning of these genes has also extended the shores of the *ben-cat* chromosomal island of catabolic genes from the previously characterised 20kbp to the now characterised 30kbp of catabolic, regulatory, and transport genes.

In addition, a further 5kbp of DNA upstream of *salA* has been cloned and sequenced (GenBank accession no. AF150928). Sequence similarity searches of this DNA revealed no homologies to genes involved with aromatic catabolism. Three complete open reading frames were identified, which have homologies to two-component regulatory systems. The function of these genes is yet to be determined, but due to their proximity to the *sal-are-ben-cat* cluster, the possibility remains that they could be involved in the regulation of aromatic catabolising enzymes.

The full extent of this genetic island of catabolic diversity is still unknown. The present knowledge lies within *sal* on one side and within *cat* on the other. If the genome sequence of ADP1 was available, the perimeter of the island of clustered catabolic genes might be identified. But in fact, the limiting amount of sequence data available might have also been helpful in focusing upon cloned genes for which functions have been readily defined. Extending the perimeter of the catabolic island should be straightforward and demands only the identification of DNA extending beyond a marker which can be selected.

It is premature to anticipate that extending the perimeter of the catabolic island will discover any additional genes with catabolic functions. But experience from this study shows that sequence comparisons of newly cloned genes will indicate a catabolic activity that will be a guide to tests with an appropriate growth substrate. Identification of the *salE* gene function is a good example where sequence comparisons suggested similarities to esterases, and the relative location of *salE* amid

aromatic-catabolising genes suggested an aromatic catabolic function. Other sequences at the limits of the island may give indications of horizontal transfer between species as a factor in acquisition of the cluster of aromatic-catabolising genes.

Appendix A. The nucleotide sequence from *salA* to *benM*.

```
promoter        186..152
                /note="putative; for salAR"
misc feature    110..130
                /note="SalR regulator protein; putative"
                /note="binding site"

gene            gene(317..1587)
                /gene="salA"
                /function="salicylate hydroxylase"
                /product="SalA"
                /translation="VGKKISIAIIGGGIAGVALAANLFKQPHLEVCLYEAAPQFSEIG
                AGISFGANAVRAIELLGLASQYTAIADQVSAPFQDVWFQWRNGYNDDYLSSSISPQVG
                QSSVHRADFLDAILGNIPQHQCKFNKKLKSIQEYDTHIELSFEDGTCAEADYVIGADG
                IHSATRDYVLQTHQFAPVRPNFTGTWAYRGIIKAAEFRQAIVAAGLDVEIADVPQMFL
                GQNKHILTFPIRQGEDINIVAFKTNPEQRTLPEHTPWTRAVDKQEMLDDFQDWSESCR
                ILLGLIERPTLWALHELAELPTYQSHSGRVILMGDAAHAMLPHQGAGAGQGLEDALTL
                KVLFEHTELTVEDLPRVSAIYEQIRKERACKVQRTSRESGQIYELNSALYPSFEAVGA
                HLQNRFDWLWQHDLAQDMLAARAAIQPVATI"

gene            1712..2602
                /gene="salR"
                /function="regulator protein"
                /product="SalR"
                /translation="VDLSLIRIFICVYENKNISKAAEILNLSQPSVTYNLNRLRKHLN
                NPLFERTQYGVEATKLSHELYPVFKESILKIEIAVDEALNFNPLTSNKTFRIGLSDIG
                EICLLPTLIEYLRAHAPKIKIEVEEIKIDQVEKWLIEGFIDVAVFNSTHLEFKHLEYE
                TLFLERYVALVNMNHPRIRSTLSFDAYLNESHVAIKSSTGHTQVDHVLKLMGHQRKIA
                LEVPHFGVLQGVLDKTDLMVTLPSRAAQQYLNQSHVRVLELPFQMSEFYVGLHWFAQT
                NEPLARIWLIQTCKKVISVL"

gene            complement (2634..3368)
                /gene="salE"
                /function="alkyl salicylate esterase"
                /product="SalE"
                /translation="MITYVLVHGSWHDGSLWEPVATHLRAQGHTVHCPTVAGHGPNAD
                RNVTHAQCSQSIADYIVKHDLSEIVLLGHSYGGTIISKVAEAIPERIQRLIYWNAFVL
                QDGENMFDNMPEAYYELFTSLAAASGDNTVLLPYEVWRHAFINDADDQMAEETYKMLT
                PEPCQPFHDRLDLKKFYTLNIPKSYLNCTEDQALPAGFWHPKMSNRLGEFKLVEMGGS
                HEAMFTRPQELATKIIEASHD"

gene            complement (3384..4592)
                /gene="salD"
                /function="putative facilitator of uptake"
                /product="SalD"
                /translation="MFPHKLSLVLIACMPLSSYAFNGVYNTGSGQVSSGMGGVSLAQG
                MDRTSIQDNPANLSFQNSGADAQLSLLNIRAKATLLNQPEGFHSDKVIPIPSLALVQH
                YNDRLSWGLSLTGAGAAVDYDEPAIHGFPSDKAKDNLAIVKASPTVSYKVLPNLSLGA
                SLDLGIQQFRAKGVLAGVDGQGTPIFLESHGNQWAYGVGASIGSTWEFQPNWWLGASY
                STEMKFSKLDDYKNDLLARSEGRINLPERYGIGIKHKINQKLTLAADIVRFNWQDADG
                LGKKGSFNWQNQTVYRAGVDYQLTPHDHLRFGYSHADSVVDAQDTLVNFYANAVANQS
                WTIGYGHDLGFATLNLSYEYAVNKTLNGTGNSTGTNLSNQNHVITLGISKDF"
```

```
gene            complement (4741..5727)
                /gene="areA"
                /function="benzyl esterase"
                /product="AreA"
                /translation="MLQSLDPQLASFFAMLENQGAPAFDIGTLNDPNINEMRNWYLNF
                GPHVAGQFIELQDVEDLWIQVDDGSLRLRLYKTQRQHSESCETLIYFHGGGWCIGSIE
                THDLLCRHLVNDLDVNVISVDYRLAPEYPYPQAAHDAIHATQWIFEHASALGIHPDHL
                YVAGDSAGGNLAAITALTFKEQTYRTSLAGQILIYPCIDLTVIHDNKYPSFEKQKLTP
                PLSAEQMSKMIQIYLKNEDFAYQWLASPLLAQSHQHLPSTLIITAEFDPLHDQGCVYF
                EKLKSSLVNTEYHEIKGVVHGFFEMYGVLNKTKETINIIRAWLKNNTKSK"

gene            complement (5805..6920)
                /gene="areB"
                /function="benzyl alcohol dehydrogenase"
                /product="AreB"
                /translation="MTKFTEITAAVTECKGADFELKQLKIREPQNDEVLVKIVATGMC
                HTDLIVRDQYYPVPLPAVLGHEGSGIIEAIGPNVKDLAVGDHVVLSYGFCGSCEQCSS
                GQPAYCQDFFGRNFSGGDPDGHSAVCSHDHHKIHDHFFAQSSFATYAISRENNAVKVS
                KDVPLELLGPLGCGIQTGAGAAINALKVAPASSFVTWGAGAVGLSALLAAKVCGATTI
                IAVDIVESRLALAKELGATHVINSKTQDPVEAIKEITGGGVKFALESTGRPEILKQGI
                DALGILGSIAVVGAPPLGTTAAFDVNDLLLGGKSIIGVVEGSGVPKKFIPALVSLYQQ
                GKFPFDKLVKFYDFKDINQAAIDSHKGITLKPILKIG"

gene            complement (6958..8504)
                /gene="areC"
                /function="benzaldehyde dehydrogenase"
                /product="AreC"
                /translation="MTLLDASIWNKKLFNGGWFESGQPYGVVEVATGEQLGQTGSASP
                TDVAQAAKEAQTAQRQWWALDYLERQAVFEKAVQIATEHQDELVNWLIRESGSLALKA
                QFEVKVSIQMLKNCIAFPQLDQGNILPSRNGKLSLAKRLPLGVVGVISPFNFPLYLAL
                RAVGPALAFGNAVVLKPDERTAVCSGYAIARIFELAGLPKGLLHVLPGGADTGEALTL
                DPHIASIQFTGSTQVGRIIGANAGKTLKKVSLELGGKNSLIILDDADLDLAAENIAWG
                AFLHSGQICMTSGKILIHEKIYDALKARVIEKVKRFVVGNPLEQNVTIGPLINEKQSK
                RVEALVQAAVEHGATLEIGGKANGPFFEPSVLSQVQADNPIFSEEIFGPVAVLIPFAS
                DEQAIELANMGDYGLSAGIISSNVGRAMQLGAQLNVGLLHINDQTVNDETINPFGGFG
                ASGNATRIGGPANPDEFTQWQWMTIQAEAPHYPF"

promoter        8630..8610
                /gene="are operon"
                /note="putative"

gene            complement (8653..10470)
                /gene="areR"
                /note="putative sigma 54-type regulator"
                /product="AreR"
                /translation="MQISSSRDWCQVKDYFYKTGQVPEQTELDPTILQMWHKAKAAGL
                SPFTPVPASTQQPLTTDDIRLSETVFPILKDIWALFKHQNICTFFLNTQFKIIAEYQN
                PDPHSPYHFLKQGQVLDIAQFGSIAPSCSILSQMPIIMMGHQHYLNEFSDFYCASVPV
                FNGQGKILGALDITSYREQLASNWLRHLLYQGYVVENEIIKKNMPAQHKLLYFQHSED
                LLKTAYTGMIEIDATGQIQKANQMALTLLNTSIDQLLYKPISNYFSSISTLASLEQQT
                QFIRSQDEALFYARLYTPQMTKKTSHLFSTSTPDFKDIAKLSKILHSDVPILITGATG
                SGKDHLARQIHEFATKDQAFISINCAAIPEHLLESELFGYEAGAFTGASAKGKRGLIE
                LAHQGILFLDEIGDLPKHLQAKLLRVLQDQVFYRVGGHKPIQSQFKLISATHQDLLAM
                IDNHDFRSDLYYRICGYQIELPQLKDRPDKIDIFKTLLANAKIYSWSVNVQEQFEHYD
                WPGNIRELTHVIQLSAALIAHDSLEHAHLEHLHVPQQRKVTPQSYLANKPQTSLNQMT
                KQLVLDILQQEKGNISQAAKRLNISRTTLYKYLQ"
```

gene 10649..111884
/gene="benP"
/note="putative porin"
/product="BenP"
/translation="MQYSTTRLYQFMLRMGVTFGMVGCIVVIPTVVNAQDNWTLSLKN
AYIDRDYDGNAVKDTGSWSQGASLFYKSDYYKTPIDGLEIGVDGSVQYAVRLSHDKGV
ADTILPFDPVSQKQARDYLKYGGTLKLKYDQTELRVGELWPDLPVTAVDRSRQLLTSY
QGVSLNSKLTSKLTGEIGIISKVSPRNEEDFRKLTFTKNGIKYVSDGLNYIDLKYQVL
PNLKLKYFNASEDNKNYNIDSQNFGILETLKYRNHTVGLGYQQIVGDAYPLPDGFLPE
TYFINWNTTGFFKSDEKSVHFIYGYDFKDYVPGLNFTFKHVYGYDFKTATGLKNKEQE
SNFILNYAFQNPKLKGVGFQWLYIPYKVRYGTDFNENRLFLTYTKKF"

```
     1  CTCAATATCG TTCAGTTATT TGAGGGGTAT AGCTACATTA AAGAGTATGA
    51  AAATTAAGAG AAAATTAGTT TTCGCTTCAA TGATTCAATT TAAATCTAAA
   101  TATCAATGTT TTAAGATCAT AATGATGGAT ATGAAAACTT AAACCTATAC
   151  CGGATAGGCA ATTAAGAAAG AAATATGATA TAATTGTTTT AAGATATTGA
   201  TTTTAATAAA TTATTTAAAA AATTAATAAC TGCTTTAAAA AACTTTAAAT
   251  TGGATTTAAT TGAGTTTGCA GACTAAAAAT AAAACAATAA ATCAAGTTCT
   301  CAAAGGAAAT GAGTCGTGGG TAAAAAAATT AGTATAGCCA TTATTGGTGG
   351  TGGTATTGCA GGGGTCGCGC TTGCAGCCAA CTTATTTAAG CAACCACATT
   401  TAGAGGTTTG CTTGTATGAA GCAGCACCTC AATTTTCTGA AATTGGTGCC
   451  GGAATTTCAT TTGGCGCGAA TGCGGTTCGT GCCATTGAGC TACTTGGTTT
   501  GGCTTCGCAA TATACCGCGA TTGCAGATCA AGTATCTGCG CCATTTCAGG
   551  ATGTGTGGTT TCAATGGCGA AATGGTTATA ACGATGATTA TTTGTCGAGT
   601  TCAATTTCTC CTCAGGTTGG TCAGTCTTCG GTGCATCGTG CCGATTTCTT
   651  AGATGCTATT TTAGGGAATA TTCCACAGCA CCAGTGTAAG TTTAATAAAA
   701  AACTCAAATC GATTCAGGAG TACGACACAC ATATTGAATT GAGTTTTGAA
   751  GATGGTACAT GCGCCGAAGC CGATTATGTG ATTGGTGCAG ATGGCATACA
   801  CTCAGCCACG CGTGATTATG TGCTGCAAAC CCATCAGTTT GCTCCAGTGC
   851  GTCCTAATTT TACTGGAACA TGGGCTTACC GAGGCATTAT TAAAGCAGCA
   901  GAATTTAGGC AAGCCATTGT CGCAGCCGGC CTAGATGTAG AAATTGCCGA
   951  CGTACCACAA ATGTTTTTAG GCCAAAACAA GCACATTTTA ACCTTTCCGA
  1001  TTCGTCAAGG TGAAGACATT AACATTGTGG CGTTTAAGAC AAACCCTGAG
  1051  CAGCGTACGC TTCCAGAACA TACCCCATGG ACACGTGCGG TAGACAAACA
  1101  GGAAATGTTG GACGATTTTC AGGACTGGAG CGAAAGCTGC CGAATTTTAC
  1151  TCGGTTTGAT TGAGCGTCCG ACCTTGTGGG CACTGCACGA ACTGGCCGAA
  1201  TTGCCGACTT ATCAAAGCCA CTCTGGCCGC GTCATATTAA TGGGAGATGC
  1251  TGCGCACGCC ATGCTTCCAC ATCAGGGTGC CGGAGCAGGA CAAGGGCTTG
  1301  AAGATGCACT CACGCTCAAA GTATTGTTTG AGCACACTGA GCTGACTGTT
  1351  GAAGATTTAC CGCGAGTTTC TGCAATTTAT GAACAGATCC GAAAAGAACG
  1401  CGCCTGCAAG GTTCAGCGTA CTTCGCGTGA GTCTGGGCAA ATATATGAAC
  1451  TTAACTCAGC ACTGTATCCA AGCTTTGAAG CAGTGGGTGC ACATTTGCAA
  1501  AACCGTTTTG ACTGGTTATG GCAGCATGAT TTAGCACAAG ACATGTTAGC
  1551  CGCGCGAGCA GCAATACAAC CTGTTGCAAC GATTTAAACG CTAAGAATTT
  1601  GGCACAAGAG TGTTTTGAAC GACTTGTGCC TTTAAAACAA TTCTATTTTG
  1651  AAAGAGTTGA ATAAAAGTGT TTATGATAGG ATTAAATTAA AATCATGGAA
  1701  GATTCTAAAA CGTGGACTTA AGCCTGATCC GTATTTTTAT TTGTGTTTAT
  1751  GAAAATAAAA ATATCAGTAA AGCCGCTGAG ATTTTAAATC TGAGTCAACC
  1801  TTCTGTTACC TATAATTTAA ATCGATTACG TAAGCATTTA AATAATCCTT
  1851  TATTTGAACG CACGCAATAT GGTGTGGAAG CAACTAAATT ATCGCATGAA
  1901  TTATATCCTG TATTTAAAGA GTCGATTTTA AAAATTGAAA TAGCAGTCGA
  1951  TGAGGCATTA AATTTTAATC CACTCACTTC AAATAAAACC TTTCGAATTG
  2001  GTTTATCAGA TATTGGTGAA ATTTGCTTGT TACCGACATT AATCGAATAC
  2051  TTACGAGCAC ATGCACCAAA AATAAAAATA GAAGTAGAAG AGATTAAAAT
  2101  TGATCAAGTC GAAAAATGGT TAATCGAAGG ATTTATTGAT GTGGCTGTTT
  2151  TTAATAGTAC ACATTTGGAG TTTAAGCATC TTGAATATGA AACTCTTTTT
  2201  TTAGAGCGAT ATGTTGCACT GGTCAACATG AATCATCCGC GAATAAGAAG
  2251  TACGCTCAGT TTTGATGCGT ATTTGAATGA ATCTCATGTG GCGATAAAGT
  2301  CATCTACCGG GCATACTCAG GTCGATCATG TATTAAAACT AATGGGGCAT
  2351  CAGCGTAAAA TTGCCTTGGA AGTGCCACAT TTTGGAGTTT TACAAGGTGT
```

```
2401  GCTGGATAAA ACAGATTTGA TGGTGACGCT GCCCAGTCGT GCTGCACAGC
2451  AATATTTAAA TCAATCCCAT GTCCGTGTTT TGGAGTTGCC GTTTCAAATG
2501  TCTGAATTTT ATGTGGGCTT ACATTGGTTT GCCCAAACCA ATGAGCCCCT
2551  TGCTCGGATA TGGTTGATTC AAACCTGTAA AAAGGTTATT TCAGTTTTGT
2601  AAGACCGCGG GCTAAGCCCG GCAGTCTAAA CAATTAATCA TGACTGGCTT
2651  CAATAATTTT AGTCGCCAAT TCCTGAGGGC GGGTAAACAT TGCTTCATGG
2701  CTTCCACCCA TTTCAACCAA CTTAAATTCT CCGAGTCTGT TCGACATTTT
2751  AGGGTGCCAA AAGCCAGCAG GTAAAGCCTG ATCTTCGGTG CAGTTCAGGT
2801  AGCTTTTTGG AATATTCAGC GTATAGAATT TTTTTAGATC CAGTCGATCA
2851  TGAAAGGGCT GACAAGGCTC AGGCGTAAGC ATCTTGTAAG TTTCCTCGGC
2901  CATTTGATCA TCTGCATCAT TGATGAAAGC ATGACGCCAT ACTTCATAAG
2951  GCAGTAAAAC AGTATTATCA CCCGATGCTG CGGCCAAACT CGTAAACAGC
3001  TCGTAGTAGG CTTCTGGCAT GTTATCAAAC ATATTTTCGC CGTCTTGTAA
3051  AACAAAAGCA TTCCAGTAAA TCAGCCTTTG AATGCGCTCA GGAATGGCCT
3101  CGGCGACTTT GGAAATAATG GTGCCACCAT AGCTGTGCCC AAGCAGTACA
3151  ATCTCGCTGA GATCATGCTT TACGATATAA TCCGCAATCG ATTGACTGCA
3201  CTGCGCGTGA GTCACATTCC GGTCGGCATT TGGGCCATGG CCTGCCACGG
3251  TCGGACAGTG CACGGTATGG CCTTGAGCAC GTAAATGAGT GGCTACAGGT
3301  TCCCAGAGAG AGCCATCGTG CCATGAGCCA TGAACAAGTA CATATGTTAT
3351  CATTTTGATC GTCCTTTAAT GTGTTGATCG GGTTTAAAAG TCTTTTGAAA
3401  TTCCGAGGGT AATCACATGG TTTTGATTAC TTAAATTGGT GCCTGTACTG
3451  TTTCCAGTAC CGTTTAAGGT TTTGTTCACC GCATATTCAT AGCTAAGATT
3501  GAGGGTTGCA AAACCAAGAT CATGGCCGTA GCCAATAGTC CATGATTGAT
3551  TGGCGACGGC ATTGGCATAA AAATTCACCA ATGTGTCTTG AGCATCTACA
3601  ACTGAGTCTG CATGGCTATA GCCAAAACGC AAATGATCAT GAGGTGTGAG
3651  CTGATAATCC ACGCCAGCTC GATAAACCGT TTGATTTTGC CAGTTAAAAG
3701  AACCTTTTTT TCCTAATCCA TCAGCATCTT GCCAATTAAA ACGTACGATA
3751  TCGGCTGCTA GAGTAAGCTT TTGATTTATT TTGTGTTTAA TTCCGATTCC
3801  ATAACGTTCT GGCAAATTAA TTCTGCCTTC ACTGCGTGCC AATAAATCAT
3851  TTTTATAGTC ATCAAGCTTC GAAAATTTCA TTTCAGTGCT ATAGCTCGCG
3901  CCAAGCCACC AATTGGGCTG AAATTCCCAA GTACTACCTA TGCTTGCACC
3951  AACACCATAA GCCCACTGAT TACCATGACT TTCTAGAAAA ATGGGCGTAC
4001  CTTGTCCATC GACTCCTGCA AGCACACCTT TTGCCCTGAA TTGCTGAATA
4051  CCTAAATCTA AACTGGCACC CAAACTGAGA TTGGGTAAAA CCTTGTACGA
4101  TACAGTCGGG CTTGCCTTTA CAATTGCCAG ATTATCTTTG GCTTTATCGG
4151  ATGGAAATCC ATGGATTGCA GGCTCATCAT AGTCTACCGC TGCACCAGCA
4201  CCCGTGAGAC TTAGTCCCCA ACTTAAGCGG TCGTTATAAT GCTGGACTAA
4251  AGCCAGACTT GGGATCGGAA TGACTTTGTC GGAATGAAAG CCTTCTGGCT
4301  GATTCAGTAA GGTCGCTTTA GCACGAATAT TTAATAAAGA AAGTTGGGCG
4351  TCTGCACCAC TGTTTTGAAA AGATAGATTG GCAGGATTAT CTTGAATCGA
4401  AGTTCGATCC ATACCCTGAG CAAGACTTAC TCCGCCCATA CCGCTAGAGA
4451  CTTGTCCTGA TCCAGTATTA TAAACGCCAT TAAATGCATA GCTTGATAAT
4501  GGCATGCAGG CAATCAGAAC CAGAGATAAC TTATGAGGAA ACATACTCAA
4551  AATCCTTGTA TTGAATAAGT CAAACCATTT GACTTTTATT TATCTAAAAT
4601  TATTTAGAGT ATTTTAAATT TTAAATGATT GGTTTTATAT TTTTAGTCAT
4651  GTGAGAGTAT ATATTTATAC ATTATTGATA TAATAATTTT ACAGGTAAAT
4701  TAAAAAGGTT CTTATTAAGA ACCTTTTTAA TTAAAACTGA TTATTTGGAT
4751  TTTGTATTGT TTTTTAACCA CGCTCTTATT ATATTGATTG TCTCTTTTGT
4801  TTTATTGAGT ACGCCATACA TTTCAAAAAA ACCATGTACC ACACCTTTTA
4851  TTTCATGATA TTCAGTATTT ACTAAACTGC TTTTTAATTT TTCAAAATAA
4901  ACACAACCCT GATCATGTAA AGGATCAAAT TCGGCTGTTA TAATTAATGT
4951  AGATGGAAGA TGTTGGTGAC TTTGAGCGAG TAGTGGCGAG GCAAGCCATT
5001  GATAAGCAAA ATCTTCATTT TTCAAATAGA TTTGTATCAT CTTTGACATT
5051  TGTTCGGCAG ATAGAGGAGG AGTGAGCTTC TGTTTTTCAA AGGATGGATA
5101  TTTGTTGTCA TGTATTACGG TAAGATCAAT ACATGGGTAG ATTAAAATTT
5151  GCCCTGCTAA AGAAGTTCTA TACGTTTGTT CTTTAAAAGT GAGAGCAGTG
5201  ATTGCTGCCA GATTGCCCCC TGCGCTGTCT CCTGCAACAT ATAAATGGTC
5251  TGGATGAATG CCAAGAGCAG ATGCATGTTC AAAAATCCAC TGTGTGGCAT
5301  GAATCGCATC ATGAGCAGCT TGAGGGTATG GGTATTCAGG CGCAAGACGA
5351  TAATCGACAC TAATGACATT AACATCTAGA TCATTCACTA AATGCCGACA
5401  AAGTAAATCA TGCGTTTCGA TACTCCCTAT ACACCAACCT CCGCCATGGA
```

```
5451  AGTAAATCAG TGTTTCACAT GATTCACTAT GCTGGCGTTG AGTTTTATAC
5501  AATCTTAAGC GTAAGCTACC ATCGTCTACT TGTATCCATA AATCTTCTAC
5551  GTCTTGAAGT TCAATGAATT GCCCCGCTAC ATGAGGACCA AAATTGAGAT
5601  ACCAGTTTCT CATTTCATTG ATATTTGGGT CATTTAGAGT ACCAATATCA
5651  AATGCAGGTG CACCTTGATT TTCAAGCATT GCAAAAAAAC TTGCCAGTTG
5701  AGGATCTAAA CTCTGTAACA TGTATTAAAA ATCCTTTTTA AATCTTTTTG
5751  ATTTATTTAC GATTTGAGAA GTTATGTATT GTTCAAGACA TAACTTCTCA
5801  ATGCTTAACC AATTTTTAAA ATGGGTTTGA GCGTGATGCC TTTATGGCTA
5851  TCAATTGCAG CCTGATTAAT ATCTTTAAAA TCATAGAACT TAACCAGTTT
5901  GTCGAACGGA AACTTACCTT GCTGATACAG CGAAACCAAA GCCGGAATAA
5951  ACTTTTTCGG TACGCCGCTG CCTTCTACAA CGCCGATAAT CGATTTGCCG
6001  CCAAGCAATA AATCATTCAC GTCAAATGCT GCGGTCGTTC CTAAAGGCGG
6051  AGCACCTACC ACTGCAATGC TGCCAAGAAT ACCGAGCGCA TCGATCCCTT
6101  GTTTCAAAAT TTCTGGGCGT CCAGTCGATT CAAGAGCAAA CTTTACACCA
6151  CCGCCTGTGA TTTCCTTGAT GGCTTCGACG GGATCTTGCG TTTTACTGTT
6201  GATAACATGG GTCGCACCAA GTTCTTTGGC TAACGCCAAA CGAGACTCGA
6251  CAATATCCAC TGCAATAATA GTGGTCGCAC CACATACTTT TGCAGCAAGA
6301  AGTGCACTGA GTCCGACTGC GCCTGCACCC CAAGTGACAA AGCTACTGGC
6351  AGGCGCGACT TTCAGCGCAT TGATTGCAGC ACCTGCACCA GTTTGTATGC
6401  CGCAGCCGAG TGGCCCCAAA AGTTCAAGCG GTACATCTTT ACTCACTTTT
6451  ACGGCATTAT TTTCACGGCT GATGGCGTAG GTTGCAAACG AAGATTGCGC
6501  AAAGAAATGA TCATGAATTT TATGATGATC GTGCGAGCAT ACGGCAGAAT
6551  GCCCATCTGG ATCTCCACCA CTGAAGTTTC GTCCAAAAAA ATCCTGACAG
6601  TAGGCTGGTT GTCCGCTGCT ACACTGCTCA CAGCTTCCAC AAAAGCCATA
6651  GCTGAGCACC ACGTGATCAC CTACAGCCAA ATCTTTGACA TTCGGGCCAA
6701  TGGCTTCAAT GATGCCTGAA CCCTCGTGAC CAAGTACCGC TGGTAATGGA
6751  ACAGGATAAT ACTGATCACG CACAATCAGA TCGGTATGGC ACATTCCCGT
6801  TGCCACAATT TTAACCAATA CTTCATCGTT TTGAGGTTCA CGAATTTTGA
6851  GCTGTTTGAG CTCAAAGTCT GCGCCTTTGC ATTCGGTCAC TGCTGCTGTG
6901  ATTTCGGTAA ACTTTGTCAT TTCTATACAT CCTGTATGAG AAAAAAATTA
6951  AAATGGATAG TGTGGGGCCT CAGCCTGAAT GGTCATCCAT TGCCACTGAG
7001  TAAATTCATC TGGATTTGCT GGCCCACCAA TTCGTGTTGC ATTACCAGAC
7051  GCGCCAAATC CGCCAAATGG ATTAATCGTT TCATCATTTA CGGTCTGATC
7101  ATTGATATGT AACAAACCAA CATTTAGCTG TGCACCGAGT TGCATGGCAC
7151  GCCCCACGTT GGATGAAATA ATGCCAGCCG AAAGCCCATA ATCTCCCATA
7201  TTGGCCAGTT CAATCGCTTG TTCATCGCTT GCAAAAGGAA TCAGTACGGC
7251  CACTGGCCCA AAGATCTCTT CACTGAAAAT TGGATTGTCT GCTTGCACCT
7301  GTGAAAGTAC GCTTGGCTCA AAGAACGGGC CATTCGCTTT ACCTCCAATT
7351  TCAAGGGTCG CGCCGTGTTC AACCGCGGCT TGAACCAATG CTTCTACGCG
7401  CTTGCTTTGC TTTTCATTGA TCAAAGGGCC AATGGTCACG TTTTGCTCTA
7451  AAGGATTACC CACCACAAAG CGTTTGACCT TTTCGATTAC ACGCGCTTTG
7501  AGGGCATCGT AAATTTTTTC ATGAATTAAG ATTTTGCCAG ATGTCATGCA
7551  AATCTGCCCA GAATGCAAAA ATGCACCCCA TGCAATATTC TCGGCTGCTA
7601  AATCGAGATC GGCATCGTCG AGAATAATCA GTGAGTTTTT GCCGCCAAGT
7651  TCTAACGAGA CTTTTTTCAA AGTTTTGCCT GCATTGGCAC CAATGATACG
7701  GCCCACCTGT GTTGAACCCG TAAACTGAAT GCTGGCAATA TGTGGATCGA
7751  GTGTGAGTGC TTCACCTGTA TCGGCGCCAC CTGGTAATAC ATGCAGTAAT
7801  CCCTTTGGTA AGCCAGCCAG TTCAAAAATG CGGGCAATTG CGTAACCACT
7851  ACACACCGCC GTTCGTTCAT CTGGCTTAAG TACCACCGCA TTACCAAATG
7901  CCAAAGCTGG CCCAACCGCA CGAAGCGCCA AATAAAGCGG AAAATTAAAA
7951  GGTGAAATCA CGCCGACCAC ACCCAACGGT AAGCGTTTGG CCAAGCTTAA
8001  TTTACCGTTG CGACTCGGTA ATATGTTCCC CTGATCTAGC TGCGGAAATG
8051  CAATACAGTT TTTAAGCATT TGAATGCTTA CTTTCACTTC AAACTGAGCT
8101  TTGAGTGCCA GTGAGCCACT TTCACGAATC AGCCAATTGA CCAGTTCATC
8151  TTGATGCTCT GTGGCAATCT GTACGGCTTT TTCAAATACG GCTTGGCGTT
8201  CTAAATAATC GAGCGCCCAC CACTGCCGCT GTGCCGTTTG CGCTTCTTTC
8251  GCTGCTTGAG CAACATCTGT AGGCGATGCA GAGCCTGTCT GGCCGAGCTG
8301  CTCACCTGTG GCAACTTCTA CCACACCATA GGGTTGTCCG CTTTCAAACC
8351  AACCGCCATT AAATAATTTT TTGTTCCAAA TTGAAGCATC AAGTAATGTC
8401  ATATAGACCT CGTAACGTCC TGTTCGTGCT TTTTCGATAA TTTCGTTCTT
8451  ACAATTGTTT AAAACCACAC AGACTAAAAT CGTATGTGTA GTATCGATAT
```

```
 8501  CTCACCTATT GCAGGGAATG TGCCACTATT TAAATTATTG TTTTTATTAA
 8551  ATTAATTTTT TTAAATTTTC GGGTAATTGT TCAGGGGCTG GACACTATTG
 8601  AATTGTACAT TGTCCAGAAT CTGAACAATG TACAATACTG TCCTTGAGCT
 8651  TATTACTGTA AATACTTATA CAGTGTTGTT CTTGAGATAT TGAGTCGCTT
 8701  AGCGGCTTGA GAGATGTTGC CTTTTTCTTG TTGCAAAATG TCTAAAACGA
 8751  GCTGTTTAGT CATCTGATTC AGTGAAGTTT GAGGTTTATT TGCTAGGTAG
 8801  GATTGGGGCG TAACTTTTCG TTGCTGTGGT ACATGTAAAT GTTCAAGATG
 8851  CGCATGTTCA AGAGAATCAT GCGCTATAAG CGCCGCAGAA AGTTGAATCA
 8901  CATGAGTCAA TTCTCGAATG TTTCCTGGCC AATCATAGTG CTCGAACTGT
 8951  TCTTGAACAT TGACCGACCA ACTGTAAATT TTGGCGTTGG CCAGCAGTGT
 9001  CTTAAAAATA TCAATTTTAT CTGGACGGTC TTTGAGTTGC GGTAACTCGA
 9051  TCTGATAGCC ACATATTCTG TAATACAAAT CACTTCTGAA GTCATGATTG
 9101  TCAATCATTG CCAATAGATC TTGATGGGTG GCAGAAATTA ATTTGAACTG
 9151  ACTTTGAATC GGCTTATGAC CGCCTACACG ATAAAAAACC TGATCTTGTA
 9201  AAACGCGTAA TAATTTTGCC TGTAAATGTT TGGGTAAATC ACCAATTTCA
 9251  TCTAAAAACA AAATGCCCTG ATGAGCAAGT TCGATCAAAC CACGTTTGCC
 9301  TTTGGCACTT GCTCCTGTAA AAGCACCTGC TTCATAGCCG AAAAGTTCAG
 9351  ATTCAAGTAA GTGTTCAGGA ATAGCTGCAC AGTTAATCGA AATAAATGCC
 9401  TGATCTTTTG TCGCAAATTC GTGAATTTGA CGTGCCAAAT GATCTTTACC
 9451  TGAGCCTGTG GCACCCGTAA TTAAAATGGG TACATCGGAA TGCAAGATTT
 9501  TGCTTAATTT GGCAATATCT TTGAAATCGG GTGTAGATGT GCTGAATAGG
 9551  TGGCTTGTTT TTTTAGTCAT TTGGGGCGTG TATAAACGTG CATAAAACAG
 9601  CGCTTCGTCT TGACTACGTA TAAATTGCGT CTGCTGCTCA AGTGAGGCGA
 9651  GGGTACTTAT CGAACTGAAA TAGTTCGAAA TGGGCTTATA CAATAATTGA
 9701  TCGATAGACG TGTTGAGTAG AGTTAGTGCC ATTTGATTGG CTTTTTGAAT
 9751  CTGCCCCGTG GCATCAATTT CAATCATGCC AGTATAGGCT GTTTTTAATA
 9801  AATCTTCTGA ATGCTGAAAA TAAAGCAGCT TATGCTGTGC TGGCATATTC
 9851  TTTTTTATAA TTTCATTTTC AACCACATAC CCCTGATAAA GCAGGTGTCT
 9901  TAGCCAGTTA CTGGCCAGTT GTTCACGATA CGAGGTGATG TCTAGCGCGC
 9951  CCAGTATTTT GCCCTGACCA TTGAACACAG GTACCGATGC ACAGTAAAAG
10001  TCAGAAAATT CATTTAAATA ATGCTGGTGG CCCATCATAA TGATCGGCAT
10051  CTGACTCAAA ATACTACACG AGGGCGCAAT ACTGCCAAAC TGTGCAATAT
10101  CCAGAACTTG ACCTTGCTTT AAAAAATGAT AGGGCGAATG AGGGTCAGGA
10151  TTTTGATATT CGGCAATAAT TTTAAATTGG GTATTAAGGA AAAATGTACA
10201  AATATTCTGA TGTTTAAACA GCGCCCAAAT ATCTTTCAAA ATAGGAAAAA
10251  CCGTTTCAGA TAGACGAATA TCGTCTGTGG TTAGGGGCTG CTGTGTCGAA
10301  GCTGGAACTG GCGTGAACGG GCTCAAGCCC GCTGCTTTGG CTTTATGCCA
10351  CATCTGTAAA ATGGTTGGAT CAAGCTCGGT TTGTTCTGGA ACTTGACCCG
10401  TTTTATAAAA ATAATCTTTA ACTTGGCACC AGTCACGACT TGATGAAATT
10451  TGCATAGTTT GAATCCTTCA AAAACAAAGA AACGTTGCTC GCGACTACAT
10501  CCGAAACAAC ATCATCTTTT CATCGTCCAT TGAATCTTTT TTGACTATAG
10551  CAAAGATACT TTTTGAGTTG TGTATAAAAG TTGATTTTAT AGATTTTAGG
10601  CGTTTTATAC GCATTAACAT ATCCTAAAGG TATGAATATA TGTATTTTTG
10651  AATATCGACA CACGTGATTA AAAAAAATAA ATTAGCGATC AATAATTATT
10701  TTTATTACAT GATCAGGATG ATCTTTTGGG AAGGACCTAT GCAATATTCA
10751  ACAACGCGTC TTTATCAGTT TATGCTACGA ATGGGGGTTA CTTTTGGCAT
10801  GGTAGGCTGT ATCGTTGTGA TACCGACAGT GGTCAATGCT CAAGATAACT
10851  GGACACTTAG CCTCAAAAAC GCTTATATCG ATCGTGATTA TGACGGCAAT
10901  GCTGTCAAGG ACACCGGTAG CTGGTCGCAA GGCGCTTCTT TATTTTATAA
10951  ATCTGACTAT TACAAAACAC CCATCGATGG ACTTGAAATT GGTGTGGATG
11001  GTTCAGTGCA ATATGCGGTT CGTTTGAGTC ATGACAAGGG GGTGGCGGAT
11051  ACCATTTTAC CTTTTGATCC GGTAAGCCAA AAGCAGGCGC GTGATTATTT
11101  GAAATATGGT GGAACACTTA AACTTAAATA TGATCAAACC GAACTTCGTG
11151  TGGGTGAGCT TTGGCCAGAT TTACCTGTCA CTGCGGTAGA CCGTAGCCGT
11201  CAGTTACTGA CTTCGTATCA GGGCGTTAGC CTCAATTCGA AACTGACCAG
11251  TAAACTTACC GGTGAAATCG GGATAATTTC TAAAGTATCA CCACGTAACG
11301  AAGAAGATTT TCGTAAACTG ACCTTCACCA AAAATGGCAT CAAATACGTT
11351  TCGGACGGTT TAAATTACAT CGATTTAAAG TATCAGGTTT TACCGAATTT
11401  AAAGCTTAAA TATTTTAATG CCAGTGAAGA CAATAAAAAT TACAACATTG
11451  ATAGCCAAAA CTTCGGAATT TTAGAAACTT TGAAATACCG AAATCATACT
11501  GTCGGATTGG GTTATCAGCA AATTGTGGGA GATGCTTACC CACTGCCAGA
```

```
11551  TGGTTTTTTA CCTGAAACCT ATTTCATCAA CTGGAACACC ACAGGCTTCT
11601  TTAAATCTGA TGAAAAATCA GTTCACTTTA TCTATGGCTA CGACTTTAAA
11651  GATTATGTAC CGGGTTTAAA TTTTACCTTT AAGCATGTGT ATGGGTACGA
11701  CTTTAAGACC GCAACAGGTC TGAAAAACAA AGAGCAAGAA AGCAACTTTA
11751  TTTTGAATTA TGCCTTTCAA AATCCAAAGC TAAAAGGCGT CGGATTCCAG
11801  TGGTTGTATA TCCCGTATAA AGTCCGTTAT GGAACGGATT TTAACGAAAA
11851  TCGTTTATTT CTAACTTACA CCAAAAAATT CTGAGTGCTT GGTGCTAAAT
11901  CCTTATGGAT GGAGGAATGT T
```

References

Abril, M-A., C. Michan, K. N. Timmis, and J. L. Ramos. 1989. Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. J. Bacteriol. **171:**6782-6790.

Alon, R. N., and D. L. Gutnick. 1993. Esterase from the oil-degrading *Acinetobacter lwoffii* RAG-1: sequence analysis and over-expression in *Escherichia coli.* FEMS Microbiol. Lett. **112:**275-280.

Alon, R. N., L. Mirny, J. L. Sussman, and D. L. Gutnick. 1995. Detection of α/β-hydrolase fold in the cell surface esterases of *Acinetobacter* species using an analysis of 3D profiles. FEBS Letters. **371:**231-235.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-410.

Amund, O. O. 1997. Utilization and degradation of an ester-based synthetic lubricant by *Acinetobacter lwoffi.* Biodegradation. **7:**91-95.

Assinder, S. J., and P. A.Williams. 1990. Plasmids determining the catabolism of toluene and xylene: the TOL plasmids. Adv. Microb. Physiol. **31**:1-69.

Ausubel, F. M., R. Brent, R. E. Kingston, D. E. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.

Averhoff, B., L. Gregg-Jolly, D. Elsemore, and L. N. Ornston. 1992. Genetic analysis of supraoperonic clustering by use of natural transformation in *Acinetobacter calcoaceticus.* J. Bacteriol. **174:** 200-204.

Bairoch, A., P. Bucher, and K. Hofman. 1997. The PROSITE database, its status in 1997. Nucl. Acid Res. **24:**217-221.

Bauchop, T., and S. R. Elsden. 1960. The growth of microorganisms in relation to energy supply. J. Gen. Microbiol. **23:**457-469.

Baumann, P., M. Doudoroff, and R. Y. Stanier. 1968. A study of the *Moraxella* group II. Oxidative-negative species (genus *Acinetobacter*). J. Bacteriol. **95:**520-541.

Biswas, G. D., S. A. Lacks, and P. F. Sparling. 1989. Transformation deficient mutants of piliated *Neisseria gonorrhoeae.* J. Bacteriol. **171:** 657-664.

Bosch, R., E. GarciaValdes, and E. R. B Moore. 1999. Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from *Pseudomonas stutzeri* AN10. Gene. **236:**149-157.

Brady, L., A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Christiansen, B. Hugejensen, L Norskov, L. Thim, and U. Menge. 1990. A serine protease triad forms the catalytic centre of a triacylglycerol lipase. Nature. **343:**767-770.

Brahimi-Horn, M. C., C. A. Mickelson, A. M. Goal, M. G. Guglielmino, and L. G. Sparrow. 1991. Lipolytic activity produced by *Pseudomonas aeruginosa* and *Acinetobacter calcoaceticus* strains grown in wool-scour effluent. Enzyme. Microbiol. Technol. **13:**740-746.

Breuil, C., and D. J. Kushner. 1975. Lipase and esterase formation by psychrophilic and mesophilic *Acinetobacter* species. Can. J. Microbiol. **21:**423-433.

Brzozowski, A. M., U. Derewenda, Z. S. Derewenda, G. Dodson, D. M. Lawson, J. P. Turkenburg, F. Bjorkling, B. Hugejensen, S. A Patkar, and L. Thim. 1991. A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. Nature. **351:**491-494.

Buck, M., and W. Cannon. 1992. Specific binding of the transcriptional factor sigma-54 to promoter DNA. Nature. **358:** 422-424.

Bundy, B. M., A. L. Campbell, and E. L. Neidle. 1998 Similarities between the *antABC*-encoded anthranilate dioxygenase and the *benABC*-encoded benzoate dioxygenase of *Acinetobacter* sp. strain ADP1. J. Bacteriol. **180:** 4466-4474.

Burgess, R., A. Travers, J. J. Dunn, and E. K. F. Bautz. 1969. Factor stimulating transcription by RNA polymerase. Nature. **221:**43-46.

Busch, S., C. Rosenplänter, and B. Averhoff. 1999. Identification and characterization of ComE and ComF, two novel pilin-like competence factors involved in natural transformation of *Acinetobacter* sp. strain BD413. Appl. Env. Microbiol. **65:**4568-4574.

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

Chalmers, R. M., A. J. Scott, and C. A. Fewson. 1990. Purification of the benzyl alcohol dehydrogenase encoded by the TOL plasmid pWW53 of *Pseudomonas putida* MT53 and their preliminary comparison with benzyl alcohol dehydrogenase I and II from *Acinetobacter calcoaceticus*. J. Gen. Microbiol. **136:**637-643.

Chari, R. V. J., C. P. Whitman, and J. W. Kozarich. 1987a. Absolute sterochemical course of muconolactone D-isomerase and of 4-carboxymuconolactone decarboxylase: a ^{1}H NMR "riochet" analysis. J. Am. Chem. Soc. **109:**5520-5521.

Chari, R. V. J., C. P. Whitman, J. W. Kozarich, J. W. Ngai, and L.N. Ornston. 1987b. Absolute sterochemical course of the 3-carboxymuconate cycloisomerase from *Pseudomonas putida* and *Acinetobacter calcoaceticus*: analysis and implications. J. Am. Chem. Soc. **109:**5514-5519.

Choo, D. W., T. Kurihara, T. Suzuki, K. Soda, and N. Esaki. 1998. A cold-adapted lipase of an Alaskan psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and characterization. Appl. Environ. Microbiol. **64:**486-491.

Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty. 1993. Nucleotide-sequence and initial functional-characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida.* J. Bacteriol. **175:**417-427.

Collier, L. S., N. N. Nichols, and E. L. Neidle. 1997. *benK* encodes a hydrophobic permease-like protein involved in benzoate degradation by *Acinetobacter* sp. strain ADP1. J.Bacteriol. **179:**5943-5946.

Collier, L. S., G. L. Gaines III, and E. L. Neidle. 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J. Bacteriol. **180:**2493-2501.

D'Argenio, D. A., A. Segura, W. M. Coco, P. V. Bünz, and L. N. Ornston. 1999. The physiological contribution of *Acinetobacter* PcaK, a transport system that acts upon protocatechuate, can be masked by the overlapping specificity of VanK. J. Bacteriol. **181:**3505-3515.

Dabnau, D. 1991. Genetic competence in *Bacillus subtilis.* Microbiological Rev. **55:**395-424

deHaseth, P. L., M. L. Zupancic, and M. T. Record Jr. 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. J. Bacteriol. **180:**3019-3025.

Delgado, A., and J. L. Ramos. 1994. Genetic evidence for the activation of the positive transcriptional regulator XylR, a member of the NtrC family of transcriptonal regulators, by effector binding. J. Biol. Chem. **269:**8059-8062.

Devereaux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387-395.

Dixon, R. 1998. The oxygen responsive NifL-NifA complex: a novel two component regulatory system controlling nitrogenase synthesis in gamma-proteobacteria. Arch. Microbiol. **169:**371-380.

Dombroski, A. J., W. A. Walter, M. T. Record Jr., D.A. Seigele, and C.A. Gross. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exibit specificity of binding to promoter DNA. Cell. **3:**501-512.

Dombroski, A. J. 1997. Recognition of the -10 promoter sequence by a partial polypeptide of sigma70 in vitro. J. Biol. Chem. **272:**3487-3494.

Doten, R. C., K.-L. Ngai, D. J. Mitchell, and L. N. Ornston. 1987. Cloning and genetic organization of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**3168-3174.

Drummond, M. H., A. Contreras, and L. A. Mitchenall. 1990. The functions of isolated domains and chimaeric proteins constructed from transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. Mol. Microbiol. **1:**29-33.

Dunn, N.W., and I. C. Gunsalus. 1973. Transmissible plasmid coding early enzymes of naphthalene oxidation in *Pseudomaonas putida*. J. Bacteriol. **114:**974-976.

Eklund, H., J-P. Samama, L. Wallen, and C-I. Branden. 1981. Structure of a triclinic ternary complex of horse liver alcohol dehydrogenase. J. Mol. Biol. **146:**561-587.

Elsemore, D. A., and L. N. Ornston. 1995. Unusual ancestry of Dehydratases associated with Quinate catabolism in *Acinetobacter calcoaceticus*. J. Bacteriol. **177:**5971-5978.

Eraso, J. M., and S. Kaplan. 1994. *prrA*, a putative response regulator involved in oxygen regulation of photosynthetic gene expression in *Rhodobacter spheroides*. J. Bacteriol. **176:**32-43.

Farres, J., T. T. Y. Wang, S. J. Cunningham, and H. Weiner. 1995. Investigation of the active-site cysteine residue of rat-liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. Biochemistry. **34:**2592-2598.

Fernández, J., A. A. DiMarco, L. N. Ornston, and S. Harayama. 1995. Purification and characterization of *Acinetobacter calcoaceticus* 4-Hydroxybenzoate 3-Hydroxylase after its overexpression in *Escherichia coli*. J. Biochemistry. **117:**1261-1266.

Fischer, S. E., and H. P. Kleber. 1987. Isolation and characterisation of the extracellular lipase of *Acinetobacter calcoaceticus* 69V. J. Basic Microbiol. **27:**427-432.

Gaines, G. L., L. Smith, and E.L. Neidle. 1996. Novel nuclear magnetic resonance spectroscopy methods demonstrate preferential carbon source utilization by *Acinetobacter calcoaceticus*. J. Bacteriol. **178:**6833-6841.

Gallegos, M. T., R. Schleif, and J. L. Ramos. 1993. AraC/XylS family of transcriptional regulators. Microbiology and Molecular Biology Reviews. **61:**393.

Geißdörfer, W., R. G. Kok, A. Ratajczak, K. J. Hellingwerf, and W. Hillen. 1999. The genes *rubA* and *rubB* for alkane degradation in *Acinetobacter* sp. strain ADP1 are in an operon with *estB*, encoding an esterase, and *oxyR*. J. Bacteriol. **181:**4292-4298.

Gerischer, U., and L. N. Ornston. 1995. Spontaneous mutations in *pcaH* and *-G*, structural genes for protocatechuate 3,4-dioxygenase in *Acinetobacter calcoaceticus*. J. Bacteriol. **177:**1336-1347.

Gerischer, U., A. Segura, and L. N. Ornston. 1998. PcaU, a transcriptional activator of genes for protocatechuate utilization in *Acinetobacter*. J. Bacteriol. **180:**1512-1524.

Gillooly, D. J., A. G. S. Robertson, and C. A. Fewson. 1998. Molecular characterization of benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II of *Acinetobacter calcoaceticus*. Biochem. J. **330:**1375-1381.

Gralton, E. M., A. L. Campbell, and E. L. Neidle. 1997. Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the *Acinetobacter* sp. strain ADP1 (BD413UE) chromosome. Microbiology. **143:**1345-1357.

Gregg-Jolly, L. A., and L. N. Ornston. 1990. Recovery of DNA from the *Acinetobacter calcoaceticus* chromosome by gap repair. J. Bacteriol. **172:**6169-6172.

Gregg-Jolly, L. A., and L. N. Ornston. 1994. Properties of *Acinetobacter calcoaceticus recA* and its contribution to intracellular gene conversion. Mol. Microbiol. **12:**985-992.

Grund, E., and H. Kutzner. 1998. Utilization of quinate and *p*-hydroxybenzoate by actinomycetes: key enzymes and taxonomic relevance. Journal of Basic Microbiology. **38:**241-255.

Gutnick, D. L., R. Allon, C. Levy, R. Petter, and W. Minas. 1991. Applications of *Acinetobacter* as an industrial microorganism. The Biology of *Acinetobacter*. Plenum Press, New York and London. pp. 411-441.

Harayama, S., M. Rekik, M. Wubbolts, K. Rose, R. A. Leppik, and K. N. Timmis. 1989. Characterisation of five genes in the upper-pathway of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products. J. Bacteriol. **171:**5048-5055.

Harayama, S., M. Rekik, A. Bairoch, E. L. Neidle and L. N. Ornston. 1991. Potential DNA slippage structures acquired during evolutionary divergence of *Acinetobacter calcoaceticus* chromosomal *benABC* and *Pseudomonas putida* Tol pWW0 plasmid *xylXYZ,* genes encoding benzoate dioxygenases. J. Bacteriol. **173:**7540-7548.

Harley, C. B., and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. Nucl. Acid Res. **15:**2343-2361.

Harwood, C. S., and R. E. Parales. 1996. The β-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50:**553-590.

Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter sequences. Nucl. Acids Res. **11:**2237-2255.

Helmann, J. D. 1994. Bacterial sigma factors, p. 1-17. In R. C. Conaway, and J. W. Conaway (ed.), Transcription: mechanisms and regulation. Raven Press, New York, N.Y.

Hempel, J., H. Nicholas, and R. Lindahl. 1993. Aldehyde Dehydrogenases – Widespread structural and functional diversity within a shared framework. Protein Sci. **2:**1890-1900.

Holloway, B. W., U. Romling, and B. Tummler. 1994. Genomic mapping of *Pseudomaonas aeruginosa* PAO. Microbiology. **140:**2907-29.

Hoover, T. R., E. Santero, S. Porter, and S. Kustu. 1990. The integration host factor stimulates interaction of RNA polymerase with NifA, the transcriptional activator for nitrogen fixation operons. Cell. **63:**11-22.

Hopper, S., and A. Bock. 1995. Effector mediated stimulation of ATPase activity by the sigma-54 dependant transcriptional FhaL from *Eschericia coli.* J. Bacteriol. **177:**2798-2803.

Hovarth, B., C. W. B. Bachem, J. Schell, and A. Kondorosi. 1987. Host specific regulation of nodulation genes in *Rhizobium* is mediated by a plant signal, interacting with *nodD* gene product. EMBO J. **6:**841-848.

Huang, J., and M. A. Schell. 1991. In vivo interactions of the NahR transcriptional activator with its target sequences: inducer mediated changes resulting in transcription activation. J. Biol. Chem. **266:**10830-10838.

Inoue, J., N. Tomioka, A. Itai, and S. Harayama. 1998. Proton transfer in benzyl alcohol dehydrogenase during catalysis: alternate proton-relay routes. Biochemistry **37:**3305-3311.

Jaeger, K. E., S. Ransac, B. W, Dijkstra, C. Colson, M. Vanheuvel, and O. Misset. 1994. Bacterial Lipases. FEMS Microbiology Reviews. **15:**29-63.

Jones, Rh. M., L. S. Collier, E. L. Neidle, and P. A. Williams. 1999. *areABC* genes determine the catabolism of aryl esters in *Acinetobacter* sp. strain ADP1. J. Bacteriol. **181:**4568-4575.

Jones, Rh. M., V. Pagmantidis, and P. A. Williams. 2000. *sal* Genes determining the catabolism of salicylate esters are part of a supraoperonic cluster of catabolic genes in acinetobacter sp. strain ADP1. J. Bacteriol. **182:**2018-2025.

Juni, E. 1972. Interspecies transformation of *Acinetobacter*: genetic evidence for a ubiquitous genus. J. Bacteriol. **112:**917-931.

Juni, E., and A. Janik. 1969. Transformation of *Acinetobacter calcoaceticus* (Bacterium anitratum). J. Bacteriol. **98:**281-288.

Kahng, H-Y, A. M. Byrne, R. H. Olsen, and J. J. Kukor. 2000. Characterization and role of *tbuX* in utilization of toluene by *Ralstonia pickettii* PK01. J. Bacteriol. **182:**1232-1242.

Kelly, M. T., and T. R. Hoover. 1999. Bacterial enhancers function at a distance. ASM News. **65:** 484-489.

Kessler, B., S. Marques, T. Kohler, J. L. Ramos, K. N. Timmis, and V. deLorenzo. 1994. Cross talk between catabolic pathways in Pseudomonas putida: XylS- dependent and-independent activation of the TOL meta operon requires the same *cis*-acting sequences within the Pm promoter. J. Bacteriol. **176:**5578-5582.

Klose, K. E., D. E. Weiss, and S. Kustu. 1993. Glutamate at the site of phosphorylation of nitrogen regulatory protein NtrC mimics aspartyl-phosphate and activates the protein. J. Mol. Biol. **1:**67-78.

Kok, R. G., V. M. Christoffels, B. Vosman, and K. J. Hellingwerf. 1993. Growth-phase dependent expression of the lipolytic system of *Acinetobacter calcoaceticus* BD413: cloning of a gene encoding one of the esterases. J. Gen. Microbiol. **139:**2329-2342.

Kok, R. G., J. J. van Thor, I. M. Nugteren-Roodzant, M. B. W. Brouwer, M. R. Egmond, C. B. Nudel, B. Vosman, and K. J. Hellingwerf. 1995. Characterization of the extracellular lipase, LipA, of *Acinetobacter calcoaceticus* BD413 and sequence analysis of the cloned structural gene. Mol. Microbiol. **15:**803-818.

Kok, R. G., D. A. D'Argenio, and L. N. Ornston. 1997. Combining localized PCR mutagenesis and natural transformation in direct genetic analysis of a transcriptional regulator gene, *pobR*. J. Bacteriol. **179:**4270-4276.

Kok, R. G, D. A. D'Argenio, and L. N. Ornston. 1998. Mutation analysis of PobR and PcaU, closely related transcriptional activators in *Acinetobacter*. J. Bacteriol. **180:**5058-5069.

Kokotek, W., and W. Lotz. 1989. Construction of a lacZ-kanamycin-resistance cassette, useful for site-directed mutagenesis and as a promoter probe. Gene. **84:**467-471.

Kowalchuk, G. A., G. B. Hartnett, A. Benson, J. E. Houghton, K-L. Ngai and L. N. Ornston. 1994. Contrasting patterns of evolutionary divergence within *Acinetobacter calcoaceticus pca* operon. Gene. **146:**23-30.

Kroos, L., and S. Cutting. 1994. Intercellular and intercompartmental communication during *Bacillus subtilis* sporulation. In P. J. Piggot, C. P. Moran Jr., and P. Youngman (eds.), Regulation of Bacterial Differentiation. American Society for Microbiology, Washington D.C. pp. 155-180.

Krüger, N., F. Berand Oppermann, H. Lorenzl, and A Steinbüchel. 1994. Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. J. Bacteriol. **176:**3614-3630.

Link, C., S. Eickernjäger, D. Porstendörfer, and B. Averhoff. 1998. Identification and characterization of a novel competence gene, *comC*, required for DNA binding and uptake in *Acinetobacter* sp. strain BD413. J Bacteriol. **180:**1592-1595.

Lonetto, M., M. Gribskov, and C. A. Gross. 1992. The sigma 70 family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843-3849.

Macintosh, R. W., and C. A. Fewson. 1988a. Benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II from *Acinetobacter calcoaceticus* - purification and preliminary characterization. Biochem. J. **250:**743-751.

Macintosh, R. W., and C. A. Fewson. 1988b. Benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II from *Acinetobacter calcoaceticus* - substrate specificities and inhibition studies. Biochem. J. **255:**653-661.

Marger, M. D., and M. H. Saier. 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport, and antiport. Trends Biochem. Sci. **18:**13-19.

Marques, S., A. Holtel, K. N. Timmis, and J. L. Ramos. 1994. Transcriptional induction kinetics from the promoters of the catabolic pathways of TOL plasmid pWW0 of *Pseudomonas putida* for metabolism of aromatics. J. Bacteriol. **176:**2517-2524.

Merrick, M. J. 1993. In a class of its own - the RNA polymerase sigma factor sigma 54 (sigma N). Mol. Microbiol. **5:**903-909.

Miller, J. H. 1972. In experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Mitsuhashi, K., M. Yamashita, Y. S. Hwan, F. Ihara, T. Nihira, and Y. Yamada. 1999. Purification and characterization of a novel extracellular lipase catalyzing hydrolysis of oleyl benzoate from *Acinetobacter nov* sp. strain KM109. Bioscience Biotechnology and Biochemistry. **63:**1959-1964.

Morett, E., and L. Segovia. 1993. The σ54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175:**6067-6074.

Murakami, S., A. Takashima, J. Takemoto, S. Takenaka, R. Shinke, and K. Aoki. 1999. Cloning and sequence analysis of two catechol-degrading gene clusters from the aniline-assimilating bacterium *Frateuria* species. Gene. **226:**189-198.

Nakai, C., K. Horiike, S. Kuramistu, H. Kagamiyama and H. Nozaki. 1990. Three isozymes of catechol 1,2 dioxygenase (pyrocatechase), αα, αβ, and ββ, from *Pseudomonas arvilla* C-1. J. Biol. Chem. **265:**660-665.

Neidle, E. L., and L. N. Ornston. 1987. Benzoate and muconate, structurally dissimilar metabolites, induce expression of *catA* in *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**414-415.

Neidle, E. L., M. K. Shapiro, and L. N. Ornston. 1987. Cloning and expression in *Escherichia coli* of *Acinetobacter calcoaceticus* genes for benzoate degradation. J. Bacteriol. **169:**5496-5503.

Neidle, E. L., C. Hartnett, and L. N. Ornston. 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. **17:**5410-5421.

Neidle, E. L., C. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama. 1991. Nucleotide-sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-Dioxygenase reveal evolutionary relationships among multi-component oxygenases. J. Bacteriol. **173:**5385-5395.

Neidle, E. L., C. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama. 1992.Cis-Diol Dehydrogenases encoded by the Tol pWW0 plasmid *xylL*-gene and the *Acinetobacter-calcoaceticus* Chromosomal *benD* gene are members of the short chain Alcohol Dehydrogenase Superfamily. Eur. J. Biochemistry. **1:**113-120.

Ng, L. C., E. O'Neill, and V. Shingler. 1996. Genetic evidence for interdomain regulation of the phenol responsive final sigma54-dependent activator DmpR. J. Biol. Chem. **271:**17281-17286.

Nguyen, L. H. and R. R. Burgess. 1997. Comparitive analysis of the interactions of *Eschericia coli* sigma S and sigma 70 RNA polymerase holoenzyme with the stationary-phase-specific bolAp1 promoter. Biochemistry. **7:**1748-1755.

Nichols, N. N., and C. S. Harwood. 1995. Repression of 4-hydroxybenzoate transport and degradation by benzoate: a new layer of regulatory control in the *Pseudomonas putida* β- ketoadipate pathway. J. Bacteriol. **177:**7033-7040.

Nielsen, K. M., K. Smalla, and J. D. vanElsas. 2000. Natural transformation of *Acinetobacter* sp. strain BD413 with cell lysates of *Acinetobacter* sp., *Pseudomonas fluorescens*, and *Burkholderia cepacia* in soil microcosms. Appl. Environ. Microbiol. **66:**206-211.

Nielsen, K. M., A. M. Bones, and J. vanElsas. 1997a. Induced natural transformation of *Acinetobacter calcoaceticus* in soil microcosms. Appl. Environ. Microbiol. **63:**4150-4157.

Nielsen, K. M., M. D. M. vanWeerelt, T. N. Berg, A. M. Bones, A. N. Hagler and J. D. vanElsas. 1997b. Natural transformation and availability of transforming DNA to *Acinetobacter calcoaceticus* in soil microcosms. Appl. Environ. Microbiol. **63:**1945-1952.

North, A. K., K. E. Klose, K. M. Steadman, and S. Kustu. 1993. Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. J. Bacteriol. **175:**4267-4273.

O'Neill, E., L. C. Ng, C. C. Sze, and V. Shingler. 1998. Aromatic ligand binding and intramolecular signalling of the phenol-responsive sigma54-dependent regulator DmpR. Mol. Microbiol. **28:**131-141.

Ohlendorf, D. H., A. M. Orville, and J. D Lipscomb. 1994. Structure of protacatechuate 3,4-dioxygenase from *Pseudomonas aeruginosa* at 2.15A resolution. J. Mol. Biol. **244:**586-608.

Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. D. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman. 1992. The α /β hydrolase fold. Protein Eng. **5:**197-211.

Ornston, L. N., and E. L. Neidle. 1991. Evolution of genes for the β-ketoadipate pathway in *Acinetobacter calcoaceticus,* pp. 201-237. In K. Tower, E. Bergogne-Berezin, and C. A. Fewson (ed.), The biology of *Acinetobacter.* Plenum Press, New York.

Osuna, J., X. Soberon, and E. Morett. 1997. A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. Protein Science. **6:**543-555.

Palmen, R., B. Vosman, R. Kok, J. R. Vanderzee, and K. J. Hellingwerf. 1992. Characterisation of transformation deficient mutants of *Acinetobacter calcoaceticus*. Mol. Microbiol. **6:**1747-1754.

Palmen, R., B. Vosman, P. Buijsman, C. Breek and K. J. Hellingwerf. 1993. Physiological characterization of natural transformation in *Acinetobacter calcoaceticus*. J. Gen. Microbiol. **139:**295- 305.

Palmen, R., P. Buijsman, and K. J. Hellingwerf. 1994. Physiological regulation of competence induction for natural transformation in *Acinetobacter calcoaceticus*. Archives of Microbiology. **162:**344-351.

Palmen, R., and K. J. Hellingwerf. 1997. Uptake and processing of DNA by *Acinetobacter calcoaceticus* - A review. Gene. **192:**179-190.

Parche, S., W. Geißdörfer, and W. Hillen. 1997. Identification and characterization of *xcpR* encoding a subunit of the general secretory pathway necessary for dodecane degradation in *Acinetobacter calcoaceticus* ADP1. J. Bacteriol. **179:**4631-4634.

Parke, D. 1995. Supraoperonic clustering of *pca* genes for catabolism of the phenolic compound protocatechuate in *Agrobacterium tumefaciens*. J. Bacteriol. **177:**3808-3817.

Parke, D., D. A. D'Argenio, and L. N. Ornston. 2000. Bacteria are not what they eat: that is why they are so diverse. J. Bacteriol. **182:**257-263.

Parsek, M. R., D. L. Shinabarger, R. K. Rothmel, and A. M. Chakrabarty. 1992. Roles of CatR and *cis,cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. J. Bacteriol. **174:**7798-7806.

Perez-Martin, J., F. Rojo, and V. de Lorenzo. 1994a. Promoters responsive to DNA bending – a common theme in prokaryotic. Microbiol. Rev. **58:**268-290.

Perez-Martin, J., K. Timmis, and V. de Lorenzo. 1994b. Co-regulation by bent DNA. Functional substitutions of the integration host factor site at sigma 54-dependant promoter Pu of the upper TOL operon by intrinsically curved sequences. J. Biol. Chem. **269:**22657-22667.

Perez-Martin, J., and V. de Lorenzo. 1995. The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. Proc. Natl. Acad. Sci. USA. **92:**9392-9396.

Porstendorfer, D., U. Drotschmann, and B. Averhoff. 1997. A novel competence gene, *comP*, is essential for natural transformation of *Acinetobacter* sp. strain BD413 Appl. Environ. Microbiol. **63:** 4150-4157.

Rainey, F. A., E. Lang, and E. Stakenbrandt. 1994. The phylogenic structure of the genus *Acinetobacter*. FEMS Microbiol. Lett. **124:**349-354.

Ramos, J. L., N. Mermod, and K. N. Timmis. 1987. Regulatroy circuits controlling transcription of TOL plasmid operon encoding *meta*-cleavage pathway for degrading of alkylbenzoates by *Pseudomonas*. Mol. Microbiol. **1:**293-300.

Rao, L., W. Ross, J. A. Appleman, T. Gaal, S. Leirmo, P. J. Schlax, M. T. Record Jr., and R. L. Gourse. 1994. Factor independent activation of *rrnB* P1. An "extended" promoter with an upstream element that dramatically increases promoter strength. J. Mol. Biol. **235:**1421-1435.

Record Jr., M. T., W. S. Reznikoff, M. L. Craig, K. L. McQuade, and P. J. Schlaz. 1996. *Escherichia coli* RNA polymerase (E70), promoters, and the kinetics of the steps of transcription initiation, p. 792-820. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. R. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.

Reddy, P. G., R. Allon, M. Mevarech, S. Mendelovitz, Y. Sato and D.L Gutnick. 1989. Cloning and expression in *E.coli* of an esterase coding gene from the oil degrading bacterium *Acinetobacter calcoaceticus* RAG-1. Gene. **76:**145-152.

Rippe, K., N. Mucke, and A. Schulz. 1998. Association states of transcriptional activator protein NtrC from *E.*coli determined by analytical ultracentrifugation. J. Mol. Biol. **278:**915-933.

Romero-Arroyo, C. E., M. A. Schell, G. L. Gaines, 3rd, and E. L. Neidle. 1995. *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus.* J. Bacteriol. **177:**5891-5898.

Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R. L. Gourse. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. Science. **262:**1407-1413.

Sabelnikov, A. G., B. Greenberg, and S. A. Lacks. 1995. An extended -10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae.* J. Mol. Biol. **250:**144-155.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schell, M. A, and E. F. Poser. 1989. Demonstration, characterization, and mutational analysis of NahR Protein-Binding to *nah* and *sal* Promoters. Journal of Bacteriology. **171:**837-846.

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

Schell, M. A., P. H. Brown, and S. Raju. 1990. Use of saturation mutagenesis to localise probable functional domains in the NahR protein, a LysR-type transcriptional activator. J. Biol. Chem. **265:**3844-3850.

Schell, M. A. 1993. Molecular-biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47:**597-626.

Schrag, J. D., and M. Cygler. 1997. Lipases and α/β-hydrolase fold. Methods Enzymol. **284:**85-107.

Segura, A., P. V. Bunz, D. A. D'Argenio, and L. N. Ornston. 1999. Genetic analysis of a chromosomal region containing *vanA* and *vanB*, genes required for conversion of either ferulate or vanillate to protocatechuate in *Acinetobacter*. J. Bacteriol. **181:**3494-3504.

Shanley, M. S., A. Harrison, R. E. Parales, G. Kowalchuk, D. J. Mitchell, and L. N. Ornston. 1994. Unusual g+c content and codon usage in *catIJF*, a segment of the *ben-cat* supraoperonic cluster in the *Acinetobacter calcoaceticus* chromosome. Gene. **138:**59-65.

Shaw, J. P., M. Reikik, F. Schwager and S. Harayama. 1993. Kinetic studies on benzyl alcohol dehydrogenase encoded by TOL plasmid pWW0. J. Biol. Chem. **268:**10842-10850.

Shingler, V., J. Powlowski, and U. Marklund. 1992. Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. J. Bacteriol. **174:**711-724.

Shingler, V., M. Bartilson, and T. Moore. 1993. Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators. J. Bacteriol. **175:**1596-1604.

Shingler, V., and T. Moore. 1994. Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp. strain CF600. J. Bacteriol. **176:**1555-1560.

Shingler, V. 1996. Signal sensing by sigma54-dependent regulators: derepression as a control mechanism. Mol. Microbiol. **19:**409-416.

Stanier, R. Y., N. J. Palleroni, and M. Doudoroff. 1966. The aerobic *Pseudomonads*: a taxonomic study. J. Gen. Microbiol. **43:**159-271.

Stewart, S. J., and C. A. Carlson. 1986. The biology of natural transformation. Ann. Rev. Microbiol. **40:**211-235.

Suzuki, M., and S. E. Brenner. 1995. Classification of multi-helical DNA-binding domains and application to predict the DBD structures of sigma-factor, LysR, OmpR/PhoB, CenP-B, Rap1, and XylS/Ada/AraC. FEBS Letters. **372:** 215-221.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weights, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22:**4673-4680.

Tirgari, S., and B. Mosely. 1980. Transformation in *Micrococcus radiodurans*: measurment of various parameters and evidence for multiple, independently segragating genomes per cell. J. Gen. Microbiol. **119:**287-296.

Travers, A., and R. R. Burgess. 1969. Cyclic re-use of the RNA polymerase sigma factor. Nature. **222:**537-540.

Trombe, M. C., C. Clave, and J. M. Manias. 1992. Calcium regulation of growth and differentiation in *Streptococcus pneumoniae*. J. Gen. Microbiol. **138:**77-84.

Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene. **19:**259-268.

Wang, Y., M. Rawlings, D. T. Gibson, D. Labbe, H. Bergeron, R. Brousseau, and P. C. Lau. 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. Mol. Gen. Genet. **246:**570-579.

Wang, X. P., and H. Weiner. 1995. Involvement of glutamate-268 in the active-site of human liver mitochondrial (class-2) aldehyde dehydrogenase as probed by site-directed mutagenesis. Biochemistry. **34:**237-243.

Wang, J. T., and J. D. Gralla. 1996. The transcription initiation pathway of sigma 54 mutants that bypass the enhancer protein requirement. Implications for the mechanism of activation. J. Biol. Chem. **51:**32707-32713.

Wang, Y. K., J. H. Lee, J. M. Brewer, and T. R. Hoover. 1997. A conserved region in the σ^{54}-dependant activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. Mol. Microbiol. **26:**373-386.

Weiss, D. S., J. Batut, K. E. Klose, J. Kreener, and S. Kustu. 1991. The phosphorylated form of the enhancer-binding protein NtrC has an ATPase activity that is essential for activation of transcription. Cell. **67:**155-167.

Williams, P. A., and K. Murray. 1974. Metabolism of Benzoate and the Methylbenzoates by *Pseudomonas arvilla* mt-2: Evidence for the existence of a TOL plasmid. J. Bacteriol. **120:**416-423.

Williams, P. A., and B. N. Zaba. 1985. EnzPack for Windows. Biosoft, Cambridge, U.K.

Williams, P. A. and L. E. Shaw. 1997. mucK, a gene in *Acinetobacter calcoaceticus* ADP1 (BD413), encodes the ability to grow on exogenous *cis,cis*-muconate as the sole carbon source. J. Bacteriol. **179:**5935-5942.

Wooton, J. C., and M. H. Drumond. 1989. The Q-linker of interdomain sequence found in bacterial multidomain regulatroy proteins. Protein Eng. **2:**535-543.

Wyman, C., I. Rombel, A. K. North, C. Bustamante and S. Kustu. 1997. Unusual oligomerisation required for activity of NtrC, a bacterial enhancer-binding protein. Science. **275:**1658-1661.

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

You, I.-S., D. Ghosal, and I. C. Gunsalus. 1988. Nucleotide sequence of plasmid NAH7 gene *nahR* and DNA binding of the *nahR* product. J. Bacteriol. **170:**5409-5415.

You, I.-S., D. Ghosal, and I. C. Gunsalus. 1991. Nucleotide sequence analysis of the *Pseudomonas putida* PpG7 salicylate hydroxylase gene (*nahG*) and its 3'-flanking region. Biochemistry. **30:**1635-1641.

JOURNAL OF BACTERIOLOGY, Aug. 1999, p. 4568–4575 Vol. 181, No. 15
0021-9193/99/$04.00+0

areABC Genes Determine the Catabolism of Aryl Esters in *Acinetobacter* sp. Strain ADP1

RHEINALLT M. JONES,[1] LAUREN S. COLLIER,[2] ELLEN L. NEIDLE,[2] AND PETER A. WILLIAMS[1]*

School of Biological Sciences, University of Wales Bangor, Bangor, Gwynedd LL57 2UW, Wales, United Kingdom,[1] and Department of Microbiology, University of Georgia, Athens, Georgia 30602-2605[2]

Received 30 March 1999/Accepted 27 May 1999

***Acinetobacter* sp. strain ADP1 is able to grow on a range of esters of aromatic alcohols, converting them to the corresponding aromatic carboxylic acids by the sequential action of three inducible enzymes: an *areA*-encoded esterase, an *areB*-encoded benzyl alcohol dehydrogenase, and an *areC*-encoded benzaldehyde dehydrogenase. The *are* genes, adjacent to each other on the chromosome and transcribed in the order *areCBA*, were located 3.5 kbp upstream of *benK*. *benK*, encoding a permease implicated in benzoate uptake, is at one end of the *ben-cat* supraoperonic cluster for benzoate catabolism by the β-ketoadipate pathway. Two open reading frames which may encode a transcriptional regulator, *areR*, and a porin, *benP*, separate *benK* from *areC*. Each *are* gene was individually expressed to high specific activity in *Escherichia coli*. The relative activities against different substrates of the cloned enzymes were, within experimental error, identical to that of wild-type *Acinetobacter* sp. strain ADP1 grown on either benzyl acetate, benzyl alcohol, or 4-hydroxybenzyl alcohol as the carbon source. The substrate preferences of all three enzymes were broad, encompassing a range of substituted aromatic compounds and in the case of the AreA esterase, different carboxylic acids. The *areA*, *areB*, and *areC* genes were individually disrupted on the chromosome by insertion of a kanamycin resistance cassette, and the rates at which the resultant strains utilized substrates of the aryl ester catabolic pathway were severely reduced as determined by growth competitions between the mutant and wild-type strains.**

Acinetobacter sp. strain ADP1, also called BD413 (22), is a saprophytic soil bacterium that has become a focus of research on the organization and evolution of genes involved in aromatic compound degradation (19, 30). ADP1 can degrade a wide variety of aromatic compounds by converting them to either catechol (1,2-dihydroxybenzene) or protocatechuate (3,4-dihydroxybenzoate), compounds whose aromatic rings can be enzymatically cleaved. Following intradiol ring cleavage, metabolites are channelled into the tricarboxylic acid cycle by enzymes encoded by the *cat* (for catechol) and *pca* (for protocatechuate) genes. Collectively, this catabolic route is referred to as the β-ketoadipate pathway.

In strain ADP1, the chromosomal *cat* and *pca* genes form two distinct supraoperonic clusters of genes each containing functionally related genes involved in funnelling aromatic compounds into the central metabolism. For example, the *cat* genes are adjacent to the *ben* genes needed for the conversion of benzoate to catechol (26). Similarly, the *pca* genes are clustered with *pob* and *qui* genes involved in converting 4-hydroxybenzoate and quinate, respectively, to protocatechuate (11, 13). Adjacent to this region are genes needed for the initial degradation of ferulate (29). The *ben-cat* and *pob-qui-pca* clusters are separated on the single circular chromosome by approximately 270 kbp (18).

Two exceptions to this clustering have recently been noted. The *ant* genes, encoding the enzyme that converts anthranilate (2-aminobenzoate) to catechol, lie distant from either cluster (6), and the genes needed for the conversion of vanillate to protocatechuate lie in a fourth distinct region of the chromosome (18a, 29). Despite these observations, the organization, regulation, and integration of the full complement of ADP1 genes for aromatic compound catabolism remain to be fully characterized. In this study, we report the sequence and characterization of an approximately 7-kbp extension of the *ben-cat* supraoperonic cluster. Included in this region are three contiguous genes whose protein products are capable of channelling the aromatic nucleus of esters of aryl alcohols into the β-ketoadipate pathway. These genes and their products have not previously been reported in strain ADP1.

* Corresponding author. Mailing address: School of Biological Sciences, University of Wales Bangor, Bangor, Gwynedd LL57 2UW, Wales, United Kingdom. Phone: (44) 1248 382363. Fax: (44) 1248 370731. E-mail: P.A.Williams@bangor.ac.uk.

MATERIALS AND METHODS

Strains and plasmids. The plasmids and bacterial strains used in this study are listed in Table 1.

Chemicals and media. Aromatic substrates were obtained from Sigma-Aldrich Co. 2-Hydroxybenzyl acetate and 4-hydroxybenzyl acetate were gifts from A. Boyes, Department of Chemistry, University of Wales Bangor, Bangor, United Kingdom. Luria-Bertani medium (LB) (32) was used to cultivate bacteria unless noted otherwise. For growth on minimal medium, single carbon sources were added to the minimal salts medium (5) at the following concentrations: benzyl acetate, benzyl alcohol, benzaldehyde, and 4-hydroxybenzyl alcohol at 2.5 mM and succinate at 10 mM. Where appropriate, ampicillin at 100 μg/ml and kanamycin at 50 μg/ml for *Escherichia coli* and at 10 μg/ml for *Acinetobacter* were used.

DNA manipulations. Standard methods for DNA manipulations were used (32). Total DNA was prepared from *Acinetobacter* sp. strain ADP1 by the method of Ausubel et al. (4). Plasmids carrying *Acinetobacter* DNA were isolated from and maintained in *E. coli* host XL1-Blue MRF′ or DH5α (Table 1) unless otherwise noted. Plasmid DNA was prepared from *E. coli* by alkaline lysis miniprep (32) or by Qiaprep columns (Qiagen). DNA fragments were recovered from agarose gels with Qiaquick columns (Qiagen). Southern blots were prepared as described by Sambrook et al. (32), and hybridizations were performed with ECL direct labelling kit (Amersham) according to the manufacturer's instructions.

PCR amplification. PCR amplifications were performed in 50-μl reaction mixture volumes containing 10 ng of template DNA, 100 pmol of each primer, 2.5 nmol of each deoxynucleoside triphosphate, 300 nmol of $MgSO_4$, and 1 U of Vent polymerase (New England Biolabs) in the reaction buffer supplied by the manufacturer. In some reaction mixtures, 200 nmol of $MgCl_2$ and 1 U of *Taq* polymerase were used in lieu of 300 nmol of $MgSO_4$ and 1 U of Vent polymerase. The mixtures were subjected to a 4-min hot start at 94°C and then to 30 cycles, with 1 cycle consisting of 1 min at 94°C, 1 min at 56°C, and 2 min at 74°C.

Cloning of *Acinetobacter* sp. strain ADP1 DNA. DNA adjacent to the *ben-cat* cluster was isolated by using the chromosomal drug resistance cassette in *benM*

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype, phenotype, or description[a]	Reference or source
Strains		
Acinetobacter		
ADP1 (BD413)	Wild type	22
ACN73	*areR*::Km^r; transformation of ADP1 with pBAC87	This study
ADPW56	*areC*::Km^r; transformation of ADP1 with pADPW22	This study
ADPW57	*areB*::Km^r; transformation of ADP1 with pADPW26	This study
ADPW58	*areA*::Km^r; transformation of ADP1 with pADPW37	This study
ADPW68	*areB*::Km^r; transformation of ADP1 with pADPW75	This study
ISA36	*benM*::Sm^r Sp^r	9
E. coli		
DH5α	F$^-$ ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 deoR recA1 endA1 hsdR17* (r_K^- m_K^+) *phoA supE44* λ$^-$ *thi-1 gyrA96 relA1*	Gibco BRL
XL1-Blue MRF′	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lacI*^{q}*Z*ΔM15 Tn*10* (Tet^r)]	Stratagene
BL21(DE3)pLysS	F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *dcm gal* (DE3) pLysS Cm^r	Promega
Plasmids		
pET5a	Ap^r; T7 expression vector	Promega
pUC18 and pUC19	Ap^r; cloning vector	40
pUC4K	Ap^r Km^r; source plasmid for Km^r cassette	36
pUI1637	Ap^r Km^r; source plasmid for Km^r cassette	14
pBAC68	7.0-kbp *Bgl*II/*Nsi*I fragment cloned from ISA36 containing Sm^r Sp^r cassette and all of *benP* and part of *areR* in pUC19	8 and this study
pBAC78	0.86-kbp *Bgl*II/*Kpn*I fragment containing part of *areR* in pUC19	This study
pBAC87	pBAC78 with Km^r cassette excised from pUI1637 with *Eco*RV and cloned into *Sna*BI site in *areR*	This study
pBAC98	6.7-kbp *Nsi*I/*Kpn*I fragment cloned from ACN73 containing Km^r cassette, all of *areCB*, and part of *areR* and *areA*	This study
pADPW21	1.6-kbp *Bam*HI*/*Sac*I fragment containing *areC* in pUC18	This study
pADPW22	pADPW21 with Km^r cassette from pUC4K cloned into *Bgl*II site in *areC*	This study
pADPW25	1.4-kbp *Bam*HI*/*Eco*RI* fragment containing *areB* in pUC18	This study
pADPW26	pADPW25 with Km^r cassette from pUC4K cloned into *Bcl*I site in *areB*	This study
pADPW29	*Nde*I*/*Eco*RI* fragment containing *areC* in pET5a	This study
pADPW30	*Nde*I*/*Eco*RI* fragment containing *areB* in pET5a	This study
pADPW31	0.65-kbp *Sac*I*/*Sal*I* fragment containing part of *areA* in pUC18	This study
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and all of *areA* in pUC18	This study
pADPW33	3.6-kbp *Hin*dIII subclone of pADPW32 in pUC18	This study
pADPW36	1.2-kbp *Eco*RI* fragment containing *areA* in pUC18	This study
pADPW37	pADPW36 with Km^r cassette from pUC4K cloned into *Nsi*I site in *areA*	This study
pADPW40	*Nde*I*/*Eco*RI* fragment containing *areA* in pET5a	This study
pADPW72	1.0-kbp *Eco*RI*/*Hin*dIII* fragment containing part of *areB* in pUC18	This study
pADPW75	pADPW72 with Km^r cassette from pUI1637 cloned into *Sac*I site in *areB*	This study

[a] Restriction sites added by PCR are indicated by superscript asterisks.

of strain ISA36 (9). Chromosomal DNA from this strain was used to make a plasmid library in pUC19 in *E. coli*, and plasmid pBAC68 (Bg_2 to N_2 in Fig. 1) was selected by appropriate drug resistance. A pBAC68 subclone, pBAC76 (Bg_2 to K in Fig. 1), was disrupted by a Km^r cassette to form pBAC87. With previously described methods using the natural transformation system of strain ADP1 (18), the corresponding chromosomal region was replaced by the disrupted *Acinetobacter* DNA of pBAC87, thereby generating ACN73 (Table 1). A chromosomal fragment from ACN73 was similarly isolated on pBAC98 (N_1 to K in Fig. 1).

To clone the additional chromosomal DNA, the *areB* gene in pADPW25, subcloned from pBAC98, was disrupted by the insertion of a Km^r cassette from pUC4K (36), creating pADPW26 (*areB*::Km^r). pADPW26 was linearized and transformed into ADP1, creating ADPW57. Chromosomal DNA from ADPW57 was cut with *Sac*I, ligated to *Sac*I-cut pUC18, and transformed into *E. coli*. Selection for Km^r Ap^r transformants yielded plasmid pADPW32 with an 8.0-kbp insert. Southern hybridization and DNA sequencing confirmed that pBAC98 and pADPW32 contained overlapping chromosomal DNA fragments. For this purpose, pADPW33 was subcloned from pADPW32 for use as a DNA sequencing template. pADPW31 was designed for use as a hybridization probe for the 5′ end of *areA* by PCR amplification of pBAC98 with the primer sequences 5′-ATGT TATGTCGACTTTGGTTTC-3′ (forward) and 5′-GCATCATGAGGAGCTCG AGGGTAT-3′ (reverse). The *Sal*I and *Sac*I sites used for cloning the amplified DNA are underlined, and the bases altered from those in the chromosome are in boldface type.

Expression of *areC*, *areB*, and *areA* in *E. coli*. Pairs of oligonucleotide primers were designed to produce PCR fragments of each *are* gene with an *Nde*I site introduced at the putative ATG start codon of each reading frame and a second constructed restriction site upstream of the *Nde*I site to allow cloning into pUC18. A natural or constructed restriction site downstream of the gene was used to clone PCR-generated fragments of the expected size first into pUC18. In this way, the *areC* and *areB* genes were amplified from pBAC98, and *areA* was amplified from pADPW32 to create pADPW21, pADPW25, and pADPW36, respectively (Table 1). These three clones were sequenced on one strand to check that mutations had not been incorporated by the PCR. Fragments from each were then excised with *Nde*I and *Eco*RI, religated into the expression vector pET5a, and transformed into *E. coli* BL21(DE3)pLysS to produce plasmids pADPW40 (*areA*), pADPW30 (*areB*), and pADPW29 (*areC*). In the following list of the PCR primers used to make these plasmids, the *Nde*I site is italicized, the *Eco*RI or *Bam*HI restriction sites used for cloning into pUC18 are underlined, and the bases that differ from those in the wild-type sequence are in boldface type (roman or italics, as appropriate): *areB* (forward), 5′-ACAGGAT GGATCC*ATATG*ACAAAGTTTACC-3′; *areB* (reverse), 5′-TTCAATGAATT CCCCCGCTACATGAG-3′; *areA* (forward), 5′-AAAGAATTCTTAA*CATATG* TTACAGAGTTT-3′; and *areA* (reverse), 5′-GGCATGCAGGGAATTCGAAC CAGAGATA-3′. In the case of *areC*, the reverse primer was chosen from downstream of the native *Sac*I site (see Fig. 1): *areC* (forward), 5′-AGGACGT TACGGGATCC*ATATG*ACATTACT-3′; and *areC* (reverse), 5′-ATGCCCATC TGGATCTCCACCACTGAAGTT-3′.

The Are proteins encoded on the expression vector pADPW29 (*areC*), pADPW30 (*areB*), or pADPW40 (*areA*) were individually expressed in *E. coli* BL21(DE3)pLysS by growth in LB to an optical density at 600 nm of 0.6 and then

induced for 4 h by the addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed on a discontinuous gel in a Mini-PROTEAN II Electrophoresis Cell (Bio-Rad) according to the manufacturer's instructions.

Chromosomal disruption of *areA*, *areB*, and *areC* in *Acinetobacter* sp. strain ADP1. Strains ADPW58, ADPW68, and ADPW56 with chromosomal disruptions in *areA*, *areB*, and *areC*, respectively, were constructed by methods similar to that used for ADPW57 (see above). As a first step, pUC18-derived plasmids that each carried one of the *are* genes disrupted by a Km^r cassette were made. To interrupt *areA* on pADPW36, the pUC4K cassette was cloned into a unique *Nsi*I site, creating pADPW37. A plasmid-borne *areB* disruption was made after first creating pADPW72, which has a PCR-generated *Eco*RI/*Hin*dIII insert in pUC18. This 1.0-kbp fragment, flanking the *Sac*I site in *areB*, was amplified from pBAC98. Primer sequences, with the *Eco*RI and *Hin*dIII sites underlined and the altered bases in boldface type, were as follows: (forward) 5′-TTGAATTCGGCGCGACCCTTGAAATTGGAG-3′ and (reverse) 5′-TTGAAAGCTTCACCTGCACCAGTTTGTATGCC-3′. The central *Sac*I site in the resultant pADPW72 was used as the insertion site for a Km^r cassette from pUI1637 (14) to create pADPW75. The *areC* gene was disrupted by the insertion of the cassette from pUC4K (36) into the unique *Bgl*II site of pADPW21, creating pADPW22. Plasmids pADPW37, pADPW75, and pADPW22 were linearized by digestion with a restriction enzyme, and each was used to transform ADP1. Southern hybridization confirmed that in strains ADPW58 (*areA*::Km^r), ADPW68 (*areB*::Km^r), and ADPW56 (*areC*::Km^r), the altered plasmid-borne allele had replaced the corresponding chromosomal wild-type region.

Competition for growth substrate between wild-type and *are* mutant *Acinetobacter* strains. ADP1 and the gene knockout mutants ADPW56, ADPW58, and ADPW68 were grown overnight in minimal medium containing 10 mM succinate to a cell density of about 10^9 cells/ml. A 5-μl aliquot of a 10^{-2} dilution of the wild type and of each of the mutants was inoculated into four tubes containing 990 μl of minimal medium supplemented with either succinate, benzaldehyde, benzyl alcohol, or benzyl acetate as the sole carbon source, respectively. The mixed cultures were grown on each substrate at 37°C for between 12 and 14 generations. Dilutions of the mixed culture were plated on LB agar. A total of 100 colonies were transferred onto LB agar containing 10 μg of kanamycin per ml and onto LB agar as a positive control. The percentage of mutants in the culture was assessed as the number of Km^r colonies. A 10-μl aliquot of each mixed culture was further inoculated into 990 μl of minimal medium containing the corresponding substrate. These cultures were grown for an additional 12 to 14 generations after which the percentage of mutants was again determined.

Preparation of cell extracts. Cells were harvested by centrifugation, washed with 100 mM phosphate buffer (pH 7.4), and stored as pellets at −20°C. Cell extracts were prepared by suspending frozen pellets in ice-cold 100 mM phosphate buffer, pH 7.4, disrupting by a pass through a French pressure cell (SLM Instruments, Inc., Urbana, Ill.), and centrifuging at 120,000 × *g* for 30 min at 4°C. The supernatant was stored frozen as 1-ml aliquots at −20°C.

Enzyme assays. Benzaldehyde dehydrogenase (BZDH) activity was measured in 3-ml assay mixtures containing 100 mM bicine (pH 9.5), 2 mM NAD^+, and 100 μM substrate. The reaction was initiated with 10 μl of cell extract, and the production of NADH was monitored spectrophotometrically at 340 nm. All three hydroxybenzaldehydes absorb at 340 nm with the following estimated molar extinction coefficients: 2-hydroxybenzaldehyde, 5,000 mol^{-1} cm^{-1}; 3-hydroxybenzaldehyde, 1,500 mol^{-1} cm^{-1}; and 4-hydroxybenzaldehyde, 18,400 mol^{-1} cm^{-1}. The difference between these values and the molar extinction coefficient of NADH at 340 nm (6,220 mol^{-1} cm^{-1}) was used to calculate activity with these substrates.

Benzyl alcohol dehydrogenase (BADH) activity was measured in a linked assay in 3 ml containing 100 mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS) (pH 10.5), 2 mM NAD^+, 200 μM substrate, and 10 μl of undiluted extract of *E. coli*(pADPW29) overexpressing BZDH. The reaction was initiated with 10 μl of enzyme extract, and the production of NADH was monitored spectrophotometrically at 340 nm. The BZDH was added in excess such that the activity of BADH was rate limiting and the rate of reaction was linear with respect to BADH added. Calculated rates took into account that two NADH molecules are produced per benzyl alcohol molecule oxidized except in the case of 2-methylbenzyl alcohol, since 2-methylbenzaldehyde is not further oxidized by BZDH. A small correction was made to the rate of 2- and 3-hydroxybenzyl alcohol oxidation to allow for the absorbances of 2-hydroxybenzoate (molar extinction coefficient, 40 mol^{-1} cm^{-1}) and 3-hydroxybenzoate (250 mol^{-1} cm^{-1}) at 340 nm.

Benzyl esterase was assayed with 4-nitrophenyl ester substrates in 1-ml assay mixtures containing 980 μl of 100 mM phosphate buffer (pH 8) and 10 μl containing various amounts of 4-nitrophenyl ester in methanol, and the reaction was initiated by the addition of 10 μl of enzyme. The activity of the esterase with benzyl ester substrates was determined in a linked assay. The rate of increase of absorbance at 340 nm was measured in 1-ml assay mixtures containing 100 mM phosphate buffer (pH 8), 2 mM NAD^+, various amounts of substrate, and 10 μl of BZDH and BADH from extracts of *E. coli* containing pADPW29 and pADPW30, respectively. The reaction was initiated by the addition of esterase. The fact that the reaction rate was a linear function of esterase concentration verified that the esterase-catalyzed reaction was the rate-limiting step. The assay produces two NADH molecules per molecule of benzyl alcohol produced. This is taken into account when determining Michaelis constant (K_m) and maximum velocity (V_{max}) values.

Determination of kinetic parameters for benzyl esterase. To obtain K_m and V_{max} values, initial velocities were measured at several nonsaturating concentrations of each compound. Preliminary experiments were used to determine the approximate value of K_m. Subsequently, accurate rate determinations using from 7 to 10 different substrate concentrations spanning the approximate K_m value were performed. Initial velocities were analyzed by Direct Linear analysis using the program EnzPack which calculates the most probable value for the kinetic parameters and the range with 68% confidence (39). Each reaction velocity was determined in triplicate using two separate extract preparations. The concentration of stock solution of substrate was accurately determined enzymatically by making the substrate limiting in the assay, while other components were in excess and the ΔA_{340} corresponding to the total conversion of added substrate was determined.

Nucleotide sequencing and sequence analysis. DNA sequences were determined by primer walking of fragments cloned in pUC18 or pUC19. In some cases, sequencing primers (Promega) were used that recognize the pUC18 and pUC19 cloning vectors. Sequencing was done by Alta Biosciences (University of Birmingham, Birmingham, United Kingdom), MWG-Biotech Ltd. (Ebersberg, Germany), and the University of Georgia Molecular Genetics Instrumentation Facility. Searches of the GenBank database were carried out by the BLASTN and BLASTP programs from the National Center for Biotechnology Information, Bethesda, Md. (3). Sequence data were aligned and edited by using the Lasergene software package (DNAStar, Inc., Madison, Wis.). Amino acid sequence alignments were performed by ClustalW (PAM350 matrix) (35). The Wisconsin Genetics Computer Group programs were also used (10).

Nucleotide sequence accession number. The DNA sequence obtained in this study has been added to GenBank under accession no. AF150928.

RESULTS

Characterization of DNA adjacent to the *ben* genes. Plasmid pBAC68 (Fig. 1) was originally isolated during an investigation of *benK*, the gene for a transport protein involved in the uptake of benzoate, located at one end of the *ben-cat* supraoperonic cluster for benzoate utilization (8). Nucleotide sequencing of the region immediately upstream of *benK* revealed an open reading frame, *benP* (Fig. 1), whose predicted protein product was found to be similar to porins, some of which are associated with genes for aromatic compound degradation (Table 2). The close proximity of *benP* to *benK*, with only 37 nucleotides separating the predicted coding regions, indicates that they may be cotranscribed. This possibility, that *benP*, like *benK*, is part of a regulon that funnels aromatic compounds through the β-ketoadipate pathway (8), prompted further exploration of the adjacent genetic region.

As described in Materials and Methods and illustrated in Fig. 1, plasmids pBAC98 and pADPW32 were isolated and the DNA sequences were determined. An open reading frame, *areR*, was found upstream of and divergently transcribed from *benP*, which is homologous to genes encoding XylR-type transcriptional activators (Table 2) (25, 31). The results of sequence analyses were consistent with the possibility that the *areR*-encoded protein acts in a σ^{54}-dependent fashion to control the expression of the genes immediately downstream (Fig. 1).

Identification of the *areCBA* genes. Downstream of *areR*, three genes whose predicted protein products have homology to bacterial dehydrogenases active against benzaldehyde and benzyl alcohol and to esterases, respectively, were identified. The genes have been designated *areCBA* in the order of their transcription. From database searches, the deduced amino acid sequence of AreC showed the most similarity to XylC, benzaldehyde dehydrogenase II from *Acinetobacter calcoaceticus* NCIMB8250 (17), and to a lesser extent to the benzaldehyde dehydrogenase encoded by the *xylC* gene on the *Pseudomonas* TOL plasmid pWW0 (20). AreB similarly was most like the corresponding BADHs encoded by the *xylB* genes from the same sources, *A. calcoaceticus* NCIMB8250 (17) and the TOL plasmid (34). AreA is similar to a number of esterases, partic-

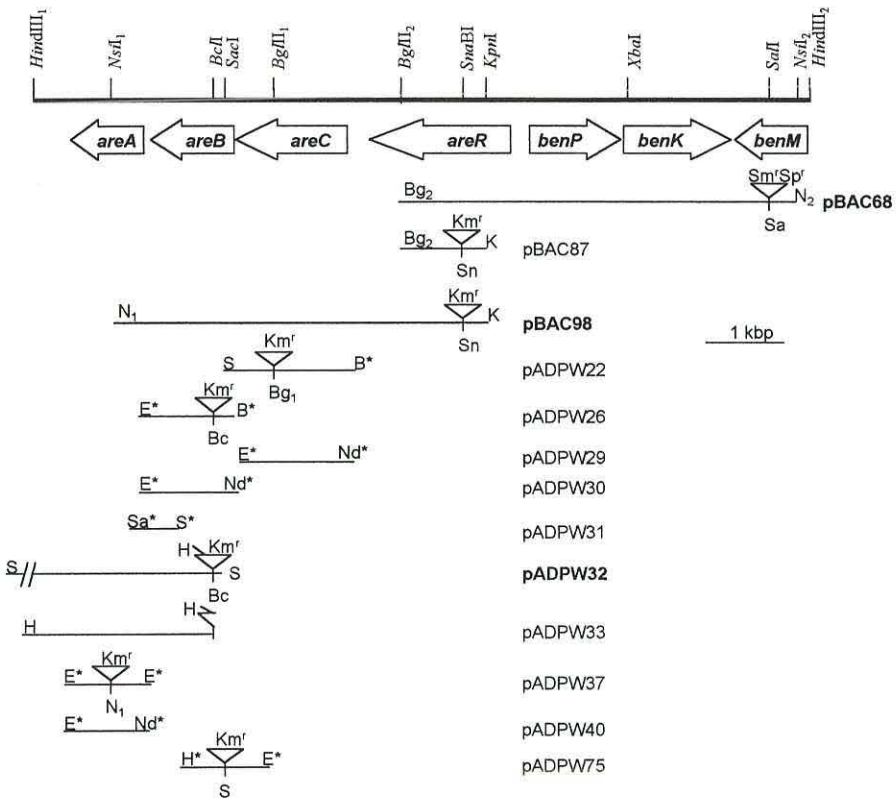


FIG. 1. Physical map of the *areCBA* genes and their locations relative to one end of the supraoperonic *ben-cat* cluster. The various inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 1. Plasmids shown in boldface type contain inserts that were cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with an asterisk were incorporated via PCR primers. The Km^r cassette insertions are not drawn to scale. Restriction site abbreviations: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Nsi*I; Nd, *Nde*I; S, *Sac*I; Sa, *Sal*I; Sn, *Sna*BI.

ularly those from *Archaeoglobus fulgidus* and *Pseudomonas putida* (7) (Table 2).

Expression of individual *are* genes. Each of the *are* genes was inserted into the expression vector pET5a such that gene expression would be optimally controlled by the T7 promoter on the plasmid when carried by an appropriate *E. coli* host. SDS-PAGE of extracts of the induced *E. coli* hosts showed high levels of expression of each protein with the predicted size (Fig. 2), whereas these proteins were not observed in either uninduced strains or those lacking the recombinant plasmids. In the case of induced AreB (lane 7), in addition to the expected band, there are also two smaller bands with molecular masses of 32 and 27 kDa: these could be degradation products of AreB but they are both prominent to a lesser extent in lanes 5 (induced AreA) and 9 (induced AreC).

Enzyme assays of cloned Are proteins. The AreC protein, expressed from plasmid pADPW29 (Table 1), showed NAD^+-dependent BZDH activity against benzaldehyde and all the methyl- and hydroxy-substituted benzaldehydes except for 2-methylbenzaldehyde, against which there was no measurable

TABLE 2. *Acinetobacter* sp. strain ADP1 genes and gene products

Gene designation	Putative function of gene product	Size of gene (bp)	% A+T	Size of gene product in:		Most similar gene products (species) (% amino acid identity/% amino acid similarity) (accession no.) (reference)
				Residues	kDa	
benP	Porin	1,146	61	381	43.8	PhaK (*P. putida*) (40/49) (AF029714)[a] (27) OprD2 (*Pseudomonas aeruginosa*) (35/43) (X63152)[a]
areR	Regulatory protein	1,803	60	600	68.4	AcoR (*P. putida*) (25/48) (3688510) AcoR (*Clostridium magnum*) (21/43) (472325) (24)
areC	BZDH	1,455	53.2	484	51.9	XylC (*A. calcoaceticus*) (77/88) (1408293) (19) XylC (*P. putida*) (44/64) (D63341)
areB	BADH	1,116	53.1	371	38.9	XylB (*A. calcoaceticus*) (82/91) (1408294) (19) XylB (*P. putida*) (56/70) (D63341)
areA	Benzyl esterase	981	63	326	37.1	Esterase (*Archaeoglobus fulgidus*) (31/48) (2648837) Esterase (*P. putida*) (28/50) (2853612) (7)

[a] In a selective alignment of 150 residues.

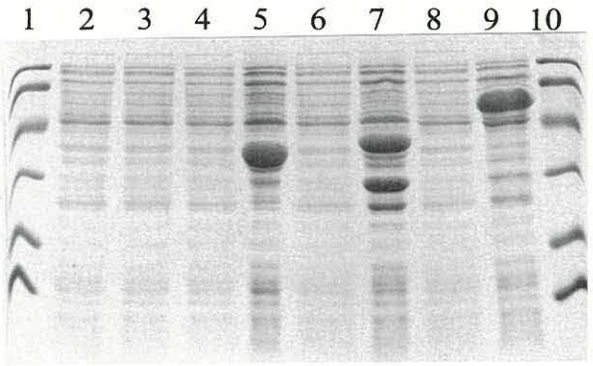


FIG. 2. SDS-PAGE of overexpressed Are proteins. The lanes contain lysates from *E. coli* BL21(DE3)pLysS (lane 2) on carrying the following plasmids: pET5a (lane 3), pADPW40 (lane 4), pADPW40 (lane 5), pADPW30 (lane 6), pADPW30 (lane 7), pADPW29 (lane 8), or pADPW29 (lane 9). The proteins were induced (lanes 5, 7, and 9) or not induced (lanes 4, 6, and 8) with IPTG. Lanes 1 and 10 contain the following molecular mass standards (values in kilodaltons): A, 97.4; B, 66.2; C, 45.0; D, 31.0; E, 21.5; and F, 14.5. The estimated molecular masses for the overexpressed bands are as follows: AreA (lane 5), 37 kDa; AreB (lane 7), 39 kDa with minor components at 32 and 27 kDa; and AreC (lane 9), 52 kDa.

activity (Table 3). The best substrate tested was 2-hydroxybenzaldehyde (salicylaldehyde), which was oxidized at a rate sixfold greater than was benzaldehyde or 3-hydroxybenzaldehyde.

AreB, expressed from pADPW30 (Table 4), was assayed for NAD^+-dependent BADH activity in the presence of excess AreC with compensation made for the twofold production of NADH in all cases except 2-methylbenzyl alcohol: by linking the alcohol dehydrogenase assay to BZDH, the initial rate of NADH production was linear over a much longer time period, expediting accurate rate determination. AreB showed significant activity against benzyl alcohol and the methyl- and hydroxy-substituted benzyl alcohols and, unlike AreC, was able to oxidize the 2-methyl analog at a rate comparable to those of the other substrates.

The esterase AreA, cloned on pADPW40 (Table 1), was assayed in two ways, against 4-nitrophenyl esters by monitoring the absorbance at 405 nm of the 4-nitrophenolate anion product and against benzyl esters by monitoring linkage to NADH production in the presence of excess amounts of both AreB and AreC. With both assays, the results (Table 5) were analyzed to determine the kinetic parameters K_m and V_{max}. For both sets of substrates, the value of K_m dropped as the chain length of the carboxylic acid moiety increased. The acetate esters were the poorest substrates in both classes, as can be judged by the lower relative specificity constants (V_{max}/K_m) resulting from both higher K_m and lower V_{max} values. The butyrate esters were the best substrates tested in both classes with higher affinities and higher V_{max} values.

Expression of *are* genes in strain ADP1. Wild-type ADP1 is able to utilize the following compounds as growth substrates: benzyl acetate, benzyl propionate, benzyl butyrate, 2-hydroxybenzyl acetate, 4-hydroxybenzyl acetate, benzyl alcohol, 2-hydroxybenzyl alcohol, 4-hydroxybenzyl alcohol, benzaldehyde, 2-hydroxybenzaldehyde, and 4-hydroxybenzaldehyde. It is unable to grow on benzyl benzoate or benzyl alcohols or benzaldehydes with 3-hydroxy or 2-, 3-, or 4-methyl substituents. To test whether the Are enzymes are expressed during growth on some of these substrates, assays of benzyl alcohol and benzaldehyde dehydrogenase activities against the same range of substrates were measured after growth on benzyl acetate, benzyl alcohol, and 4-hydroxybenzyl alcohol (Tables 3 and 4). In all cases, the relative activities against different substrates were essentially identical to those of the overexpressed AreC and AreB activities, although the induced specific activities in ADP1 were lower than in the recombinant *E. coli* by factors of 10 and 2.5, respectively. AreB and AreC were induced to considerably higher specific activity when grown on benzyl acetate or benzyl alcohol than when grown on succinate (Tables 3 and 4); similarly AreA was induced >500-fold when grown on the aromatic substrates compared with growth on succinate (data not shown).

Insertional inactivation of *are* genes. The chromosomal copy of *areA*, *areB*, or *areC* was specifically disrupted by a Km^r cassette in *Acinetobacter* strains ADPW58, ADPW68, and ADPW56, respectively (Table 1). These strains were tested for plate growth on benzyl acetate, benzyl alcohol, and benzaldehyde. In ADPW56 (*areC*::Km^r), the rate of growth on all three substrates was severely reduced, but not completely eliminated, and patches took from 2 to 3 days to achieve significant size compared with overnight growth for ADP1. ADPW68 (*areB*::Km^r) showed a reduced growth rate on benzyl acetate and benzyl alcohol but appeared to grow normally on benzaldehyde. The growth rate of ADPW58 (*areA*::Km^r) was reduced only on benzyl acetate. The results were therefore consistent with disruption of a pathway involving sequential action of AreA, AreB, and AreC.

To confirm these qualitative assessments of growth rates, growth competition experiments in liquid medium on all three substrates were set up between each of the knockout mutants and ADP1. In these assays, the initial cell densities of wild type and mutants were set to be approximately equal and the proportion of the mutants during two sequential subcultures each of about 12 to 14 generations was determined by plating out serial dilutions and determining the percentage of Km^r colonies. In all cases where the plate growth rate appeared to have been reduced, the percentage of mutant colonies dropped from approximately 50 to below 6 after the two subcultures.

TABLE 3. Relative activities of AreC BZDH against different benzaldehydes in crude extracts of cells grown on different substrates

Strain	Growth substrate	Relative activity[a] (sp act [U/mg of protein]) against benzaldehyde with substituent:						
		None	2-Methyl	3-Methyl	4-Methyl	2-OH	3-OH	4-OH
E. coli (pADPW29)	LB + Amp[b] + IPTG	100 (3.0)	<1	67	29	600	102	80
ADP1	Benzyl alcohol	100 (0.29)	<1	80	33	600	101	117
ADP1	Benzyl acetate	100 (0.16)	<1	80	29	600	101	96
ADP1	4-OH benzyl alcohol	100 (0.27)	<1	80	28	600	105	107
ADP1	Succinate	(<0.001)	ND	ND	ND	ND	ND	ND

[a] Activity relative to benzaldehyde, which is set at 100. All reaction rates upon which the relative activities were based were the averages of three measurements, none of which varied by >5%. ND, not determined.

[b] Amp, ampicillin.

TABLE 4. Relative activities of AreB BADH against various benzyl alcohols in crude extracts of cells grown on different substrates

Strain	Growth substrate	Relative activity[a] (sp act [U/mg of protein]) against benzyl alcohol with substituent:						
		None	2-Methyl	3-Methyl	4-Methyl	2-OH	3-OH	4-OH
E. coli (pADPW30)	LB + Amp[b] + IPTG	100 (2.3)	78	105	107	47	41	18
ADP1	Benzyl alcohol	100 (0.94)	62	99	97	42	32	14
ADP1	Benzyl acetate	100 (0.90)	63	103	106	43	33	11
ADP1	4-OH benzyl alcohol	100 (0.92)	57	106	110	50	51	14
ADP1	Succinate	(<0.001)	ND	ND	ND	ND	ND	ND

[a] Activity relative to benzyl alcohol, which is set at 100. All reaction rates upon which the relative activities are based were the averages of three measurements none of which varied by >5%. ND, not determined.
[b] Amp, ampicillin.

This was the case for ADPW56 on all three substrates, for ADPW68 on benzyl acetate and benzyl alcohol and for ADPW58 on benzyl acetate. Neither ADPW68 nor ADPW58 was outcompeted by ADP1 during growth on benzaldehyde nor was the latter outcompeted during growth on benzyl alcohol. Control experiments with succinate as the sole carbon source showed no outcompetition of the mutants by ADP1.

DISCUSSION

Contribution of *areABC* to the catabolic breadth of *Acinetobacter* sp. strain ADP1. The *areABC* genes of ADP1, which were identified and characterized in this study, encode enzymes for the sequential conversions of the esters of benzyl alcohol to benzyl alcohol, benzaldehyde, and benzoic acid (Fig. 3). The *areABC* genes can thereby funnel metabolites into the catechol branch of the β-ketoadipate pathway encoded by the chromosomally adjacent *ben* and *cat* genes. The substrate preferences of the Are proteins were tested after individually overexpressing *areA*, *-B*, and *-C* in an *E. coli* host. The resultant substrate profiles of AreB and AreC, consistent with the proposed enzyme functions, matched those of the wild-type strain grown on benzyl alcohol, benzyl acetate, or 4-OH benzyl alcohol. Although the specific inactivation of each gene on the *Acinetobacter* chromosome resulted in slow growth on benzyl acetate, it did not completely eliminate the ability to utilize this carbon source. It is likely that other esterases and dehydrogenases within the cell can partially complement the defective *are* genes. Nevertheless, the catabolic deficiencies of the *are* mutants were clearly demonstrated by the rapid ability of the wild-type strain to outcompete each mutant during growth on benzyl acetate.

A number of results suggest that the AreCBA proteins are responsible for the catabolism of a wider range of esters than just benzyl esters. First, the substrate ranges of all three enzymes are broad. Both dehydrogenases are able to utilize 2- and 4-hydroxy-substituted substrates, producing salicylate and 4-hydroxybenzoate, respectively. Both of these carboxylic acids are growth substrates for ADP1 and are also metabolized via the β-ketoadipate pathway. Salicylate is converted by salicylate 1-hydroxylase to catechol and thence dissimilated by the *cat* gene products, whereas 4-hydroxybenzoate is dissimilated by the *pca-pob* gene products. Growth of ADP1 on 4-hydroxybenzyl alcohol induces both BZDH and BADH activities with the same substrate preference as when grown upon benzyl alcohol or benzyl acetate, indicating that the same genes are expressed and are functional during growth on both substrates. Moreover, the chromosomal disruption of either *areA*, *-B*, or *-C* severely reduces the wild-type ability to grow on the 2-OH- and 4-OH-substituted benzyl acetates as carbon sources.

Given the wide range of natural aromatic acids which ADP1 can utilize (29), ADP1 may degrade a much wider range of aryl esters than those that were tested here. This possibility is difficult to test, since such compounds are not currently commercially available. It is also noteworthy that the esterase itself is nonspecific as far as the acid moiety is concerned and is able to hydrolyze 4-nitrophenyl esters of alkanoic acids with chain lengths of up to 16 carbons (data not presented). Esterase activity on a novel substrate might produce an aromatic alcohol that is not further metabolized, and yet the alkanoic acid itself could serve as a carbon source. In the natural environment, the Are enzymes may therefore be responsible for supporting growth on a wide range of substrates.

Factors beyond substrate specificity clearly contribute to carbon source utilization as well. For example, benzyl benzoate acts as a substrate for AreA to produce two products, benzoic acid and benzyl alcohol, both of which are excellent growth substrates. Despite this, benzyl benzoate is not itself a growth substrate, perhaps because of toxicity or blocks in uptake or regulation.

Comparisons of the Are proteins with similar enzymes. The deduced amino acid sequence of AreB shows all the characteristics of group 1 long-chain zinc-dependent alcohol dehydrogenases (ADH) as typified by horse liver ADH (12). AreB

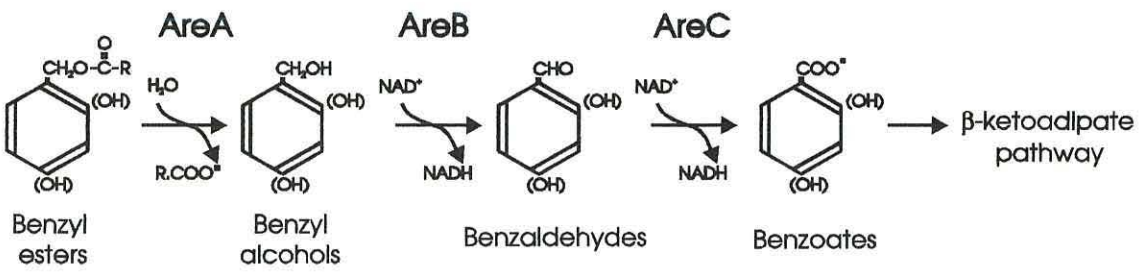


FIG. 3. Proposed pathway for catabolism of aryl esters by *Acinetobacter* sp. strain ADP1.

TABLE 5. Kinetic parameters for AreA benzyl esterase

Substrate	K_m (μM) (range)[a]	V_{max} (μmol/min/mg) (range)[a]	V_{max}/K_m
4-Nitrophenyl acetate	2,410 (2,260–2,720)	220 (207–236)	0.11
4-Nitrophenyl butyrate	61 (53–68)	960 (912–1,020)	16
4-Nitrophenyl hexanoate	11 (10.7–12.3)	430 (424–443)	39
Benzyl acetate	227 (213–238)	690 (670–712)	3
Benzyl propionate	67 (63–70)	2,090 (2,020–2,140)	31
Benzyl butyrate	42.4 (40–45)	2,050 (2,010–2,090)	47
Benzyl benzoate	47 (43.6–49)	1,100 (1,070–1,130)	23

[a] The range in parentheses after each value specifies the 68% confidence limits.

has the conserved residues which bind the catalytic Zn^{2+} (Cys44, His65, and Cys173) and the four Cys residues which bind the structural Zn^{2+} (at positions 94, 97, 100, and 108). It also contains the characteristic motif GHEXXGXXXXXG XXV from residues 64 to 78. Three bacterial sequences that are similar to AreB include two BADHs, both designated XylB, from *A. calcoaceticus* NCIMB8250 (17) and from the TOL plasmid of *P. putida* (34), and the XylW protein of unknown function from the TOL plasmid (38). In AreB, as in each of these bacterial sequences, there is a replacement of the His51 (as numbered in the horse liver enzyme) by a hydrophobic residue, in this case Ile49. This His is conserved in almost all of the long-chain zinc-dependent ADHs and is believed to act as a general base during the catalytic reaction. Inoue et al. (21), however, have suggested that in the TOL XylB an adjacent His residue takes this role, and in AreB there is an equivalent adjacent residue (His46).

AreC has an amino acid sequence which is typical of a large family of NAD^+-linked aldehyde dehydrogenases (class II). In particular, it contains the consensus sequence FTGSTXVG (residues 228 to 235, FTGSTQVG). Also conserved are the residues Cys285 and Glu251, homologous to Cys302 and Glu268 in the human liver mitochondrial aldehyde dehydrogenase that have been shown to be involved in the catalytic reaction (15, 37).

AreA appears to be a member of the α/β hydrolase fold family of enzymes (2, 28, 33). Despite a characteristic GXSXG consensus sequence, many members of this family display relatively low primary sequence conservation and are identified mainly by a method based on predicted structural similarities (2). In the case of AreA, the primary sequence can be aligned well with two other *Acinetobacter* esterases in this family. In the *Acinetobacter lwoffii* RAG-1 esterase, the residues that comprise the catalytic triad were suggested to be Ser149, either Asp196 or Glu244, and either His265, His274, or His298 (1). Three of these residues, Ser149, Glu244, and His274, align with those in AreA, indicating that in AreA, Ser166, Asp266, and His296 may form the catalytic triad. These AreA residues also align, respectively, with Ser201, Asp295, and His325 of a carboxylesterase of *Acinetobacter* sp. strain BD413 (ADP1) (23).

Genetic organization and regulation. The *areCBA* genes are all transcribed in the same direction, raising the possibility that these genes are cotranscribed. However, the order of the genes is opposite to that in which the gene products are needed, making it difficult to draw conclusions about whether the chromosomal insertion of Km^r cassettes in the *are* genes had polar effects on gene expression. For example, insertion of the cassette in the upstream gene *areC*, encoding BZDH activity, caused reduced growth on all the ester, alcohol, and aldehyde substrates, as would be expected by a polar effect of the inserted cassette on expression of the downstream genes. Nevertheless, this phenotype would also be expected if the genes were independently expressed, since the AreC gene product is the last of the three Are enzymes used in the formation of benzoic acids (Fig. 3). However, the relatively short intergenic regions between *areC* and *-B* (27 bp) and between *areB* and *-A* (83 bp) suggest coexpression of *areCBA* as an operon.

A sequence upstream of *areC* indicates that *areC* is at the start of a transcriptional unit. From bases −123 to −109 upstream of the start codon of *areC*, there is a possible σ^{54}-dependent (−12, −24) promoter sequence (**TGG**CAC-N_5-CT **GC**) that is well matched to the consensus sequence T**GG**CA C-N_5-TT**GC**. This putative promoter is of particular interest, since the upstream gene, *areR*, appears to encode a member of the σ^{54}-dependent XylR transcriptional regulator family. The deduced AreR amino acid sequence contained the conserved regions found in members of this family that function in ATP binding, σ^{54} interaction, and DNA binding (16, 25, 31).

The *xylBC* genes of *A. calcoaceticus* NCIMB8250 (17) are even more similar to *areBC* of ADP1 than are *xylBC* of the TOL plasmid. The two *A. calcoaceticus* genes are in the same order as in ADP1, although no *areA* homolog has been identified downstream of these *xyl* genes. In the published NCIMB8250 *xylCB* sequence, there are 200 bp downstream of *xylB* which show negligible homology to that downstream of *areB*. Furthermore, we have probed Southern blots of a plasmid containing 1 to 2 kb downstream of *xylBC* (a gift of D. J. Gillooly and C. A. Fewson) with an *areA* probe with negative results (data not shown). In the NCIMB8250 strain, the *xylCB* genes may therefore be involved in catabolism of only benzyl alcohol and not of benzyl esters.

The aromatic supraoperonic cluster in ADP1. The known size of the genetic region involved in the catabolism of compounds by the catechol branch of the β-ketoadipate pathway was extended by the identification of the *are* genes. The characterized region of this supraoperonic cluster now encompasses the region from *areA* at one end (Fig. 1) to *catD* (accession number AF00924) at the other, a total distance of 25 kbp. The linkage of genes encoding functionally related Are, Ben, and Cat proteins suggests that AreR and BenP also participate in aromatic compound degradation. The possibilities that *areR* encodes a XylR-like transcriptional regulator and that *benP* encodes a porin remain to be tested, however.

ACKNOWLEDGMENTS

This research was funded in part by a BBSRC research studentship (to R.M.J.) and by National Science Foundation grant MCB-9808784 (to E.L.N.).

REFERENCES

1. **Alon, R. N., and D. L. Gutnick.** 1993. Esterase from the oil-degrading *Acinetobacter lwoffii* RAG-1: sequence analysis and over-expression in *Escherichia coli*. FEMS Microbiol. Lett. **112**:275–280.

2. **Alon, R. N., L. Mirny, J. L. Sussman, and D. L. Gutnick.** 1995. Detection of α/β-hydrolase fold in the cell surface esterases of *Acinetobacter* species using an analysis of 3D profiles. FEBS Lett. **371:**231–235.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. E. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
5. **Bauchop, T., and S. R. Elsden.** 1960. The growth of microorganisms in relation to energy supply. J. Gen. Microbiol. **23:**457–469.
6. **Bundy, B. M., A. L. Campbell, and E. L. Neidle.** 1998. Similarities between the *antABC*-encoded anthranilate dioxygenase and the *benABC*-encoded benzoate dioxygenase of *Acinetobacter* sp. strain ADP1. J. Bacteriol. **180:** 4466–4474.
7. **Choo, D. W., T. Kurihara, T. Suzuki, K. Soda, and N. Esaki.** 1998. A cold-adapted lipase of an Alaskan psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and characterization. Appl. Environ. Microbiol. **64:**486–491.
8. **Collier, L. S., N. N. Nichols, and E. L. Neidle.** 1997. *benK* encodes a hydrophobic permease-like protein involved in benzoate degradation by *Acinetobacter* sp. strain ADP1. J. Bacteriol. **179:**5943–5946.
9. **Collier, L. S., G. L. Gaines III, and E. L. Neidle.** 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J. Bacteriol. **180:**2493–2501.
10. **Devereaux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
11. **Doten, R. C., K.-L. Ngai, D. J. Mitchell, and L. N. Ornston.** 1987. Cloning and genetic organization of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**3168–3174.
12. **Eklund, H., J.-P. Samama, L. Wallen, and C.-I. Branden.** 1981. Structure of a triclinic ternary complex of horse liver alcohol dehydrogenase. J. Mol. Biol. **146:**561–587.
13. **Elsemore, D. A., and L. N. Ornston.** 1994. The *pca-pob* supraoperonic cluster of *Acinetobacter calcoaceticus* contains *quiA*, the structural gene for quinate-shikimate dehydrogenase. J. Bacteriol. **176:**7659–7666.
14. **Eraso, J. M., and S. Kaplan.** 1994. *prrA*, a putative response regulator involved in oxygen regulation of photosynthetic gene expression in *Rhodobacter sphaeroides*. J. Bacteriol. **176:**32–43.
15. **Farres, J., T. T. Y. Wang, S. J. Cunningham, and H. Weiner.** 1995. Investigation of the active-site cysteine residue of rat-liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. Biochemistry **34:**2592–2598.
16. **Fernández, S., V. Shingler, and V. De Lorenzo.** 1994. Cross-regulation by XylR and DmpR activators of *Pseudomonas putida* suggests that transcriptional control of biodegradative operons evolves independently of catabolic genes. J. Bacteriol. **176:**5052–5058.
17. **Gillooly, D. J., A. G. S. Robertson, and C. A. Fewson.** 1998. Molecular characterization of benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II of *Acinetobacter calcoaceticus*. Biochem. J. **330:**1375–1381.
18. **Gralton, E. M., A. L. Campbell, and E. L. Neidle.** 1997. Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the *Acinetobacter* sp. strain ADP1 (BD413UE) chromosome. Microbiology **143:**1345–1357.

18a. **Gralton, E. M., and E. L. Neidle.** Unpublished observation.

19. **Harwood, C. S., and R. E. Parales.** 1996. The β-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50:**553–590.
20. **Inoue, J., J. P. Shaw, M. Rekik, and S. Harayama.** 1995. Overlapping substrate specificities of benzaldehyde dehydrogenase (the *xylC* gene product) and 2-hydroxymuconic semialdehyde dehydrogenase (the *xylG* gene product) encoded by TOL plasmid pWW0 of *Pseudomonas putida*. J. Bacteriol. **177:** 1196–1201.
21. **Inoue, J., N. Tomioka, A. Itai, and S. Harayama.** 1998. Proton transfer in benzyl alcohol dehydrogenase during catalysis: alternate proton-relay routes. Biochemistry **37:**3305–3311.
22. **Juni, E.** 1972. Interspecies transformation of *Acinetobacter*: genetic evidence for a ubiquitous genus. J. Bacteriol. **112:**917–931.
23. **Kok, R. G., V. M. Christoffels, B. Vosman, and K. J. Hellingwerf.** 1993. Growth-phase dependent expression of the lipolytic system of *Acinetobacter calcoaceticus* BD413: cloning of a gene encoding one of the esterases. J. Gen. Microbiol. **139:**2329–2342.
24. **Krüger, N., F. B. Oppermann, H. Lorenzl, and A. Steinbüchel.** 1994. Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. J. Bacteriol. **176:**3614–3630.
25. **Morett, E., and L. Segovia.** 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175:**6067–6074.
26. **Neidle, E. L., M. K. Shapiro, and L. N. Ornston.** 1987. Cloning and expression in *Escherichia coli* of *Acinetobacter calcoaceticus* genes for benzoate degradation. J. Bacteriol. **169:**5496–5503.
27. **Olivera, E. R., B. Minambres, B. Garcia, C. Muniz, M. A. Moreno, A. Ferrandez, E. Diaz, J. L. Garcia, and J. M. Luengo.** 1998. Molecular characterization of the phenylacetic acid catabolic pathway in *Pseudomonas putida* U: the phenylacetyl-CoA catabolon. Proc. Natl. Acad. Sci. USA **95:** 6419–6424.
28. **Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman.** 1992. The α/β hydrolase fold. Protein Eng. **5:**197–211.
29. **Ornston, L. N.** Personal communication.
30. **Ornston, L. N., and E. L. Neidle.** 1991. Evolution of genes for the β-ketoadipate pathway in *Acinetobacter calcoaceticus*, p. 201–237. *In* K. Towner, E. Bergogne-Berezin, and C. A. Fewson (ed.), The biology of *Acinetobacter*. Plenum Press, New York, N.Y.
31. **Osuna, J., X. Soberon, and E. Morett.** 1997. A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. Protein Sci. **6:**543–555.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Schrag, J. D., and M. Cygler.** 1997. Lipases and α/β hydrolase fold. Methods Enzymol. **284:**85–107.
34. **Shaw, J. P., M. Rekik, F. Shwager, and S. Harayama.** 1993. Kinetic studies on benzyl alcohol dehydrogenase encoded by TOL plasmid pWW0. J. Biol. Chem. **268:**10842–10850.
35. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weights, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
36. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
37. **Wang, X. P., and H. Weiner.** 1995. Involvement of glutamate-268 in the active-site of human liver mitochondrial (class-2) aldehyde dehydrogenase as probed by site-directed mutagenesis. Biochemistry **34:**237–243.
38. **Williams, P. A., L. M. Shaw, C. W. Pitt, and M. Vrecl.** 1997. *xylUW*, two genes at the start of the upper pathway operon of TOL plasmid pWW0, appear to play no essential part in determining its catabolic phenotype. Microbiology **143:**101–107.
39. **Williams, P. A., and B. N. Zaba.** 1997. EnzPack for Windows. Biosoft, Cambridge, United Kingdom.
40. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.

JOURNAL OF BACTERIOLOGY, Apr. 2000, p. 2018–2025 Vol. 182, No. 7
0021-9193/00/$04.00+0

sal Genes Determining the Catabolism of Salicylate Esters Are Part of a Supraoperonic Cluster of Catabolic Genes in *Acinetobacter* sp. Strain ADP1

RHEINALLT M. JONES, VASSILIS PAGMANTIDIS, AND PETER A. WILLIAMS*

School of Biological Sciences, University of Wales Bangor, Bangor, Gwynedd LL57 2UW, Wales, United Kingdom

Received 4 October 1999/Accepted 4 January 2000

A 5-kbp region upstream of the *are-ben-cat* genes was cloned from *Acinetobacter* sp. strain ADP1, extending the supraoperonic cluster of catabolic genes to 30 kbp. Four open reading frames, *salA*, *salR*, *salE*, and *salD*, were identified from the nucleotide sequence. Reverse transcription-PCR studies suggested that these open reading frames are organized into two convergent transcription units, *salAR* and *salDE*. The *salE* gene, encoding a protein of 239 residues, was ligated into expression vector pET5a. Its product, SalE, was shown to have esterase activity against short-chain alkyl esters of 4-nitrophenol but was also able to hydrolyze ethyl salicylate to ethanol and salicylic acid. A mutant of ADP1 with a Km^r cassette introduced into *salE* had lost the ability to utilize only ethyl and methyl salicylates of the esters tested as sole carbon sources, and no esterase activity against ethyl salicylate could be detected in cell extracts. SalE was induced during growth on ethyl salicylate but not during growth on salicylate itself. *salD* encoded a protein of undetermined function with homologies to the *Escherichia coli* FadL membrane protein, which is involved in facilitating fatty acid transport, and a number of other proteins detected during aromatic catabolism, which may also function in hydrocarbon transport or uptake processes. A Km^r cassette insertion in *salD* deleteriously affected cell growth and viability. The *salA* and *salR* gene products closely resemble two *Pseudomonas* proteins, NahG and NahR, respectively encoding salicylate hydroxylase and the LysR family regulator of both salicylate and naphthalene catabolism. *salA* was cloned into pUC18 together with *salR* and *salE*, and its gene product showed salicylate-inducible hydroxylase activity against a range of substituted salicylates, with the same relative specific activities as found in wild-type ADP1 grown on salicylate. Mutations involving insertion of Km^r cassettes into *salA* and *salR* eliminated expression of salicylate hydroxylase activity and the ability to grow on either salicylate or ethyl salicylate. Studies of mutants with disruptions of genes of the β-ketoadipate pathway with or without an additional *salE* mutation confirmed that ethyl salicylate and salicylate were channeled into the β-ketoadipate pathway at the level of catechol and thence dissimilated by the *cat* gene products. SalR appeared to regulate expression of *salA* but not *salE*.

Acinetobacter sp. strain ADP1 is capable of utilizing a range of aromatic compounds as sole sources of carbon. These compounds are dissimilated via the β-ketoadipate pathway either through benzoate and catechol or, alternatively, through 4-hydroxybenzoate and 3,4-dihydroxybenzoate (protocatechuate) (15, 22). Although the two branches have three terminal reactions in common, from β-ketoadipate enol lactone to the coenzyme A esters of succinate and acetate, each branch has its own genes and enzymes. The genes for the two branches are located in two supraoperonic clusters, the *ben-cat* cluster (7, 8, 20) (GenBank accession no. AF009224 and AF150928) and the *pob-qui-pca* cluster (9, 10) (GenBank accession no. L05770), which both extend for >20 kbp but are separated on the chromosome by approximately 270 kbp (13). Recently we reported that at one end of the *ben-cat* cluster is a group of four genes, *areA*, *-B*, *-C*, and *-R*, encoding an esterase, two dehydrogenases, and a regulator protein, that are responsible for the catabolism of alkanoate esters of benzyl alcohol, 2-hydroxybenzyl (salicyl) alcohol, and 4-hydroxybenzyl alcohol to benzoate, salicylate, and 4-hydroxybenzoate, respectively, and thus feeding these substrates into one of the two branches of the β-ketoadipate pathway (17). In this paper, we report a further extension, by about 5 kbp, of the *ben-cat* supraoperonic cluster beyond the *are* genes to include previously unreported genes responsible for the catabolism of alkyl salicylates through salicylic acid to catechol and thus channeling new substrates into the catechol branch of the pathway.

* Corresponding author. Mailing address: School of Biological Sciences, University of Wales Bangor, Bangor, Gwynedd LL57 2UW, Wales, United Kingdom. Phone: (44) 1248 382363. Fax: (44) 1248 370731. E-mail: P.A.Williams@bangor.ac.uk.

MATERIALS AND METHODS

Strains and plasmids. The plasmids and bacterial strains used in this study are listed in Table 1. Isolates ADPW1 and ADPW38 were spontaneous mutants selected by the procedure outlined in reference 31.

Chemicals and media. Aromatic substrates were obtained from Sigma-Aldrich Co. Ethyl salicylate was redistilled under reduced pressure to remove small amounts of contaminating ethanol. Luria-Bertani (LB) medium (23) was used for the cultivation of bacteria unless otherwise noted. For growth on defined carbon sources in liquid medium, the substrates were added to minimal salts medium (4) at the following concentrations: ethyl salicylate, sodium salicylate, benzyl acetate, benzoate, and 4-hydroxybenzoate at 2.5 mM and succinate at 10 mM. For growth on solid medium, a single nonvolatile carbon source (succinate, benzoate, salicylate, or 4-hydroxybenzoate) was added to minimal agar at the same concentrations, but volatile compounds (i.e., all of the esters) were presented in small tubes in the lids of inverted petri dishes containing minimal medium. When appropriate, ampicillin was incorporated at 100 μg/ml and kanamycin was added at 50 μg/ml for *Escherichia coli* and at 10 μg/ml for *Acinetobacter*.

DNA manipulations. Standard methods were used for DNA manipulations (23). Total DNA was prepared from *Acinetobacter* sp. strain ADP1 by the cetyltrimethylammonium bromide method (3). Plasmids carrying insertions of *Acinetobacter* DNA were isolated from and maintained in *E. coli* host strain XL1-Blue MRF′ or DH5α (Table 1) unless otherwise noted. Plasmid DNA was prepared from *E. coli* by the alkaline lysis miniprep method (23) or by using Qiaprep columns (Qiagen). DNA fragments were recovered from agarose gels by using Qiaquick columns (Qiagen). Southern blots were prepared as described by

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Description[a]	Reference or source
Acinetobacter strains		
ADP1 (BD413)	Wild type	18
ADPW1	*catA*	31
ADPW38	*ben*	31
ADP6	*pcaG*	12
ADPW57	*areB*::Km^r; obtained by transformation of ADP1 with pADPW26	17
ADPW67	*salA*::Km^r; obtained by transformation of ADP1 with pADPW44	This study
ADPW70	*salE*::Km^r; obtained by transformation of ADP1 with pADPW76	This study
ADPW72	*salR*::Km^r; obtained by transformation of ADP1 with pADPW79	This study
ADPW78	*salD*::Km^r; obtained by transformation of ADP1 with pADPW86	This study
ADPW86	*salE*::Km^r *catA*; obtained by transformation of ADPW1 with pADPW76	This study
ADPW87	*salE*::Km^r *ben*; obtained by transformation of ADPW38 with pADPW76	This study
ADPW88	*salE*::Km^r *pcaG*; obtained by transformation of ADP6 with pADPW76	This study
E. coli strains		
DH5α	F$^-$ ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17*(r_K^- m_K^+) *phoA supE44* λ$^-$ *thi-1 gyrA96 relA1*	Gibco BRL
XL1-Blue MRF′	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lacI*^{q}Z ΔM15 Tn*10* (Tet^r)]	Stratagene
BL21(DE3)pLysS	F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *dcm gal* (DE3) pLysS Cm^r	Promega
Plasmids		
pET5a	Ap^r; T7 expression vector	Promega
pUC18	Ap^r; cloning vector	32
pUC4K	Ap^r Km^r; source plasmid for Km^r cassette	28
pUI1637	Ap^r Km^r; source plasmid for Km^r cassette	11
pADPW32	8.0-kbp *Sac*I fragment cloned from ADPW57 containing Km^r cassette and the entire *areA* in pUC18	17
pADPW33	3.6-kbp *Hin*dIII subclone of pADPW32 in pUC18	17
pADPW34	4.1-kbp *Sac*I-*Xba*I subclone of pADPW32 in pUC18	This study
pADPW40	*Nde*I*-*Eco*RI* fragment containing *areA* in pET5a	17
pADPW41	1.6-kbp *Sac*I-*Hin*dIII subclone of pADPW34 in pUC18	This study
pADPW44	pADPW41 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salA*	This study
pADPW49	1.0-kbp *Eco*RI* fragment containing *salE* in pUC18	This study
pADPW70	*Nde*I*-*Eco*RI* fragment containing *salE* in pET5a	This study
pADPW76	pADPW49 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salE*	This study
pADPW78	1.0-kbp *Eco*RI*-*Hin*dIII* fragment containing part of *salR* in pUC18	This study
pADPW79	pADPW78 with Km^r cassette from pUI1637 cloned into *Cla*I site in *salR*	This study
pADPW82	1.2-kbp *Eco*RI* fragment containing part of *salD* in pUC18	This study
pADPW86	pADPW82 with Km^r cassette from pUC4K cloned into *Nsi*I site in *salD*	This study

[a] Asterisks denote restriction sites added by PCR.

Sambrook et al. (23), and hybridizations were carried out with an ECL direct labeling kit (Amersham) in accordance with the manufacturer's instructions.

PCR amplification. PCR amplifications were carried out in 50-μl volumes of reaction buffer (New England Biolabs) containing 10 ng of template DNA, 100 pmol of each primer, 2.5 nmol of each deoxynucleoside triphosphate, 300 nmol of $MgSO_4$, and 1 U of Vent polymerase (New England Biolabs). In some reactions, 200 nmol of $MgCl_2$ and 1 U of *Taq* polymerase were used in place of the $MgSO_4$ and Vent polymerase. The mixtures were subjected to a 4-min hot start at 94°C and then to 30 cycles of 1 min at 94°C, 1 min at 56°C, and 2 min at 74°C.

RT-PCR. Cells were grown on minimal medium containing either ethyl salicylate, salicylate, or succinate until they reached a density of about 10^8/ml. Total RNA was prepared from 10 ml of the culture by using RNeasy Mini columns (Qiagen), with elution in 50 μl of water. To remove any contaminating genomic DNA, the RNA was incubated with 1 U of RNase-free DNase (Promega) and 1 U of RNasin (Promega) in 40 mM Tris-HCl (pH 7.9) containing 10 mM NaCl (10 mM), $CaCl_2$ (10 mM), and $MgSO_4$ (6 mM) for 30 min at 37°C. The RNA was cleaned by passage through an RNeasy Mini column prior to use in reverse transcription (RT)-PCR. RT-PCR was carried out with an Access RT-PCR kit (Promega). Amplifications were carried out across the *salD-salE* and *salA-salR* intergenic regions by using primer pair EDf (5′-AGATTTAGGTATTCAGCA ATTCAGGGCAAAAGGTG-3′)-EDr (5′-AAGGCTCAGGCGTAAGCATCT TGTAAGTTTCCTC-3′) and ARf (5′-CCATGGACACGTGCGGTAGAC-3′)-ARr (5′-TTTTTGGTGCATGTGCTCGTAAGT-3′), respectively. PCRs were carried out in 50-μl volumes of reaction buffer (Promega) containing 0.5 μg of total RNA, 50 pmol of each primer, 50 μM (each) deoxynucleoside triphosphate, 1 mM $MgSO_4$, 5 U of avian myeloblastosis virus reverse transcriptase, and 5 U of *Tfl* DNA polymerase. After RT at 48°C for 1 h, the reaction mixtures were heated to 94°C for 2 min and subjected to 40 cycles of 30 s at 94°C, 1 min at 55°C, and 2 min at 68°C. Negative-control reactions, designed to ensure that residual genomic DNA was not amplified, were performed in the same way, except that the reverse transcriptase was omitted from the reaction mixtures.

Cloning of *Acinetobacter* sp. strain ADP1 DNA. DNA adjacent to *areABC* was isolated by using the chromosomal drug resistance cassette in *areB* of strain ADPW57 (17). A plasmid library in *E. coli* (pUC18) was made from chromosomal DNA of ADPW57, from which plasmid pADPW32 with an 8.0-kbp *Sac*I insert (Fig. 1) was selected by screening for Km^r Ap^r colonies. pADPW34 was constructed as a *Sac*I-to-*Xba*I subclone of pADPW32 (Fig. 1) so as to have sequence overlap with the previously sequenced plasmid pADPW33 (17). pADPW34 was used as the DNA sequencing template. Sequence alignment confirmed its overlap with pADPW33.

Expression of *salE* in E. coli. Oligonucleotide primers were designed to produce a PCR fragment of the *salE* gene with (i) an *Nde*I site introduced at the putative start site of the reading frame, (ii) a constructed *Eco*RI site upstream of the *Nde*I site, and (iii) an *Eco*RI site downstream of the gene. The PCR fragment generated from pADPW34 was cut with *Eco*RI and first ligated into *Eco*RI-cut pUC18 to create pADPW49 (Table 1). The insert was sequenced on one strand to ensure that mutations had not been incorporated during the PCR. A fragment was excised with *Nde*I and *Eco*RI, religated into the expression vector pET5a, and transformed into *E. coli* BL21(DE3)pLysS to produce plasmid pADPW70 (Fig. 1). Since *salE* contained a *Nde*I restriction site from bp 8 to 14, the forward primer was designed with a mutation in the ninth base pair (A→G) that destroyed the native *Nde*I site but did not change the amino acid encoded (Thr). The primers used were 5′-AGGA**GAATTC***ATATG*ATAAC**G**TATGTACTTG TTC-3′ (forward) and 5′-AGC**GAATTC**CTCGGATATGGTTGATTCAAAC-3′ (reverse) (the *Nde*I site is italicized, the *Eco*RI restriction sites used for cloning into pUC18 are underlined, and the bases which differ from the wild-type sequence are in boldface type). The SalE protein encoded on the expression vector

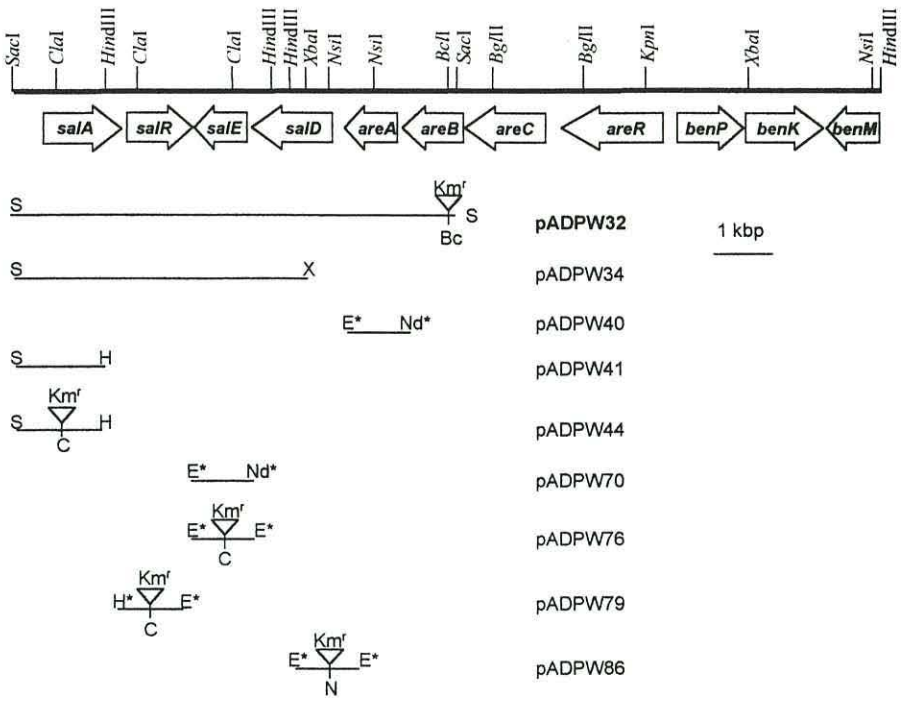


FIG. 1. Physical map of the *salA*, *salR*, *salE*, and *salD* genes and their location relative to *areABC* at the left-hand end (as drawn) of the supraoperonic *ben-cat* cluster. The inserts of the plasmids produced from cloning genomic DNA into vectors are specified in Table 1. pADPW32 was cloned directly from genomic DNA. All other plasmids were produced by PCR from genomic DNA (denoted by asterisks) or by subcloning from plasmids containing genomic DNA. Sites at the termini of the inserts marked with asterisks were incorporated via PCR primers. The Km^r cassette insertions are not to scale. The abbreviations for the restriction sites are as follows: Bc, *Bcl*I; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; N, *Nsi*I; Nd, *Nde*I; S, *Sac*I; and X, *Xba*I.

pADPW70 was expressed in *E. coli* BL21(DE3)pLysS by growth of the bacterium in LB medium to an optical density at 600 nm of 0.6 and subsequent induction for 4 h by addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out in a discontinuous gel in a Mini-PROTEAN II electrophoresis cell (Bio-Rad) in accordance with the manufacturer's instructions.

Chromosomal disruption of *salA*, *salR*, *salE*, and *salD* in *Acinetobacter* sp. strain ADP1. As a first step in the construction of gene knockouts, pUC18-derived plasmids carrying part or all of gene *salA*, *salR*, *salE*, or *salD* disrupted by a Km^r cassette were constructed. For *salA*, a *Sac*I-*Hin*dIII fragment of pADPW34 was cloned to create pADPW41 (Fig. 1). The Km^r cassette of pUI1637 (11) was cloned into a unique *Cla*I site of pADPW41, creating pADPW44. The *salE* gene was disrupted by the insertion of the cassette from pUI1637 into the unique *Cla*I site of pADPW49 (see above) to form pADPW76. Plasmid-borne *salR* disruption was achieved after first creating pADPW78, which has a PCR-generated *Eco*RI insert in pUC18. This 1.0-kbp fragment, internal to *salR* and flanking its *Cla*I site, was amplified from pADPW34. Primer sequences (with the *Eco*RI sites underlined and the altered bases in boldface) were as follows: 5′-TG**G**AATTCATGAACAGATCCGAAAAGAACG-3′ (forward) and 5′-CATG**AATTC**CCTGAGTATGCCCGGTA-3′ (reverse). The central *Cla*I site in the pADPW78 was used as the insertion site for the Km^r cassette from pUI1637 (11) to create pADPW79. Disruption of plasmid-borne *salD* was performed after first creating pADPW82, which has a PCR-generated *Eco*RI insert in pUC18. This 1.2-kbp fragment, flanking the *Nsi*I site in *salD*, was amplified from pADPW32. Primer sequences (with the *Eco*RI sites underlined and the altered bases in boldface) were as follows: 5′-AGGGGG**A**ATTCTGGCAG CAATCACTG-3′ (forward) and 5′-GGGCTG**G**AATTCCCAAGTACTACC TAT-3′ (reverse). The central *Nsi*I site in pADPW82 was used as the insertion site for a Km^r cassette from pUC4K (28) to create pADPW86. Plasmids pADPW44, pADPW76, pADPW79, and pADPW86 were linearized by digestion with an appropriate restriction enzyme, and each was used to transform ADP1 via natural transformation. Southern hybridization confirmed that in strains ADPW67 (*salA*::Km^r), ADPW70 (*salE*::Km^r), ADPW72 (*salR*::Km^r), and ADPW78 (*salD*::Km^r) the altered plasmid-borne allele had replaced the corresponding chromosomal wild-type region (data not shown).

Preparation of cell extracts. Cells were harvested by centrifugation, washed with 100 mM phosphate buffer (pH 7.4), and stored as pellets at −20°C. Cell extracts were prepared by disrupting frozen pellets, suspended in ice-cold 100 mM phosphate buffer (pH 7.4), with a French pressure cell (SLM Instruments, Inc., Urbana, Ill.) and centrifuging the broken cells at 120,000 × *g* for 30 min at 4°C. The supernatant was stored frozen as 1-ml portions at −20°C.

Transformation of metabolites. Transformation of ethyl salicylate into salicylate by cell extracts of SalE was monitored spectrophotometrically. The measuring cell contained 100 μM substrate, 100 μM Tris (pH 7.5), and 10 μl of cell extract, while the reference cell contained only buffer and enzyme.

Enzyme assays. Salicylate hydroxylase (SalA) activity was measured in 3-ml reaction mixtures containing 50 mM Tris (pH 7.5), 100 μM NADH, and 100 μM salicylate. The reaction was initiated by addition of 20 μl of cell extract, and the rate of oxidation of NADH was determined spectrophotometrically at 340 nm (extinction coefficient, 6,220 mol^{-1} cm^{-1}). Salicylate esterase (SalE) activity was assayed by spectrophotometric monitoring (405 nm) of the hydrolysis of 4-nitrophenyl ester substrates in 1-ml reaction mixtures containing 50 mM phosphate buffer (pH 8) and 2 mM 4-nitrophenyl ester. The 4-nitrophenyl esters with an aliphatic moiety of six or less carbon atoms were dissolved in methanol, and an aliquot was added to the assay mixture such that the final concentration of the ester was 2 mM. The 4-nitrophenyl esters with an aliphatic moiety of eight carbon atoms or longer were first dissolved in 2-propanol at 60°C and then added dropwise to 50 mM Tris-HCl (pH 8.0), prewarmed to 60°C, to a final ester concentration of 2 mM. The assay reactions were initiated by the addition of 10 μl of enzyme. The molar extinction coefficient of 4-nitrophenol was taken as 14,800 mol^{-1} cm^{-1}. The activity of the esterase with ethyl salicylate as the substrate was determined in a linked assay. The rate of increase of absorbance at 340 nm was measured in 1-ml reaction mixtures containing 100 mM phosphate buffer (pH 8), 2 mM NAD^+, 100 μM substrate, and 10 U of yeast alcohol dehydrogenase (Sigma-Aldrich Co.). The reaction was initiated by the addition of esterase. A linear response of rate to added esterase verified that the esterase-catalyzed reaction was the rate-limiting step. The assay produced 1 mol of NADH per mol of ethyl salicylate utilized.

Determination of kinetic parameters for salicylate esterase. To obtain K_m and maximum velocity (V_{max}) values, initial velocities were measured at several nonsaturating concentrations of each compound. Preliminary experiments determined the approximate value of K_m, and accurate rate determinations were then performed with from 7 to 10 different substrate concentrations spanning the approximate K_m value. Initial velocities were analyzed by direct linear analysis using the program EnzPack, which calculates the most probable values for the kinetic parameters with their 68% confidence limits (30). Each reaction velocity was determined in triplicate with two separate extract preparations. The concentration of the substrate stock solution was accurately determined enzymatically by making the substrate limiting in the assay while other components were in excess, and the change in absorbance at 405 nm, corresponding to the total conversion of added substrate, was determined.

TABLE 2. *Acinetobacter* sp. strain ADP1 genes and gene products

Gene designation	Putative function of gene product	Size of gene (bp)	% (A+T)	Size of gene product		Most-similar gene products (% amino acid identity/similarity)[a] [GenBank/Swissprot accession no.]
				Residues	kDa	
salD	Putative membrane protein	1,161	58	386	41.8	XylN (*P. putida* pWW0) (19%/36%) [D63341], TodX (*P. putida*) (21%/36%) [U18304]
salE	Salicylate esterase	720	54	239	27.0	Putative esterase/lipase 2 (*Mycoplasma pneumoniae*) (17%/34%) [P75311]
salR	Regulatory protein of LysR family	891	65	296	34.1	NahR (*P. stutzeri*) (33%/53%) [AF039534] NahR (*P. putida* NAH7) (31%/52%) [M22723]
salA	Salicylate hydroxylase	1,272	54	423	46.9	NahG (*P. putida*) (48%/65%) [X83926] NahG (*P. putida* NAH7) (47%/65%) [M60055]

[a] Measured over whole protein.

Nucleotide sequencing and sequence analysis. DNA sequences were determined by primer walking of fragments cloned in pUC18 by MWG-Biotech Ltd. (Ebersberg, Germany). Searches of the GenBank database were carried out with the BLASTN and BLASTP programs from the National Center for Biotechnology Information, Bethesda, Md. (2). Sequence data were aligned and edited by using the Lasergene software package (DNAStar, Inc., Madison, Wis.). Amino acid sequence alignments were performed with the program ClustalW (PAM350 matrix) (27).

Nucleotide sequence accession number. The DNA sequence obtained in this study has been added to the GenBank database (accession no. AF150928).

RESULTS

Analysis of nucleotide sequences of protein products. Analysis of the nucleotide sequence downstream of *areA* revealed the presence of four open reading frames (Table 2). Immediately downstream of *areA*, but separated by 195 bp and an inverted repeat, which could serve as a termination loop for *areA* expression (17), is a gene designated here as *salD*. SalD shows 13 to 17% amino acid sequence similarity to putative proteins encoded by other operons of aromatic catabolism in different bacteria, which include TodX (29), XylN (GenBank accession no. D63341), and TbuX (H.-Y. Kahng, A. M. Byrne, R. H. Olsen, and J. J. Kukor, submitted for publication), which have been suggested to be membrane proteins involved in the transport of hydrocarbons (J. J. Kukor, personal communication). Downstream of *salD* is a gene, which we have called *salE*, whose product shows homology to serine esterases of the α/β hydrolase family of enzymes (21, 26) but which is not closely related to the benzyl esterase AreA-encoding gene lying upstream of it (26% similar, 15% identical). Transcribed convergently toward *salE* is a gene, which we have called *salR*, apparently encoding a regulatory protein of the LysR family. The closest matches to SalR are two NahR proteins involved in naphthalene catabolism, one from *Pseudomonas stutzeri* (6) and one from the *Pseudomonas putida* NAH7 plasmid (25). The latter protein has a dual role as a positive regulator of two functionally related operons, for the conversion of naphthalene to salicylate and for the further conversion of salicylate to central metabolites via catechol and the subsequent extradiol (*meta*) cleavage pathway (24, 25, 33). The final gene in this cluster is *salA*, whose putative product's closest relatives have both been named NahG. These enzymes are salicylate hydroxylases, converting the salicylate produced from naphthalene to catechol. The genes encoding both enzymes head the salicylate catabolic operon on the NAH7 plasmid (34) and in *P. stutzeri* (6), respectively. However, it has been noted that ADP1 does not grow on naphthalene as a sole carbon source. The alignments of both *salA* and *salR* with their *Pseudomonas* homologues strongly indicate that unlike the *Pseudomonas* genes, they both have a GTG start codon.

Insertional inactivation of single genes. The chromosomal copies of *salA*, *salE*, *salR*, and *salD* were specifically disrupted, individually, by the insertion of a Km^r cassette into each of the genes in plasmid constructs (pADPW44, pADPW76, pADPW79, and pADPW86, respectively [Fig. 1]). The disrupted genes were introduced into ADP1, using the high frequency of natural transformation of which this strain is capable. Two of the resulting strains, ADPW67 (*salA*::Km^r) and ADPW72 (*salR*::Km^r), failed to grow on salicylate, unlike ADP1, which grows vigorously on salicylate overnight. By contrast, ADPW70 (*salE*::Km^r) grew on salicylate as well as did ADP1. ADPW78 (*salD*::Km^r) was a very unhealthy strain; it grew slowly compared with ADP1 even on succinate minimal medium and rich (LB) medium and lost viability when maintained on agar after 2 to 3 days. However, despite these limitations it, too, like ADPW70, grew on salicylate. Because SalE showed amino acid sequence homologies with other esterases, we compared the abilities of ADPW70 and ADP1 to grow on a range of esters containing aromatic components both as the alcohol and as the acid moiety, as well as a number of exclusively aliphatic esters. Thirteen esters (*n*-propyl acetate, benzyl acetate, *n*-butyl propionate, ethyl propionate, benzyl propionate, ethyl valerate, ethyl butyrate, benzyl butyrate, ethyl caproate, ethyl benzoate, *n*-propyl benzoate, *n*-butyl benzoate, and ethyl benzoylacetate) were growth substrates for both strains, and 6 (ethyl acetate, *n*-butyl acetate, *n*-butyl butyrate, benzyl benzoate, *n*-propyl cinnamate, and benzyl salicylate) were substrates for neither strain. However, ethyl salicylate and methyl salicylate supported growth of ADP1 but not ADPW70, suggesting that SalE is a hydrolase specific for these esters. Neither ADPW67, ADPW72, nor ADPW78 was able to grow on either of the two alkyl salicylates.

Expression of cloned *salE*. The *salE* gene was cloned into expression vector pET5a as plasmid pADPW70 with its start codon (ATG) located in the optimal position for expression. SDS-PAGE of the induced *E. coli* BL21(DE3)pLysS containing pADPW70 revealed a strong protein band with an expected molecular mass of 27 kDa (Fig. 2). The esterase activity in extracts of induced cells against a range of 4-nitrophenyl esters was measured and compared with the activity of the upstream benzyl esterase AreA (17) (Table 3). Whereas AreA shows a broad specificity for the alkanoate side chain, going up to C_{16}, SalE shows a much more restricted range, up to only C_6. The K_m values of SalE for 4-nitrophenyl acetate and 4-nitrophenyl butyrate were 106 and 77 μM, respectively, whereas the relative V_{max} dropped 10-fold, from 28 to 2.8 μmol/min/mg, for the same two substrates.

We attempted to set up a SalE assay using ethyl salicylate as a substrate by linkage to salicylate hydroxylase SalA overex-

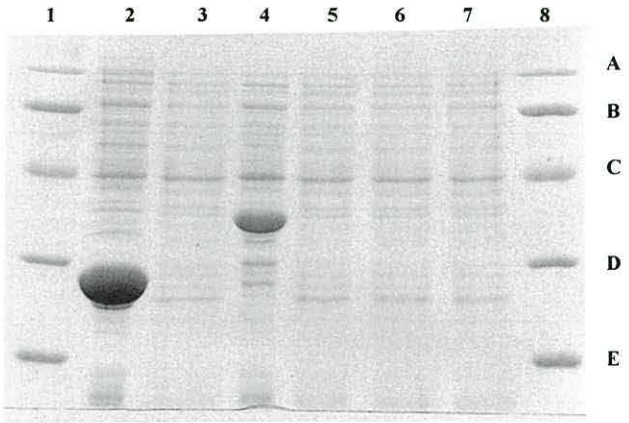


FIG. 2. SDS-PAGE of overexpressed SalE and AreA proteins. The lanes contain lysates from *E. coli* BL21(DE3)pLysS carrying the following plasmids, induced with IPTG (I) or uninduced (U): lane 2, pADPW70 (I); lane 3, pADPW70 (U); lane 4, pADPW40 (I); lane 5, pADPW40 (U); and lane 6, pET5a. Lane 1 contained molecular mass standards (A, 97.4 kDa; B, 66.2 kDa; C, 45.0 kDa; D, 31.0 kDa; and E, 21.5 kDa). The estimated molecular masses for the overexpressed bands are 27 kDa for SalE (lane 2) and 37 kDa for AreA (lane 4).

pressed from a pET5a-derived plasmid. Unfortunately, the SalA construct failed to show activity for as-yet-undiagnosed reasons. However, we did set up an alternative assay with ethyl salicylate as a substrate by linking the assay to yeast alcohol dehydrogenase, acting on the ethanol produced by SalE action. This was successfully carried out and showed that after IPTG induction the *E. coli*(pADPW70) exhibited high-level hydrolytic activity against ethyl salicylate (Table 4).

Expression of salicylate hydroxylase. Using the standard NADH-linked assay procedure, we were able to detect salicylate hydroxylase activity in a number of strains (Table 5). For ADP1, activity was not detectable in succinate-grown cells but was induced by growth on both salicylate and ethyl salicylate, and when grown on succinate in the presence of both salicylate and ethyl salicylate the activity was significantly induced, although at a lower level than in the absence of succinate. For both ADPW67 (*salA*::Km^r) and ADPW72 (*salR*::Km^r), no activity was detected when the cells were grown on succinate plus salicylate or, for the latter, succinate in the presence of ethyl salicylate. However, for ADPW70 (*salE*::Km^r) there was high-level induction when grown on salicylate and a low, but significant, level of induction when grown in the presence of ethyl salicylate, which it is unable to transform.

TABLE 3. Relative specific activities of *Acinetobacter* esterases against 4-nitrophenyl alkanoates

No. of carbons in alkyl chain	Relative specific activity[a]	
	SalE[b]	AreA[c]
2	450	1.9
4	100	100
6	0.5	87
8	<0.1	6.5
10	<0.1	1.4
12	<0.1	0.9
14	<0.1	0.2
16	<0.1	<0.1

[a] Relative to the activity against 4-nitrophenyl butyrate (C_4) (set at 100). All reaction rates were based on the average of three measurements, none of which varied by >5%.

[b] Measured in cell extracts of *E. coli* BL21(pADPW70). The specific activity against 4-nitrophenyl butyrate was 11.2 U/mg of protein.

[c] Measured in cell extracts of *E. coli* BL21(pADPW40). The specific activity against 4-nitrophenyl butyrate was 38.0 U/mg of protein.

TABLE 4. Specific activities of SalE salicylate esterase in crude extracts of cells grown on different media

Strain	Growth substrate(s)	Specific activity[a] (U/mg of protein)
ADP1	Succinate	<0.04
ADP1	Salicylate	<0.04
ADP1	Ethyl salicylate	3.44
ADP1	Succinate + ethyl salicylate	0.67
ADPW70	Succinate	<0.04
ADPW70	Salicylate	<0.04
ADPW70	Succinate + ethyl salicylate	<0.04
ADPW72	Succinate	<0.04
ADPW72	Succinate + salicylate	<0.04
ADPW72	Succinate + ethyl salicylate	0.45
E. coli(pADPW70)	LB + IPTG	52.5
E. coli(pADPW70)	LB	<0.04

[a] All reaction rates were based on the average of three measurements, none of which varied by >5%.

Although we were unable to measure salicylate hydroxylase activity encoded by the gene cloned into expression vector pET5a (see above), activity was detected in *E. coli* DH5α(pADPW34), which carries *salA*, *salR*, and *salE* (Fig. 1). Moreover, the activity was induced only when 2 mM salicylate was added to the LB medium. The specific activity was twofold higher in salicylate-induced *E. coli*(pADPW34) than in salicylate-grown ADP1. The relative activities exhibited by the strain with this cloned gene against a range of substituted salicylates were compared with those expressed in the wild-type ADP1 and found to be identical, within experimental error, with a broad substrate specificity except against the only available position 3-substituted salicylate (Table 6).

Phenotypes of mutants. To demonstrate that the aromatic moiety arising from salicylate and ethyl salicylate is channeled down the β-ketoadipate pathway, we checked the phenotypes of mutants blocked both in the *sal* genes and at three different points in the β-ketoadipate pathway (Table 7). The three β-ketoadipate pathway mutants were ADP6 (*pcaG*), which does not grow on 4-hydroxybenzoate or any substrate that feeds into the protocatechuate branch of the pathway; ADPW1 (*catA*), with a functionless catechol 1,2-dioxygenase which blocks the utilization of benzoate and catechol and which accumulates catechol as demonstrated by the brown coloration on agar plates from any substrate that feeds into the catechol branch; and ADPW38 (*ben*), which has an uncharacterized lesion in *benABC* and is unable to convert benzoate to catechol. The Km^r cassette insertion mutation in *salE* was also individually introduced into ADP6, ADPW1, and ADPW38 by natural transformation to produce double mutants, all of which were tested for growth on the appropriate carbon sources (Table 7). Only ADPW1 of the single β-ketoadipate pathway mutants failed to grow on salicylate or ethyl salicylate, with catechol accumulating in both media. The growth phenotypes of the double mutants were consistent with the proposed pathway (Fig. 3) in which the salicylate nucleus is fed into the benzoate branch at the level of catechol.

RT-PCR analysis of ADP1 transcripts. To confirm that the *sal* genes are transcribed during both ethyl salicylate and salicylate catabolism and that the operon structure is as implied by the gene organization (Fig. 1), transcripts from cells grown

TABLE 5. Specific activities of salicylate hydroxylase in crude extracts of cells grown on different carbon sources

Growth medium[a]	Specific activity[b] (U/mg of protein) in extracts of:					
	ADP1 (wild type)	ADPW67 (*salA*::Km^r)	ADPW70 (*salE*::Km^r)	ADPW72 (*salR*::Km^r)	DH5α(pADPW34) (pUC18::*salARE*)	DH5α(pUC18)
Salicylate	0.70	ND[c]	0.50	ND	ND	ND
Ethyl salicylate	0.30	ND	ND	ND	ND	ND
Salicylate + succinate	0.55	<0.01	ND	<0.01	ND	ND
Ethyl salicylate + succinate	0.11	ND	0.09	<0.01	ND	ND
Succinate	<0.01	<0.01	<0.01	<0.01	ND	ND
LB + salicylate	ND	ND	ND	ND	1.37	<0.01
LB	ND	ND	ND	ND	<0.01	ND

[a] *Acinetobacter* strains were grown in minimal medium containing aromatic substrates at 2 mM and/or succinate at 10 mM as indicated.
[b] All reaction rates were based on the average of three measurements, none of which varied by >5%.
[c] ND, not done.

on both of these substrates and on succinate as a noninducing negative control were examined. Two primer sets, spanning from *salD* through to *salE* and from *salA* through to *salR*, were constructed (Fig. 4A). The expected RT-PCR product sizes for the *salD-salE* and *salA-salR* amplicons were 1,195 and 988 bp, respectively. The PCR products obtained, together with restriction digests chosen to confirm the presence of expected restriction sites, were analyzed by agarose gel electrophoresis. The *salAR* products obtained from the total RNA of cells grown on both ethyl salicylate and salicylate were of the expected sizes (Fig. 4B). Also, a *salED* product of the expected size was obtained from the total RNA of cells grown on ethyl salicylate. The presence of restriction sites in the expected positions within the fragments was confirmed by digestion with *Cla*I (*salED*) and *Hin*dIII (*salAR*). No products were obtained from total RNA of succinate-grown cells or from reaction mixtures from which the reverse transcriptase had been omitted (data not shown).

DISCUSSION

Catabolic role of *sal* genes and proteins. The two enzymes salicylate esterase (SalE) and salicylate hydroxylase (SalA) reported in this paper are involved in the sequential catabolism of alkyl salicylates via salicylate to catechol. They represent another route into the β-ketoadipate pathway for substrates which are likely to be found as natural products, either directly of plant origin or as microbial breakdown products of plant compounds. In this study, both enzymes were expressed from cloned genes, SalE from the expression vector pET5a, from which it is expressed to a very high specific activity, and SalA from a pUC18 clone carrying *salERA*. The SalA salicylate hydroxylase activity shows the same relative substrate preferences as does the activity found in wild-type ADP1 grown on salicylate alone. In addition, insertional *salE* and *salA* knockout mutants, whose construction was facilitated by the natural transformation of ADP1, show the phenotype expected from the proposed pathway (Fig. 3). The further catabolism of the aromatic moiety into the *ben-cat* branch of the β-ketoadipate pathway at the level of catechol was also confirmed by the fact that only mutations in *catA* (for catechol 1,2-dioxygenase), and not those in the *ben* or the *pca* genes, eliminated the ability to utilize either the ester or the free salt of salicylate.

It is interesting that the location of these genes is directly adjacent to the *areCBA* operon, which we have recently described (17), whose role is to channel benzyl alkanoates into the β-ketoadipate pathway by hydrolysis of the esters to benzyl alcohol and two sequential dehydrogenase-catalyzed oxidations of benzyl alcohol to benzoate. The two sets of genes thus appear complementary in that the *are* genes are responsible for the catabolism of esters in which the alcohol moiety is aromatic whereas the *sal* genes encode proteins that function in the catabolism of esters in which the acid moiety is aromatic.

Comparisons of Sal proteins. Examination of the deduced amino acid sequence of SalE in the PROSITE database (16) shows that from residues 68 to 77 (IVLLGHSYGG) it has the signature characteristic of serine lipases, [LIV]-x-[LIVFY]-[LIVMST]-G-[HYWV]-S-x-G-[GSTAC], in which the serine is the active-site nucleophile. Its primary sequence does not align closely (<26% similarity) with other reported *Acinetobacter* esterases, one from *Acinetobacter lwoffii* RAG-1 (1) and two (a carboxylesterase [19] and the adjacent benzyl esterase, AreA [17]) from strain ADP1, all of which are longer and have their putative active site serines further from the N terminus than SalE.

The regulator protein SalR is clearly a member of the LysR family of regulator proteins, with the family signature motif containing a helix-turn-helix at residues 17 to 47 (NISKAA EILNLSQPSVTYNLNRLRKHLNNPL) according to the PROSITE database (16). The high degree of similarity of both SalR and SalA to the two *Pseudomonas* isofunctional proteins, NahR and NahG, from the *P. putida* NAH7 plasmid (24, 33) and *P. stutzeri* (6) implies the occurrence of past intergeneric exchange of genes by horizontal transfer and yet, within the context of conservation of amino acid sequence, an equilibrium of DNA composition, in terms of AT/GC ratio, with that of the host. Whereas the four *Pseudomonas* genes have G+C ratios of between 60 and 65%, the *salA* and *salR* genes have a composition more characteristic of the A+T-rich *Acinetobacter* genome (Table 2).

TABLE 6. Relative activities of SalA salicylate hydroxylases in crude extracts of cells

Assay substrate	Relative activity[a] in cell extracts of:	
	ADP1[b]	DH5α(pADPW34)[c]
Salicylate	100	100
3-Methyl salicylate	<0.1	<0.1
4-Methyl salicylate	56	54
5-Methyl salicylate	35	33
4-Chlorosalicylate	45	40
5-Chlorosalicylate	27	22

[a] Activity relative to that on salicylate, which is set at 100. All reaction rates were based on the average of three measurements, none of which varied by >5%.
[b] ADP1 was grown on minimal medium containing 2 mM salicylate; the specific activity was 0.71 U/mg of protein.
[c] *E. coli* DH5α(pADPW34) was grown on LB containing 2 mM salicylate; the specific activity was 1.37 U/mg of protein.

TABLE 7. Growth phenotypes of *Acinetobacter* strains

Growth substrate	Growth of strain[a] (relevant genotype)									
	ADP1 (wild type)	ADPW67 (*salA*)	ADPW70 (*salE*)	ADPW72 (*salR*)	ADPW38 (*ben*)	ADPW1 (*catA*)	ADP6 (*pcaG*)	ADPW88 (*salE ben*)	ADPW86 (*salE catA*)	ADPW87 (*salE pcaG*)
Benzoate	+	+	+	+	−	−B	+	−	−B	+
4-Hydroxybenzoate	+	+	+	+	+	+	−	+	+	−
Salicylate	+	−	+	−	+	−B	+	+	−B	+
Ethyl salicylate	+	−	−	−	+	−B	+	−	−	−
Succinate	+	+	+	+	+	+	+	+	+	+

[a] Growth was assessed on agar plates: +, good growth; −, no growth; −B, no growth, with accumulation of brown or black coloration (catechol).

The open reading frame *salD* appears to encode a member of a family of proteins of which FadL from *E. coli* (5) is perhaps the archetype. A number of these proteins encoded by open reading frames within gene clusters associated with aromatic catabolism have been reported, including TodX (29), XylN (GenBank accession no. D63341), TbuX (H.-Y. Kahng, A. M. Byrne, R. H. Olsen, and J. J. Kukor, submitted for publication), and CumH (14), but their function in this context has yet to be definitively determined. The overall level of similarity within the family is low, only 13 to 17%, but there are 23 conserved residues, which are also found in SalD (J. J. Kukor, personal communication). We have created a mutant of SalD, with a Km^r insertion, which is unable to grow on ethyl salicylate, but because this insertion will exert a polar effect on *salE* expression, this does not prove that SalD is essential for its catabolism. However RT-PCR has shown that *salD* and *salE* are cotranscribed during growth on salicylate and ethyl salicylate, so it is reasonable to assume that they have related functions. Definitive proof of this would be obtained by constructing a *salD* deletion mutant without a concomitant polar effect on *salE*, and so far we have been unable to make such a mutant. What is clear is that the growth rate of the *salD*::Km^r mutant is severely reduced, even on noninducing substrates, and its cell viability is also impaired, with plate cultures of ADPW78 dying after 2 to 3 days, whereas cultures of ADPW70 with a *salE*::Km^r insertion remain as viable as those of wild-type ADP1.

Regulation of *sal* genes. SalA activity is induced by growth of ADP1 on salicylate or on ethyl salicylate (Table 5). It is probable that SalR is the protein that regulates expression of SalA, since (i) insertion of a Km^r cassette into *salR* in strain ADPW72 stops induction of salicylate hydroxylase activity by salicylate (Table 5); (ii) salicylate hydroxylase is also salicylate inducible in *E. coli* from pADPW34, which carries only *salARE*

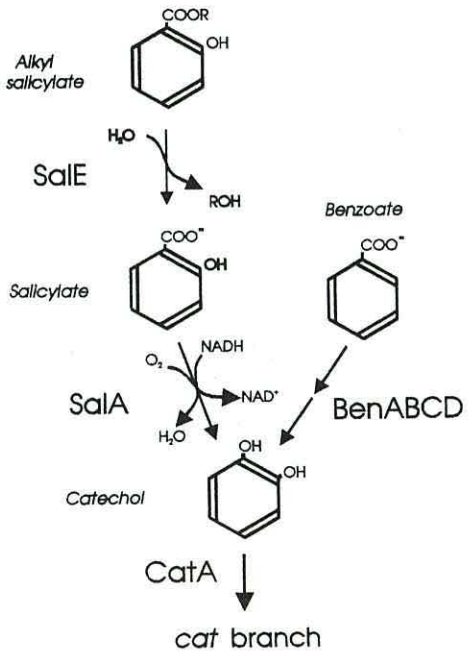


FIG. 3. Proposed pathway for the catabolism of alkyl salicylates linked to the *ben-cat* branch of the β-ketoadipate pathway in *Acinetobacter* sp. strain ADP1.

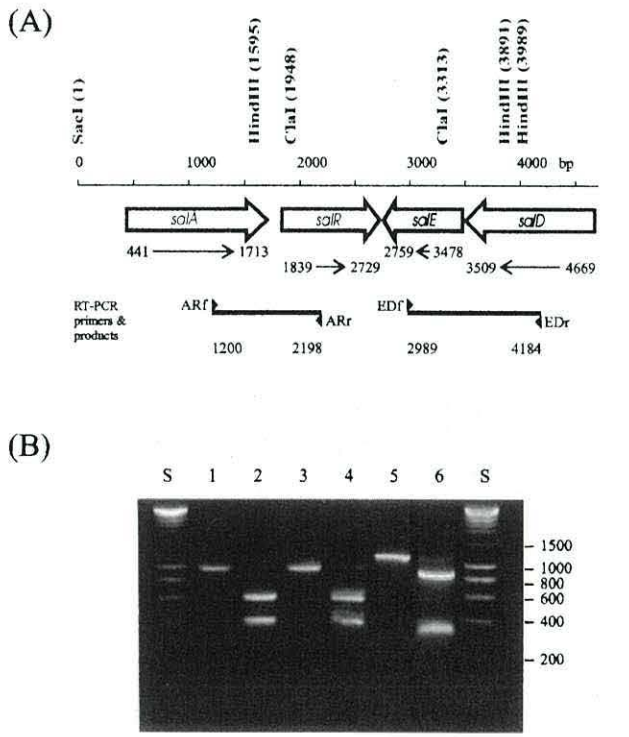


FIG. 4. RT-PCR of *sal* genes. (A) Positions of the genes relative to the *Sac*I site (at bp 1), the primers used for the RT-PCR, and the *Hin*dIII and *Cla*I restriction sites. (B) Agarose gel electrophoresis of RT-PCR products amplified from ADP1 grown on ethyl salicylate and salicylate. The sizes of molecular size markers (in base pairs) in lanes S (HyperLadder I; Bioline, London, United Kingdom) are indicated on the right. Lanes: 1, *salAR*, salicylate-grown cells (expected size, 988 bp); 2, *salAR*, salicylate-grown cells digested with *Hin*dIII (582 and 406 bp); 3, *salAR*, ethyl salicylate-grown cells (expected size, 988 bp); 4, *salAR*, salicylate-grown cells digested with *Hin*dIII (582 and 406 bp); 5, *salED*, ethyl salicylate-grown cells (expected size, 1195 bp); 6, *salED*, ethyl salicylate-grown cells digested with *Cla*I (860 and 335 bp). No detectable products were obtained in control reactions, with each pair of primers, from which reverse transcriptase had been omitted or in reactions carried out on succinate-grown cells (data not shown).

(Table 5); and (iii) there are close amino acid sequence homologies between SalR, SalA, and the *Pseudomonas* homologues NahR and NahG. A further point of comparison with the *Pseudomonas* genes is that there is also homology with the regions upstream of *salA* identified by Schell (24) as being the promoter sites at which NahR interacts; upstream of *nahAa* is sequence TCA-N_6-TGA, and upstream of *nahG* is sequence TCA-N_3-TGATGA (24) (GenBank accession no. M11863). The sequence upstream of *salA* is TCA-N_3-TGATGG. Just downstream of this there is also a −35 sequence, TAGGCAA TT, which has 5 bases that correspond to those of the −35 sequence identified for *nahAa*, TGGTGTATT (24) (GenBank accession no. M11863), but the putative −10 region shows no similarity. A major difference between the *sal* and *nah* genes is that whereas *nahR* is transcribed divergently from its adjacent catabolic gene, *nahAa*, *salA*, and *salR* appear to be cotranscribed as a single regulatory unit, as shown by the RT-PCR results. This implies that SalR controls its own expression.

Similarly, RT-PCR suggests that *salD* and *salE* are cotranscribed, although there remains a remote possibility, as is also the case with *salAR*, that the two genes are separately transcribed on two overlapping mRNAs. The induction results show that growth on ethyl salicylate is necessary for induction of SalE activity but that salicylate does not act as the inducer (Table 4). It is also clear that (i) when *salR* is inactivated in ADPW72 (*salR*::Km^r), SalE remains inducible by ethyl salicylate; and (ii) there are no obvious potential binding sites upstream of *salD* similar to *nahR* promoter sites. This points to the possibility that the regulatory mechanism for *salDE* differs from that of *salAR* and implies that there might be an additional regulator gene on the ADP1 chromosome that is involved in the induction of *salDE* but is not located in the immediate vicinity; we have sequenced about 5 kb further upstream of *salA* and found no obvious regulator gene present. A further possibility which needs to be tested is that *salD* and *-E* are cotranscribed with *areCBA* under the control of AreR.

ACKNOWLEDGMENTS

This research was funded by a BBSRC research studentship (to R.M.J.).

We thank Weiske Pool for technical help.

REFERENCES

1. **Alon, R. N., and D. L. Gutnick.** 1993. Esterase from the oil-degrading *Acinetobacter lwoffii* RAG-1: sequence analysis and over-expression in *Escherichia coli*. FEMS Microbiol. Lett. **112:**275–280.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
4. **Bauchop, T., and S. R. Elsden.** 1960. The growth of microorganisms in relation to energy supply. J. Gen. Microbiol. **23:**457–469.
5. **Black, P. N.** 1991. Primary sequence of the *Escherichia coli fadL* gene encoding an outer membrane protein required for long-chain fatty acid transport. J. Bacteriol. **173:**435–442.
6. **Bosch, R., E. Garcia Valdes, and E. R. B. Moore.** 1999. Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from *Pseudomonas stutzeri* AN10. Gene **236:**149–157.
7. **Collier, L. S., N. N. Nichols, and E. L. Neidle.** 1997. *benK* encodes a hydrophobic permease-like protein involved in benzoate degradation by *Acinetobacter* sp. strain ADP1. J. Bacteriol. **179:**5943–5946.
8. **Collier, L. S., G. L. Gaines III, and E. L. Neidle.** 1998. Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. J. Bacteriol. **180:**2493–2501.
9. **Doten, R. C., K.-L. Ngai, D. J. Mitchell, and L. N. Ornston.** 1987. Cloning and genetic organization of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**3168–3174.
10. **Elsemore, D. A., and L. N. Ornston.** 1994. The *pca-pob* supraoperonic cluster of *Acinetobacter calcoaceticus* contains *quiA*, the structural gene for quinate-shikimate dehydrogenase. J. Bacteriol. **176:**7659–7666.
11. **Eraso, J. M., and S. Kaplan.** 1994. *prrA*, a putative response regulator involved in oxygen regulation of photosynthetic gene expression in *Rhodobacter sphaeroides*. J. Bacteriol. **176:**32–43.
12. **Gerischer, U., and L. N. Ornston.** 1995. Spontaneous mutations in *pcaH* and *-G*, structural genes for protocatechuate 3,4-dioxygenase in *Acinetobacter calcoaceticus*. J. Bacteriol. **177:**1336–1347.
13. **Gralton, E. M., A. L. Campbell, and E. L. Neidle.** 1997. Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the *Acinetobacter* sp. strain ADP1 (BD413UE) chromosome. Microbiology **143:**1345–1357.
14. **Habe, H., K. Kasuga, H. Nojiri, H. Yamane, and T. Omori.** 1996. Analysis of cumene (isopropylbenzene) degradation genes from *Pseudomonas fluorescens* IP01. Appl. Environ. Microbiol. **62:**4471–4477.
15. **Harwood, C. S., and R. E. Parales.** 1996. The β-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50:**553–590.
16. **Hofmann, K., P. Bucher, L. Falquet, and A. Bairoch.** 1999. The PROSITE database, its status in 1999. Nucleic Acids Res. **27:**215–219.
17. **Jones, R. M., L. S. Collier, E. L. Neidle, and P. A. Williams.** 1999. *areABC* genes determine the catabolism of aryl esters in *Acinetobacter* sp. strain ADP1. J. Bacteriol. **181:**4568–4575.
18. **Juni, E.** 1972. Interspecies transformation of *Acinetobacter*: genetic evidence for a ubiquitous genus. J. Bacteriol. **112:**917–931.
19. **Kok, R. G., V. M. Christoffels, B. Vosman, and K. J. Hellingwerf.** 1993. Growth-phase dependent expression of the lipolytic system of *Acinetobacter calcoaceticus* BD413: cloning of a gene encoding one of the esterases. J. Gen. Microbiol. **139:**2329–2342.
20. **Neidle, E. L., M. K. Shapiro, and L. N. Ornston.** 1987. Cloning and expression in *Escherichia coli* of *Acinetobacter calcoaceticus* genes for benzoate degradation. J. Bacteriol. **169:**5496–5503.
21. **Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman.** 1992. The α/β hydrolase fold. Protein Eng. **5:**197–211.
22. **Ornston, L. N., and E. L. Neidle.** 1991. Evolution of genes for the β-keto-adipate pathway in *Acinetobacter calcoaceticus*, p. 201–237. *In* K. Towner, E. Bergogne-Berezin, and C. A. Fewson (ed.), The biology of *Acinetobacter*. Plenum Press, New York, N.Y.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Schell, M. A., and E. F. Poser.** 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. J. Bacteriol. **171:**837–846.
25. **Schell, M. A., and M. Sukordhaman.** 1989. Evidence that the transcription activator encoded by the *Pseudomonas putida nahR* gene is evolutionarily related to the transcription activators encoded by the *Rhizobium nodD* genes. J. Bacteriol. **171:**1952–1959.
26. **Schrag, J. D., and M. Cygler.** 1997. Lipases and α/β hydrolase fold. Methods Enzymol. **284:**85–107.
27. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weights, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
28. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
29. **Wang, Y., M. Rawlings, D. T. Gibson, D. Labbe, H. Bergeron, R. Brousseau, and P. C. Lau.** 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. Mol. Gen. Genet. **246:**570–579.
30. **Williams, P. A., and B. N. Zaba.** 1997. EnzPack for Windows. Biosoft, Cambridge, United Kingdom.
31. **Williams, P. A., and L. E. Shaw.** 1997. *mucK*, a gene in *Acinetobacter calcoaceticus* ADP1 (BD413), encodes the ability to grow on exogenous *cis,cis*-muconate as the sole carbon source. J. Bacteriol. **179:**5935–5942.
32. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
33. **You, I.-S., D. Ghosal, and I. C. Gunsalus.** 1988. Nucleotide sequence of plasmid NAH7 gene *nahR* and DNA binding of the *nahR* product. J. Bacteriol. **170:**5409–5415.
34. **You, I.-S., D. Ghosal, and I. C. Gunsalus.** 1991. Nucleotide sequence analysis of the *Pseudomonas putida* PpG7 salicylate hydroxylase gene (*nahG*) and its 3′-flanking region. Biochemistry **30:**1635–1641.