DOCTOR OF PHILOSOPHY

Gene expression, exchange and distribution of salicylate 1-hydroxylase genes in Acinetobacter species

Kay, Catherine Melanie

Award date:
2005

Awarding institution:
University of Wales, Bangor

Gene Expression, Exchange and Distribution of Salicylate 1-Hydroxylase Genes in *Acinetobacter* Species

A Thesis submitted to the University of Wales, Bangor

by

Catherine Melanie Kay

in candidature for the Degree of Philosophiae Doctor.

Department of Biological Sciences
University of Wales, Bangor
Bangor.
September 2005

Acknowledgements

I would like to thank first and foremost Prof. P.A. Williams for his patience, kindness, and supervision during the course of this PhD, the writing of this PhD thesis and for providing me the opportunity to study for a research degree. I would also like to thank and acknowledge Prof. M. J. Bailey of CEH Oxford for CASE award funding. Many, many thanks go to my "Lab Mother", Dr. L. Shaw, for her friendship, instruction, support and proof-reading of this thesis.

For the provision of environmental isolates of *Acinetobacter* spp. I would like to thank and acknowledge Dr. D. Young and Prof. L. N. Ornston (Yale University USA) and Prof. J. Fry (University of Wales, Cardiff, UK). Thanks also to Dr. G. Elliot, (University of Wales, Aberystwyth) for providing *Acinetobacter baylyi* strain ADPW249, and offering help and advice. Thanks also to Dr C. Pook (SBS) for assistance with PAUP data handling.

Thanks also to members of G51 past and present, especially Dr. R. Jones for his molecular biology instruction and Dr. B. Stallwood for her support and enthusiasm. Thanks to Mr Chris Gwenin, for his consideration, understanding and friendship; Miss Guo Wei for her kindness, friendship and Chinese plum candy; Mr Xiao-Feng Sun, my co-worker, for his kind support; Dr D. Pryce for his advice and encouragement. Thanks also to all members of SBS.

I also want to thank my family, my partner, Brian, and my daughter, Kela, who have supported me through every twist, turn and step of the way of this journey. Especially thanks to Kela who gave up precious Sims time to let me write my thesis.

Summary

The objective of this study was to examine and quantify the ability of environmental isolates of *Acinetobacter* spp. to restore salicylate hydroxylase ($SalA^+$) phenotypes to five *Acinetobacter baylyi* strain ADP1 *salA*$^-$ mutants, via natural transformation, and to investigate the organization and diversity of salicylate gene homologues within *Acinetobacter*. The ADP1 *salA* allele was disrupted using overlap extension PCR mutagenesis (OEPM) so that small 4 bp deletions were incorporated within different gene regions, creating five mutant strains, ADPW257 - ADPW261. Out of 34 *Acinetobacter* isolates examined, three *salA* homologues that had 60% DNA identity with the ADP1 *salA*, were isolated from *Acinetobacter* sp. strains AD3-1, FS50 and BS6. The *salA* genes from *Acinetobacter* sp. strains AD3-1 and FS50, shared 98% identity. A DNA fragment cloned from the genome of *Acinetobacter* sp. strain AD3-1 contained three open reading frames, identified from the nucleotide sequence as: *salA* (encoding a salicylate 1-hydroxyalse), *salK* (encoding a putative salicylate transport protein with closest homology to the ADP1 BenK benzoate transport protein) and *salR* (encoding a LysR-type transcriptional regulator).

A quantitative assessment of transformation frequency was conducted using homologous ADP1 and heterologous AD3-1 donor DNA: consisting of plasmid-borne salicylate operons, cloned copies of truncated *salA* and total genomic DNA digests. Transformation frequencies resulting from natural transformation and homologous recombination with homologous DNA were affected by the allelic position of the mutation and the length of flanking DNA homology. Recombination using heterologous donor DNA appeared to be confined to small regions of micro-homology resulting in overall transformation frequencies of $\sim 10^{-7}$. Significantly reduced transformation

frequencies ($<10^{-8}$) were observed in strain ADPW259, where the deletion was centrally located, indicating that either recombination events occurring in this region proved lethal to protein activity or that possibly the central region of the gene was resistant to homologous recombination.

Chapter 1 Introduction

page

Chapter 2 Materials and methods

Chapter 3 Creation of *Acinetobacter* mutants

Chapter 4 Identification, isolation and characterization of salicylate hydroxylase genes from *Acinetobacter* spp.

Chapter 6 Characterization of the *Acinetobacter* sp. strain AD3-1 salicylate operon

Figures **page**

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix B Amino acid alignments

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendix A

Antibiotics

Km – kanamycin (Km^R kanamycin resistance)

Ap – ampicillin (Ap^R ampicillin resistance)

Media

LBA – Luria Bertani agar

Phenotypic

Sal – salicylate

Succ – succinate

E-sal – ethyl salicylate

p-HBA – *p*-hydroxybenzoate

Ben - benzoate

Others

ADP – adenine diphosphate

AGE – agarose gel electrophoresis

ARDRA - amplified rDNA restriction analysis

β-ME - 2-mercaptoethanol

CFE – cell free extract

DNA - deoxyribonucleic acid

FAD – flavin adenine dinucleotide

FMN – flavin mononucleotide

GMO – genetically modified organism

h - hour

HFIR – homology facilitated illegitimate recombination

HGT – horizontal gene transfer

IPTG – isopropylthiogalactopyranoside

kbp - kilobase-pairs

kDa - kilodaltons

L - litre

LB - Luria Bertani medium

LTTR – LysR-type transcriptional regulator

MFS – major facilitator superfamily

mg - milligram

min - minute

ml - milliliter

M - molar

mM - millimolar

NAD^+ - nicotinamide adenine dinucleotide

NADH – (NADH + H^+) reduced nicotinamide adenine dinucleotide

NADPH – (NADPH + H^+) reduced nicotinamide adenine dinucleotide phosphate

nm - nanometre

OEPM – overlap extension PCR mutagenesis

RFLP - restriction fragment length polymorphism

SalA – salicylate 1-hydroxylase

SDS - sodium dodecyl sulphate

SDS-PAGE - sodium dodecyl sulphate-polyacrylamide gel electrophoresis

TCA – tricarboxylic acid cycle

UV - ultraviolet

µg - microgram

µl - microlitres

µM - micromolar

UP - ultra pure

X-Gal – 5-bromo-4-chloro-3-indolyl-β-D-galactoside

Kinetic terms

V_{max} – maximal velocity

K_m – Michaelis constant

1.1 Microbial induced soil mineralization

Microbial soil populations possess versatile nutritional capabilities that allow them to metabolize a wide range of organic compounds as carbon and energy sources. An enormous range of organic compounds exist within the soil biosphere originating from both natural and anthropogenic sources providing microorganisms with a complex environmental niche. The cycling of organic compounds containing in particular, carbon, nitrogen, phosphorus and sulphur in the soil biosphere depends primarily on the abilities of microorganisms to participate in the mineralization of organic compounds. Microbial populations in soil, mediate the re-cycling and transformation of organic compounds to CO_2 and H_2O as a result of their diverse biodegradative capabilities. Quantitatively, the most important series of transformations in which microbial populations are concerned are those involving carbon compounds. The major sources of carbon compounds deposited in soil are derived directly from plants and microrganisms. In plants, the main types of high molecular weight compounds are cellulose, hemicellulose and lignin polymers together with smaller quantities of fats, waxes, oils, proteins and nucleic and amino acids (Wackett and Hershberger, 2001).

Man-made organic compounds present within the biosphere result mainly from industrial and agricultural sources and are often toxic to higher organisms. These compounds are often formed directly or indirectly from fossil fuels. The range of compounds is enormous and groups with established toxicity and /or recalcitrance, e.g. polyaromatic hydrocarbons such as naphthalene, phenanthrene, biphenyls, polychlorinated biphenyls (PCBs), phenolics, substituted-phenolics and hydrocarbons such as toluenes, xylenes, cresols and methylbenzenes are deposited within the biosphere. Dissimilation of these

compounds by biodegradative enzymes and the adaptive evolution of microbial species enables partial or complete mineralization of such pollutants in the environment.

1.1.1 Aromatic hydrocarbons

An important class of organic compounds is represented by the aromatic hydrocarbons that exist in both mono- and poly-nuclear forms. These compounds are characterised by the unsaturated benzene ring structure, C_6H_6, which has a large negative resonance stability termed its aromaticity. These compounds possess a relatively stable structure that does not readily react to the addition of water and other nucleophiles, as a direct result of the strength of the carbon-carbon bonds. Because the six carbon-carbon bonds are all of equal strength and distance the double bond electrons do not belong to any particular bonds but are delocalized about the ring (Fig.1.1.1). The result of this is that the strength of each bond is between that of a double and single bond producing the relatively unreactive structure (Wackett and Hershberger, 2001).

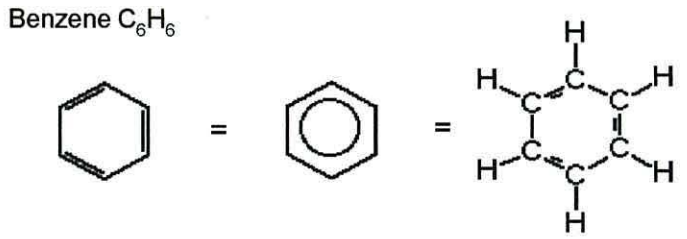


Fig.1.1.1. The structure of benzene showing C-C and C-H bonds and delocalised double bonds (Wackett and Hershberger, 2001).

Despite this integral resonance the ability to degrade these aromatic compounds is widely distributed among microorganisms, occurring in eukaryotic and prokaryotic phylogenetic and taxonomic divisions. Many biodegradative organisms described in the literature are found in the α, β and γ subdivisions of the proteobacteria and the high and low G+C mole % Gram-positive bacteria. Representative Gram-positive organisms involved in biodegradation in aerobic soils are *Rhodococcus*, *Bacillus*, *Arthrobacter, Acintomycetes,* and *Nocardia* species. Gram-negative organisms in soil are represented by *Pseudomonas*, *Acinetobacter, Enterobacter, Agrobacterium* and *Rhizobium* species (Harwood and Parales, 1996).

The aromatic ring structure presents a considerable barrier to biodegradation and metabolic strategies must overcome a higher activation energy than that required in the dissimilation of non-aromatic compounds (Wackett and Hershberger, 2001; Gibson and Subramanian, 1984). As a result microorganisms have evolved to use specific pathways and enzymes for the biodegradation of aromatic compounds.

1.1.2 Aerobic oxidation of aromatic compounds

The aerobic degradation of aromatic compounds by bacteria is catalysed by oxygenase or hydroxylase enzymes that activate molecular dioxygen and incorporate it directly into the chemical structure of the organic molecule (Mason and Cammack, 1992). Oxygenases incorporating one atom of molecular oxygen into the substrate are termed monooxygenases. This group reduces the other atom of oxygen to water, functioning therefore as part oxygenase and part oxidase and may be referred to as *mixed function oxygenases.* They are also referred to as hydroxylases since the main substrate

becomes hydroxylated. Oxygenases incorporating both atoms of molecular dioxygen into the substrate are termed dioxygenases (Mason and Cammack, 1992; Harayama *et al.*, 1992). Dioxygenases are categorised into two groups - ring hydroxylating dioxygenases and ring cleavage dioxygenases.

Both monooxygenases and dioxygenases require cofactors capable of reacting with and activating molecular dioxygen. Dioxygen exists in the atmosphere as a triplet molecule, a diradical, with its two unpaired electrons possessing the same orbital-spin orientation. The incorporation of dioxygen into the organic molecule requires the inversion of one of the orbital-spin electrons. As the activation energy of this reaction is very high, oxygenases overcome this barrier through the utilisation of a transition metal co-factor, typically iron, as a catalyst. In some cases, other transition metals are used or, as in the case of monooxygenases, flavins or pteridines serve as cofactors in place of or in addition to metal ions (Mason and Cammack, 1992; Wackett and Hershberger, 2001). The binding of dioxygen to the enzyme-coordinated iron atom facilitates spin conversion of one of the paramagnetic electrons due to the effect of some of dioxygen's electron density overlapping with the d-orbitals of the transition metal. This serves to move electrons between energy states, making dioxygen more reactive with organic substrates (Wackett and Hershberger, 2001).

1.1.3 Flavin reactions with molecular oxygen

Monooxygenases are flavoproteins where the flavin adenine dinucleotide (FAD) cofactor is tightly bound to the enzyme as a prosthetic group. The catalytically functional portion of the enzyme is the isoalloxazine ring system, specifically N-5 and C-4a (Fig.1.1.2). The bright yellow colour of flavoproteins is due to the presence of the oxidized form of the isoalloxazine chromophore.

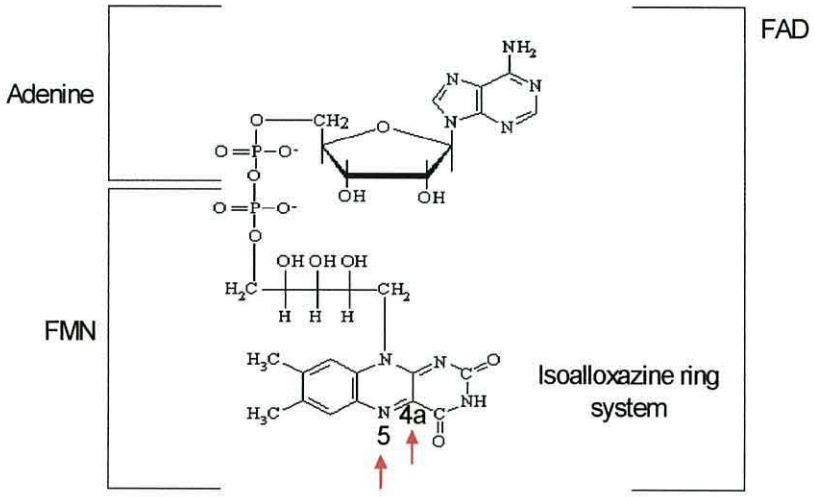


Fig.1.1.2. Flavin adenine dinucleotide molecule. The catalytically functional regions of the isoalloxazine ring N5 and C4a are indicated by arrows.

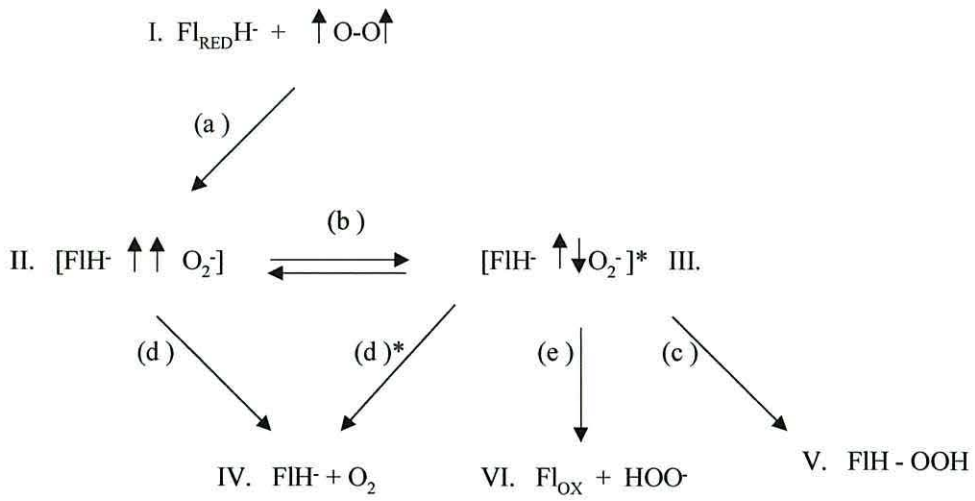


Fig.1.1.3. Possible routes for the reduction of oxygen by reduced monooxygenase FAD containing flavoproteins (Ghisla and Massey, 1989).

Flavoproteins require NAD(P)H as electron donors to reduce the FAD cofactor. The oxidase reaction of NAD(P)H to $NAD(P)^+$ and FAD reduction in these enzymes occurs in the absence of enzyme substrate. Possible routes for the reduction of oxygen are described in Fig.1.1.3. Interaction of $FADH_2$ with molecular oxygen facilitates the transfer of an electron from reduced flavin (I) to O_2 to form a paramagnetic complex of O_2 and the flavin radical (II). Spin inversion of the unpaired flavin electron leads to the biradical complex (III), in which the two electrons have paired spins and can combine to give the covalent flavin 4a-hydroperoxide (V). The hydroperoxide itself can dissociate heterolytically to yield hydrogen peroxide and oxidized flavin (VI). Reduced flavin reacts in a second-order reaction with O_2 to give a neutral radical FlH^- and O_2^- (IV). The FlH^- formed does not react further efficiently. The hydroperoxide moiety (VI) requires activation and/or modulation of its

chemical activity for reaction with a second substrate (i.e. an aromatic substrate). Three possible mechanisms can be formulated for the reactions of the hydroperoxide, (a) nucleophilic activation, (b) radical activation, or (c) electrophilic activation. The suggested mode of activation for aromatic hydrocarbon hydroxylase enzymes is that in the presence of substrate the hydroperoxide complex is converted to an intermediate charge transfer complex which is subsequently transformed to the complex of enzyme bound flavin 4a-hydroxide and product. This complex can then dissociate into the components oxidized flavin, hydroxylated substrate and water (Ghisla and Massey, 1989).

1.1.4 Monooxygenases

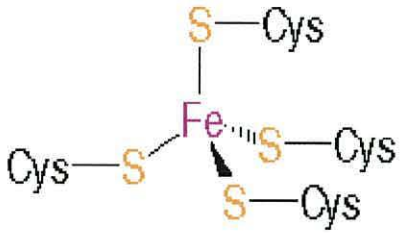
Bacterial monooxygenases that hydroxylate aromatic compounds via the insertion of one atom of molecular oxygen into the substrate are very commonly observed, especially for substrates that are already substituted e.g. salicylate and *p*-hydroxybenzoate. These enzymes catalyze the formation of aromatic compounds with two hydroxyl groups at positions either *ortho* or *para* to each other, serving as an intermediate substrate that is essential in the metabolic flow of many biodegradative pathways. As will be discussed later the formation of these intermediate dihydroxylated substrates is a prerequisite for oxidative ring cleavage. The majority of monooxygenases that catalyze monohydroxylation of the aromatic ring are single component flavoprotein enzymes, e.g. *p*-hydroxybenzoate hydroxylase and salicylate 1-hydroxylase, although multicomponent monooxygenases such as toluene-4 monooxygenase have been identified (Harayama *et al.,* 1992). According to their sequence similarities, oxygenases can be grouped into several protein families. Local sequence similarities within the groups are observed, for instance, an ADP binding fingerprint associated with a βαβ fold, that binds FAD, NADH or

NADPH is highly conserved among monooxygenases of the same family group and similar among oxygenases of different families (Wirenga *et al.*, 1986). Aromatic flavoprotein monooxygenases are classified into subgroups according to their subunit sizes. All flavin-containing monooxygenases possess approx. 20% overall amino acid identity. The strongest sequence conservation occurs in and adjacent to the FAD binding regions.

1.1.5 Dioxygenases

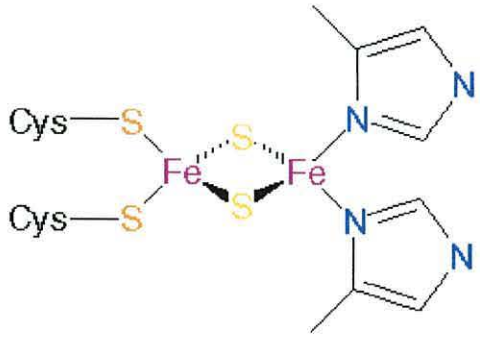
Ring hydroxylating dioxygenases are soluble, multicomponent enzyme systems requiring NADH and non-haem iron as co-factors, and are comprised of the same types of redox components (Cerniglia, 1992; Mason and Cammack, 1992). The redox components are composed of iron-sulphur proteins, ferredoxins, which mediate electron transfer within a wide range of metabolic reactions. They are classified into several sub-groups according to the nature of their iron-sulphur cluster(s).

The simplest iron-sulphur cluster is the rubredoxin-type centre (Fig.1.1.4) which has a mononuclear iron centre, where the iron is tetrahedrally co-ordinated to the protein with 4 cysteine residues [$Fe(cys)_4$]. Polymetallic iron-sulphur proteins, in which the iron ions have variable oxidation states, are placed into two subgroups, chloroplast or plant -like and Rieske-type [2Fe-2S] ferredoxins. Chloroplast-like [2Fe-2S] ferredoxins (Fig.1.1.4) are small, hydrophilic, iron-sulphur containing proteins functioning solely in the transfer of electrons. These [2Fe-2S] clusters contain two atoms of ferric iron (Fe^{3+}) bridged by two labile sulphide atoms co-ordinated to the protein by four sulphide ligands contributed by four cysteine residues.



Rubedoxin-type mononuclear iron-sulphur cluster

Chloroplast-like iron-sulphur protein [2Fe-2S] cluster



Rieske-type iron-sulphur protein [2Fe-2S] cluster

Fig.1.1.4. Examples of ferredoxin iron-sulphur clusters mediating electron transfer reactions.

Rieske-type [2Fe-2S] ferredoxins (Fig.1.1.4) contain two iron ions bridged by two sulphide ligands but coordinated to the protein by two cysteine and two histidine residues. Both chloroplast and Rieske ferredoxin types contain two Fe^{3+} ions when oxidized and are reddish in colour. Upon reduction, one of the Fe^{3+} iron atoms is reduced to ferrous iron (Fe^{2+}) and the visible absorption spectra decreases (Mason and Cammack, 1992; Fukuyama *et al.,* 1980; Gubriel *et al.*, 1989).

Depending on the substrate and source of the enzyme system, ring-hydroxylating dioxygenases differ in the distribution of the redox components and the number and the size of the protein subunits (Cerniglia, 1992). A typical arrangement of the components comprises a reductase site, an electron transfer chain and a terminal oxygenase site containing a Rieske-type [2Fe-2S] iron-sulphur cluster. The reductase system of different ring hydroxylating dioxygenases may occur in three different arrangements, depending on the number of subunits and the location of the Rieske type [2Fe-2S] cluster.

1.1.5.1 Two-component dioxygenases

Benzoate 1,2-dioxygenase is an example of a two-component dioxygenase (Fig. 1.1.5). The flavin moiety and the iron-sulphur cluster are combined in the same protein with the second component being the terminal electron acceptor. Two classes of two-component dioxygenases have been identified depending on whether they contain FAD or FMN. The sequences of the iron-sulphur flavoproteins of the two component dioxygenases bear homologies with chloroplast-type ferredoxins.

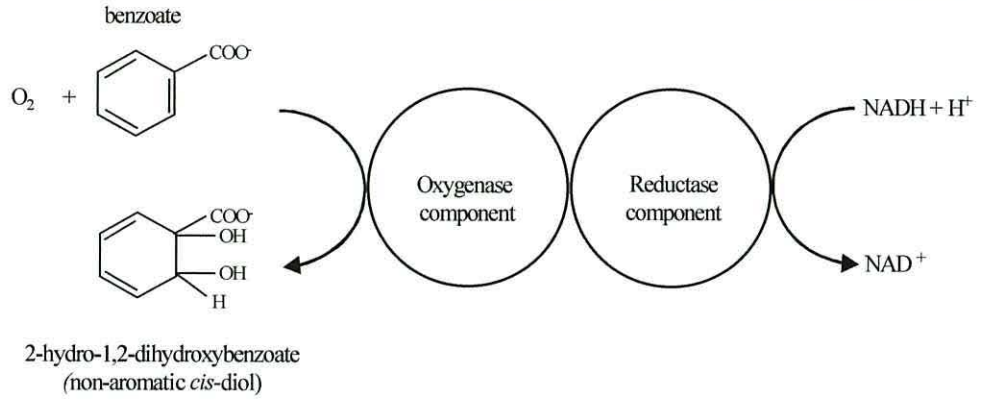


Fig.1.1.5. Two-component benzoate 1,2-dioxygenase system. The terminal oxygenase component is composed of two non-identical subunits encoded by two genes, such as *benAB* from *Acinetobacter baylyi* strain ADP1. The reductase component contains both FAD and a feredoxin iron-sulphur type cluster. The reductase component, encoded by a single gene *benC*, mediates electron transfer and can catalyze NADH dependent cytochrome c reduction (Neidle *et al.*, 1987).

1.1.5.2 Three-component dioxygenases

In three component dioxygenases the electron transfer chain is comprised of a flavoprotein and a separate iron-sulphur protein. These dioxygenases are categorised into two groups depending on the nature of the ferredoxin component, which may be either a chloroplast-type or a Rieske-type iron-sulphur cluster. Naphthalene dioxygenase is a good example of a three component dioxygenase (Fig.1.1.5.1) in which the electron transport chain is comprised of both an iron-sulphur flavoprotein and a ferredoxin (Haigler and Gibson, 1990).

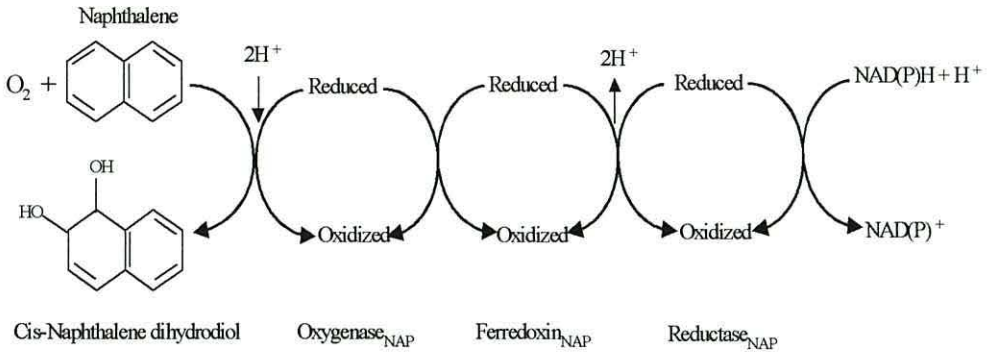


Fig.1.1.5.1. The reaction catalysed by the three component naphthalene dioxygenase showing the electron transport chain, (Haigler and Gibson, 1990; Parales *et al.,* 2000).

1.1.6 Aerobic degradation of aromatic compounds

In general, the aerobic oxidation of aromatic compounds is a two-stage process (Harwood and Parales, 1996). The initial stage involves a series of ring modification enzyme reactions to destabilise the aromatic structure and prepare the substrate for ring cleavage reactions. The general mechanism is carried out via ring hydroxylating mono- or dioxygenase enzymes forming a *cis*-dihydrodiol intermediate compound (Fig.1.1.6). The aromaticity of the substrate is restored via a dehydrogenase-catalyzed conversion of the cis-dihydrodiol to the dihydroxylated ring cleavage substrate.

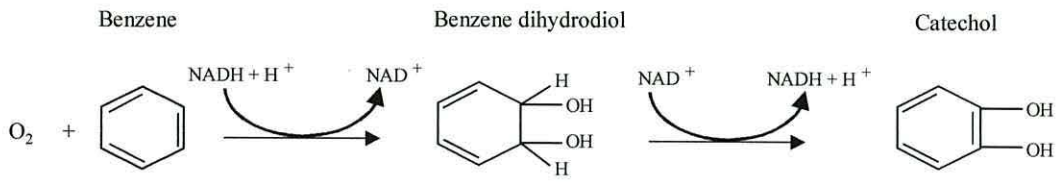


Fig.1.1.6. The oxidation of benzene to catechol showing the formation of the cis-dihydrodiol intermediate, cis-1,2-dihydroxycyclohexa-3,5-diene, and the catechol end-product of the reaction (Gibson and Subramamian, 1984).

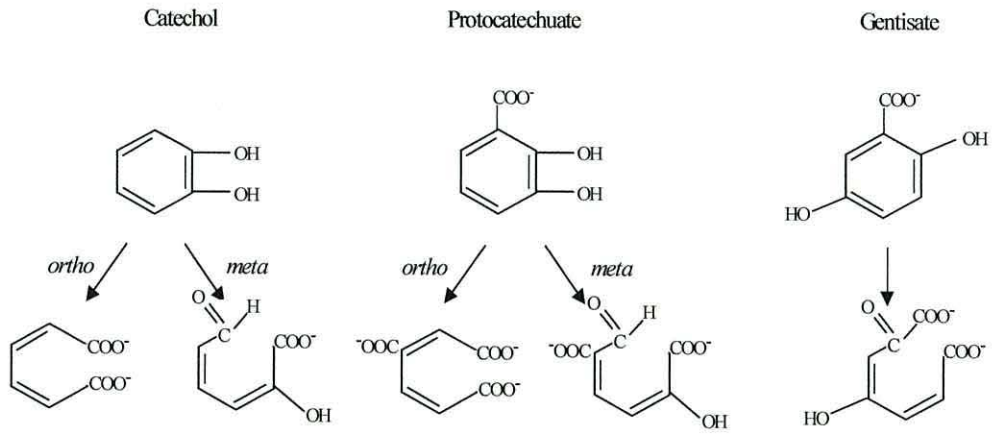


Fig.1.1.6.1. *Ortho*-, *meta*- and *para*-cleavage of the dihydroxylated aromatic ring substrate showing intradiol and extradiol cleavage reactions that occur in the biodegradation of the three possible central intermediate compounds, catechol, protocatechuate and gentisate (Harwood and Parales, 1996).

The second stage of oxidative degradation involves catalysis of ring cleavage prior to further degradation steps culminating in the production of central metabolic intermediates that enter the tricarboxylic acid cycle (TCA). The enzymes that carry out these reaction steps may differ in terms of substrate specificity, organisation and regulation but collectively they follow the same general process (Stainer and Ornston, 1973; Harwood and Parales, 1996). Ring cleavage dioxygenases have no co-factor requirement and cleave the aromatic ring through either intradiol or extradiol cleavage (Fig.1.1.6.1).

1.1.6.1 Hydroxylated ring cleavage pathways

The initial dihydroxylation of the substrate results in the formation of one of three possible ring cleavage substrates, protocatechuate, catechol or gentisate (Fig.1.1.6.1). Which catecholic substrate is formed depends on the initial aromatic compound and the specificity of the oxygenase and dehydrogenase enzymes involved (Gibson and Subramanian, 1984; Harwood and Parales, 1996; van der Meer *et al.*, 1992; Mason and Commack, 1992). A wide variety of aromatic compounds are funnelled through these central intermediates (Figs. 1.1.6.1.2 - 1.1.6.1.4).

Further dissimilation of the dihydroxylated catechol intermediate, catechol, protocatechuate or gentisate is carried out via one of three possible pathways in aerobic degradation. Catechol and protocatechuate may be degraded via the β-ketoadipate pathway (*ortho*-cleavage) or the *meta*-cleavage pathway. Gentisate is further dissimilated via the gentisate-pathway. The nomenclature derives from the type of cleavage induced by the relevant dioxygenase. O*rtho*- cleavage occurs at a position between the two hydroxyl groups on the aromatic ring. *Meta*- cleavage occurs at a position adjacent to the one of the hydroxyls on the

aromatic ring. In the gentisate route, the hydroxyl groups are at a position *para*- to each other on the aromatic ring and cleavage occurs between the carboxylated-carbon and the adjacent hydroxyl-carbon group (Fig.1.1.6.1) (Harwood and Parales, 1996; Stainer and Ornston, 1973).

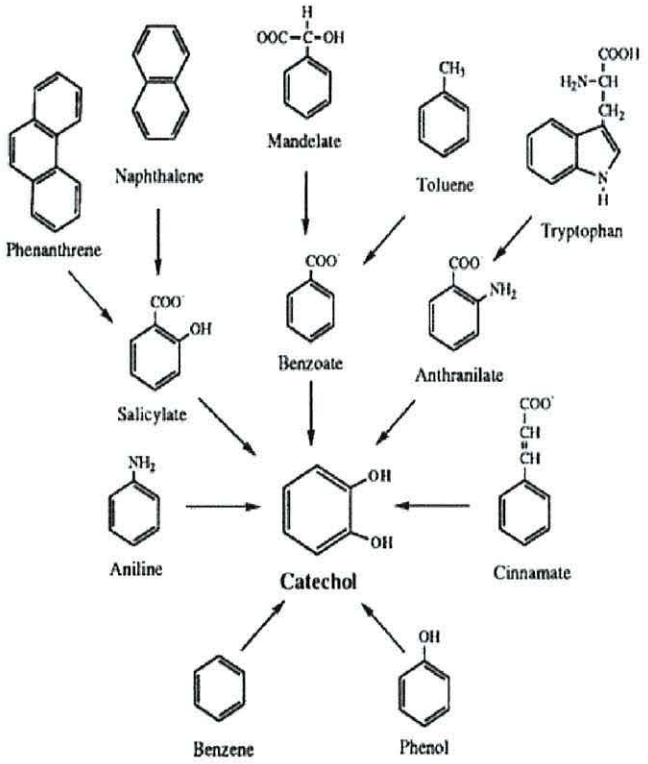


Fig.1.1.6.1.2. The range of compounds channelled through the central intermediate catechol.

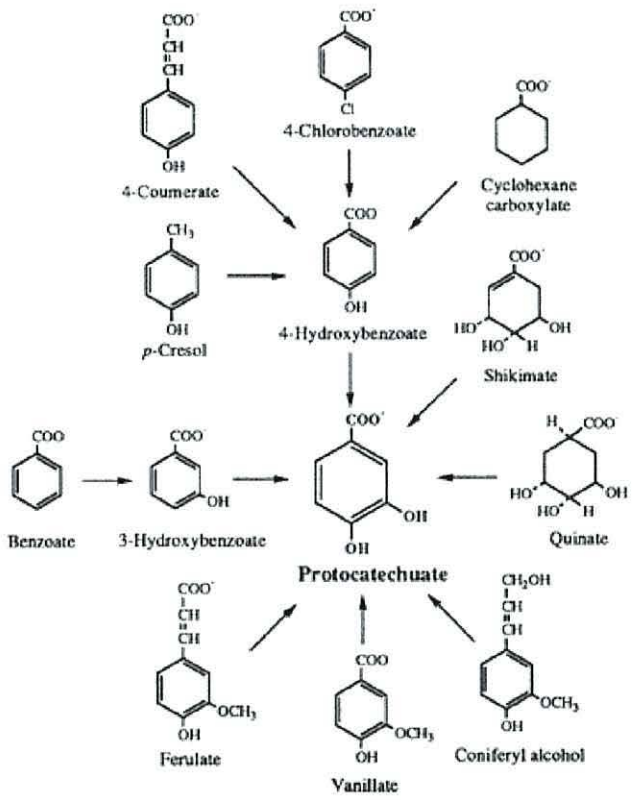


Fig.1.1.6.1.3. The range of compounds channelled through the central intermediate protocatechuate.

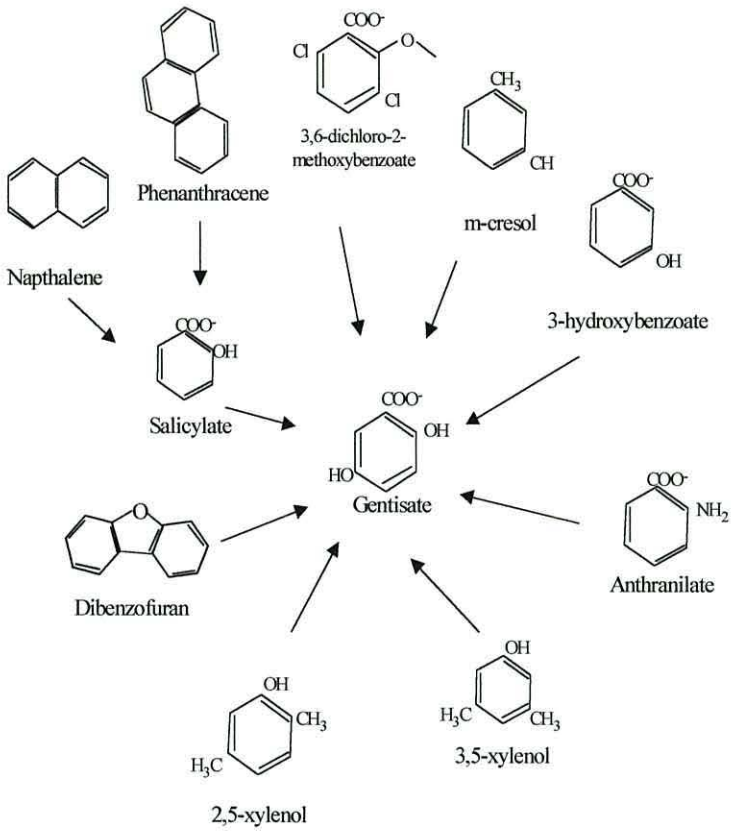


Fig.1.1.6.1.4. A selection of compounds channelled through the central intermediate gentisate.

Protocatechuate and catechol are the substrates for both the *ortho*, β-ketoadipate pathway and *meta* cleavage pathways. The β-ketoadipate pathway (Fig.1.1.6.2) is so named due to the central metabolite of the ortho pathway at which point both the protocatechuate and the catechol enzyme branches converge, prior to three further enzymic reactions ultimately forming succinate and acetyl-CoA (Harwood and Parales, 1996; Stainer and Ornston, 1973). The *ortho* pathway enzymes are almost exclusively chromosomally encoded and are widely distributed within natural microbial populations (Aldrich *et al.,* 1987; Doten *et al.*, 1987; Hughes *et al.*, 1988). A modified *ortho* pathway has evolved encoded on catabolic plasmids that exhibits similarity to the *ortho* cleavage pathway but which has substrate specificity for chlorinated catechols (Schlomann, 1994; van der Meer *et al.*, 1991; Dorn and Knackmuss, 1978).

The *meta* pathway (Fig.1.1.6.3) enzymes differ from the β-ketoadipate pathway and lead to the production of formate, acetaldehyde and pyruvate. They are generally plasmid encoded and have evolved to handle methyl-substituted substrates such as toluenes, xylenes, cresols and cumene (van der Meer, 1992; Muller, 1992; Gibson and Subramanian, 1984). The *para* cleavage pathway (Fig.1.1.6.4) possesses enzymes specific to the catabolism of gentisate producing malate and pyruvate or fumarate and pyruvate (Gibson and Subramanian, 1984).

1.1.6.2 *Ortho*-pathway

Protocatechuate and catechol β-ketoadipate pathways (Fig.1.1.6.2) share an analogous set of reactions in the production of the TCA cycle metabolites,

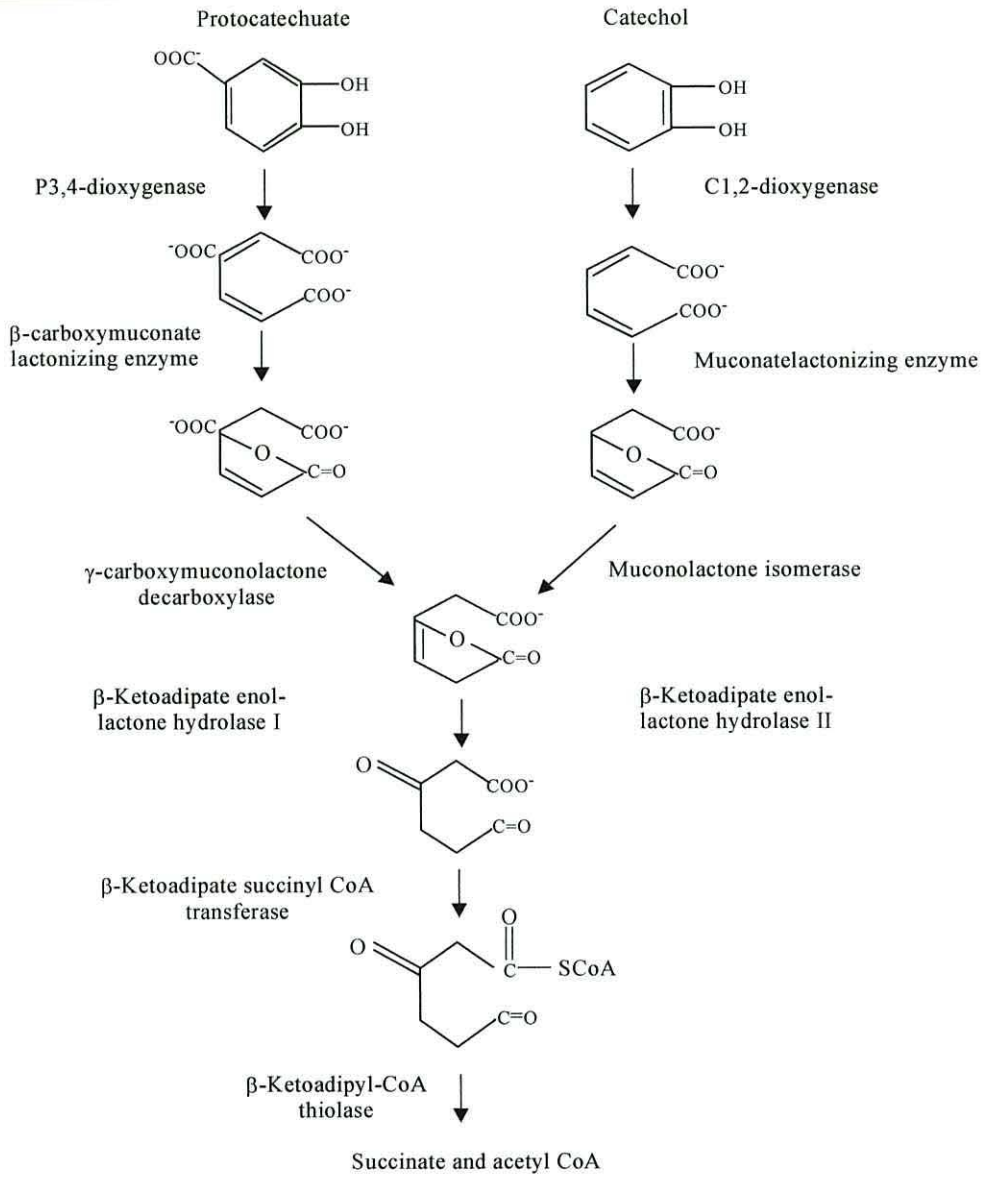


Fig.1.1.6.2. The β-ketoadipate pathway. The analogous protocatechuate and catechol branches of the β-ketoadipate pathway converge at β-ketoadipate enol lactone and are converted to the TCA cycle intermediates succinate and acetyl CoA by three common enzyme steps (Stainer & Ornston 1973; Harwood and Parales, 1996).

succinyl-CoA and acetyl-CoA (Harwood and Parales, 1996; Stainer and Ornston (1973). The pathway and intermediates were resolved in studies conducted in the 1950's and 60's (Stainer and Ornston, 1973). The enzymes of this pathway have been studied in detail to determine the specificity of function, homology and relatedness of the enzymes involved in either branch. In *Pseudomonas putida* and *Pseudomonas arvilla* biochemical studies have revealed that the ring fission enzymes, catechol-1,2-dioxygenase (C1,2O) and protocatechuate-3,4-dioxygenase (P3,4O), catalysing the production of *cis,cis*-muconate and β-carboxy-*cis,cis*-muconate respectively, share an amino acid sequence identity of ~25% (Nakai *et al.,* 1995). Catechol 1,2-dioxygenases may consist of identical or nonidentical α- and β-sub-units depending on the bacterial genera. Protocatechuate 3,4-dioxygenases contain equal numbers of two different subunits, α and β (Harayama *et al.,* 1992). The two enzymes are dimeric and enzyme activity in both depends on nonheme ferric iron (Fe^{3+}) as a cofactor which is ligated by two tyrosyl and two histidyl side chains binding the iron atom at the catalytic site of the enzyme (Que, 1985). The ligation of ferric iron within the catalytic subunit by histidyl and tyrosyl residues has a conserved amino acid sequence in intradiol dioxygenases suggesting a common ancestry for ring cleavage dioxygenases (Hartnett *et al.,* 1990; Neidle, 1988).

Two further enzymic steps that lead to the convergence of the two branches at β-ketoadipate enol lactone are carried out by analogous reactions. Muconate and β-carboxymuconate undergo lactonisation reactions to form muconolactone and γ-carboxymuconolactone respectively. These products in turn undergo specific reactions catalysed by muconolactone isomerase and γ-carboxymuconolactone decarboxylase respectively, resulting in the convergent metabolite β-ketoadipate enol lactone. Comparisons of the enzymes catalysing these steps reveal little amino acid sequence identity between the parallel

reactions which exhibit mechanistic and structural differences in enzyme activities, suggesting a separate evolutionary origin (Harwood and Parales, 1996; Chaki *et al.*, 1987; Katti *et al.,* 1989). However, although γ-carboxy-muconolactone decarboxylase shows little similarity to its counterpart in the catechol branch, it is a widely conserved enzyme of the pathway among different Gram-negative genera, revealing evolutionary relatedness with other bacterial species (Yeh *et al.*, 1981).

The remaining steps in the pathway convert β-ketoadipate enol lactone to the TCA cycle intermediates succinate and acetyl-CoA. β-Ketoadipate enol lactone hydrolase forms β-ketoadipate the substrate for the final reaction steps catalysed by β-ketoadipate:succinyl-CoA transferase and β-ketoadipyl-CoA thiolase. These two final enzymes of the pathway are representative of tightly conserved groups involved in Coenzyme A transfer and thiolytic cleavage identified in both eukaryotes and prokaryotes (Harwood and Parales, 1996; Courthesy-Thelulaz *et al.*, 1997; Kaschabek *et al.*, 2002). Mechanistically and functionally similar, they catalyse the reversible transfer of one CoA from one carboxylic acid to another. This conserved group of enzymes differs in substrate range and metabolic roles. Most bacterial CoA transferases are made up of two subunits and range from dimers to hetero-octamers. Thiolases may be grouped into two sets - thiolase-I and thiolase-II. The thiolase-I group is distinguished by broad chain length specificity for its substrates and is involved in degradative metabolism. Thiolase-II is specific for the thiolysis of acetoacetyl-CoA and is involved in biosynthetic metabolism (Igual *et al.*, 1992). Sequence comparisons of the known protein sequence of *P. putida* β-ketoadipyl-CoA thiolase with other bacterial groups showed high sequence identities (Kaschabek *et al.,* 2002).

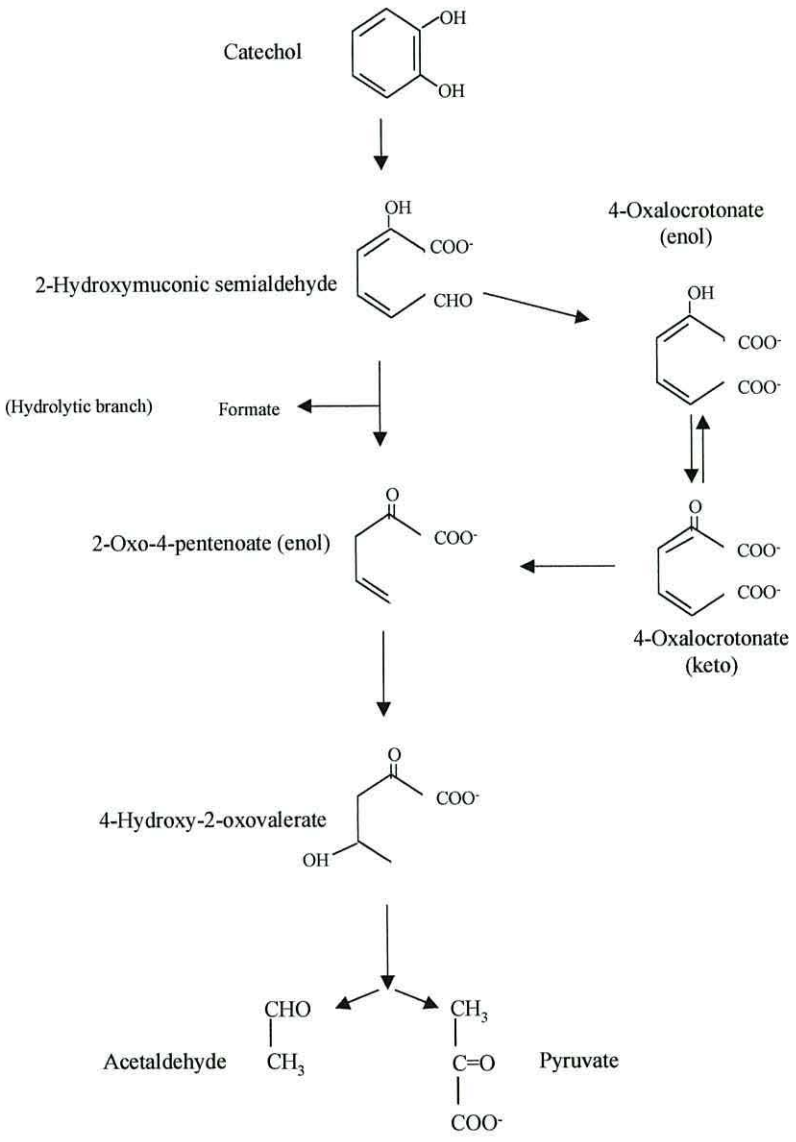


Fig.1.1.6.3. The reactions of the *meta* pathway starting at the central intermediate catechol and resulting in the formation of the TCA cycle intermediates acetaldehyde and pyruvate (Williams and Murray 1974).

1.1.6.3 *Meta-p*athway

The *meta* cleavage pathway is usually plasmid encoded and acts on the same intermediate substrates as the *ortho* pathway, protocatechuate and catechol. The degradation of these compounds to TCA cycle intermediates begins with an extradiol dioxygenase reaction resulting in ring fission. Substrates of the catechol *meta* pathway (Fig.1.1.6.3) are methylated aromatic rings i.e., toluene, xylene, naphthalene, phenanthracene, anthracene, *o*-, *m*- *p*- cresol and phenols (Bayly & Barbour, 1984). The enzymes of the *meta* pathway differ from those of the *ortho* pathway in terms of the mode of cleavage, enzyme mechanism, broad substrate specificity and the ultimate cycle products. The ability of the *meta* cleavage enzymes to utilise a wider range of products than the *ortho* cleavage enzymes may be attributed to the low specificity of induction and the function of several enzymes of the pathway (Bayly and Barbour, 1984).

The catechol intermediate is cleaved by catechol 2,3-dioxygenase, undergoing extradiol fission with the formation of the bright yellow product 2-hydroxymuconic-semialdehyde (2HMS). 2HMS may undergo two routes of dissimilation, depending on the initial substrate. A direct hydrolysis reaction catabolises the ring fission products of 3-substituted catechols (ketones) to 2-Oxo-4-pentenoate with the production of formate. An alternative pathway may occur, a branched oxidation step catalyzed by aldehyde dehydrogenase acting on the aldehyde products of catechols and 4-substituted catechols forms 4-oxalcrotonate (enol). The formation of 4-oxalocrotonate (enol) is followed by an isomerisation to the keto form of 4-oxalcrotonate which is then decarboxylated to 2-Oxo-4-pentenoate with the concomitant release of CO_2 (Sala-Trepat *et al,* 1972; Wigmore *et al.,* 1974). 2-Oxo-4-pentenoate is converted to 2-keto-4-hydroxy-valerate by a hydratase catalyzed addition of

water and then cleaved by an aldolase to form the TCA cycle intermediates pyruvate and acetaldehyde (Bayly and Barbour, 1984; Muller, 1992).

For the conversion of protocatechuate two *meta* dioxygenases exist, protocatechuate 2,3-dioxygenase (P2,3O) and protocatechuate 4,5-dioxygenase (P4,5O). The P4,50 *meta* cleavage pathway has been identified in bacteria such as *Micrococcus* species in the catabolism of phthalate (Ribbons *et al.*, 1984), *Sphingomonas paucimobilis* in the degradation of lignin derivatives (Masai *et al.*, 2000; Noda *et al.*, 1990) and *Comamonas testosteroni,* (Providenti *et al.,* 2001). Cleavage of protocatechuate by P4,5O produces 4-carboxy-2-hydroxymuconic-6-semialdehyde which is then converted to 4-carboxy-2-hydroxymuconate by an NAD-dependent dehydrogenase. A hydration step and aldol fission step yields the pathway products oxaloacetate and pyruvate (Ribbons *et al.*, 1984; Kernsten *et al.*, 1982; Masai *et al.*, 2000). The Protocatechuate 2,3-dioxygenase has been identified in Gram-positive bacteria such as *Bacillus macerans* and *Brevibacterium fuscum* (Wolgel *et al.,* 1993; Miller and Lipscomb, 1996) and catalyzes the conversion of protocatechuate to 2-hydroxy-5-carboxymuconic semialdehyde.

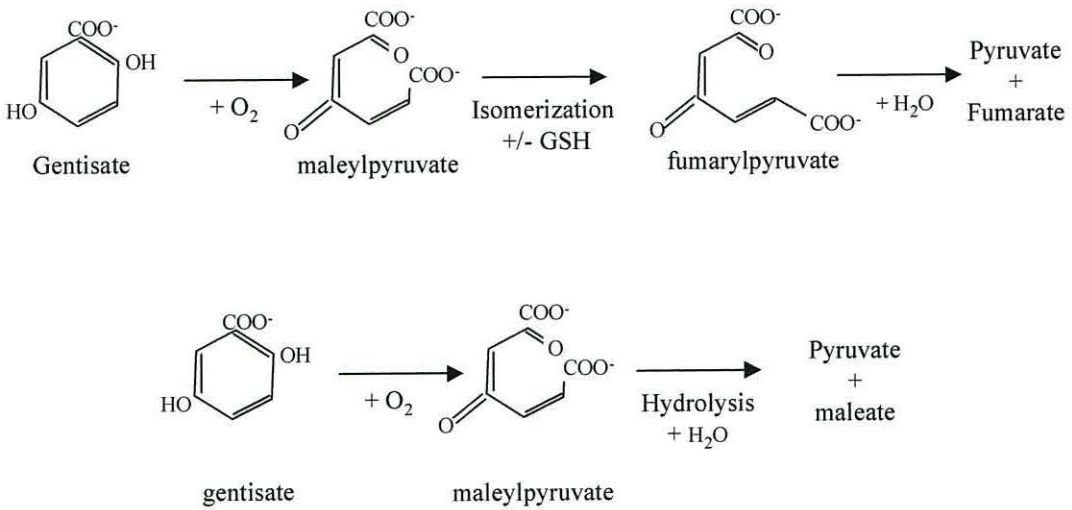


Fig.1.1.6.4 Pathways of gentisate metabolism in bacteria (Hagedorn *et al.*, 1985).

1.1.6.4 Gentisate-pathway

Gentisate and homogentisate, undergo cleavage between the carboxylated carbon and the adjacent hydroxylated carbon on the aromatic ring (Fig.1.1.6.1). Gentisate may be formed as an intermediary in the degradation of compounds such as *m*-cresol, 3-hydroxybenzoate, 3,6-dichloro-2-meth-oxybenzoate, salicylate, anthranilate, naphthalene and alkyl-substituted gentisates, intermediaries in the catabolism of 2,5-xylenol and 3,5-xylenol. The gentisate pathway has been recorded in many bacterial genera such as *Klebsiella pneumoniae* (Jones and Cooper, 1990), *Rhodococcus* sp. Strain B4 (Grund *et al.,* 1992), *Pseudomonas alcaligenes,* (Poh and Bayly 1988) *Pseudomonas putida* (Feng *et al.*, 1999), *Ralstonia* sp. strain U2 (Fuenmayor *et al.,* 1999; Zhou *et al.,* 2001), *Pseudomonas aeruginosa* (Hickey *et al.*, 2001) and *Streptomyces* sp. strain WA46 (Ishiyama *et al.*, 2004).

The scheme for the catabolism of gentisate was determined in studies using *Pseudomonas ovalis* in the latter part of the 1950s (Lack, 1959; Tanaka *et al.,* 1957). The dissimilation of gentisate is catalysed by gentisate-1,2-dioxygenase (GDO), a single polypeptide species of ~40 KDa, forming maleylpyruvate. Three mechanisms of maleylpyruvate dissimilation have been shown to occur (Fig.1.1.6.4). A glutathione (GSH) dependent isomerisation reaction forming fumarylpyruvate which is subsequently hydrolysed to fumarate and pyruvate (Lack, 1959). This isomerisation mechanism has been described for Gram-negative organisms (Johnson and Stainer, 1971; Wheelis *et al.,* 1967). A GSH independent isomerisation of maleylpyruvate to fumarylpyruvate again forming fumarate and pyruvate has been identified in mainly Gram-positive bacteria (Hagedorn *et al.*, 1985). The third route catalyzes direct hydrolysis of maleylpyruvate to form maleate and pyruvate.

1.1.7 Degradation of polycyclic aromatic hydrocarbons (PAHs)

Bacteria degrade PAHs such as naphthalene, phenanthrene, dibenzothiophene, biphenyls, chrysene and fluorene via induction of broad substrate specificity enzymes that may be plasmid or chromosomally located. The degradation of naphthalene was first identified and characterized in *P. putida* PpG7 encoded on the NAH7 plasmid (Yen and Serdar 1988) and in *P. putida* strain NCIB 9816 and derivatives (Davies and Evans 1964; Cane and Williams 1982; 1986; Yang *et al.*, 1994). The naphthalene pathway in these strains can be divided into two parts, an upper pathway and a lower pathway. The upper pathway (*nahAaAbAcAdBDCED*) encodes the enzymes required for the degradation of napthalene to salicylate. The lower pathway (*nahGTHINLOMKJ*) contains the enzymes for the conversion of salicylate to catechol initiated by catechol 2,3-oxygenase catalyzed extradiol cleavage and subsequent dissimilation via *meta* pathway enzymes to produce the TCA cycle intermediates acetaldehyde and pyruvate.

The upper and lower pathways are organized as two separate operons under the control of a single regulatory protein encoded by NahR, a LysR-type transcriptional regulator. The upper pathways identified within *Pseudomonas* spp., for the dissimilation of naphthalene, phenanthrene and dibenzothiophene are highly similar in amino acid sequence and gene organization to that of NAH7 (Denome *et al.*, 1993; Menn *et al.*, 1993; Yang *et al.*, 1994; Kiyohara *et al.*, 1994). The reactions catalyzed by the upper pathway initiate metabolism of naphthalene by naphthalene dioxygenase catalyzed incorporation of dioxygen into the aromatic nucleus to form *cis*-1,2-dihydroxy-1,2-dihydronaphthalene

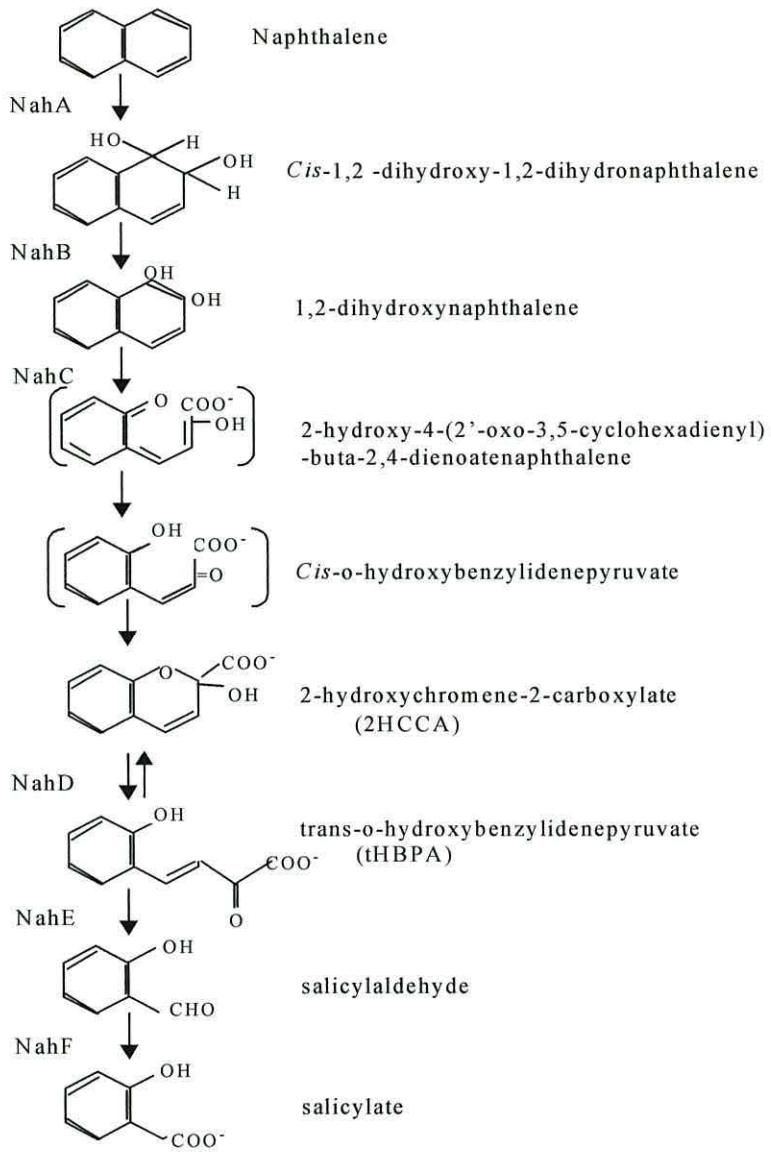


Fig.1.1.7. Upper pathway of naphthalene degradation as encoded on NAH7 and pTDG1 plasmids (Eaton and Chapman, 1992).

which is dehydrogenated to form the dihydroxy substrate. This is an initial reaction series analogous to that of the hydroxylation of single ring compounds.

Eaton and Chapman (1992) proposed the naphthalene degradation reaction mechanism (Fig.1.1.7) where 1,2-dihydroxynaphthalene is cleaved by 1,2-dihydroxynaphthalene dioxygenase, encoded by *nahC*, yielding an unstable compound 2-hydroxy-4-2-(2'-oxocyclohexa-3,5-dienyl)-buta-2,4-dienoate that recyclises forming another unstable intermediate, *cis*-o-hydroxy benzylidene-pyruvate (cHBPA). CHBPA undergoes spontaneous conversion to 2-hydroxychromene-2-carboxylate (2HCCA). Catalyzed and spontaneous isomerization of 2HCCA occurs leading to the production of 2HCCA and *trans*-o-hydroxybenzylidenepyruvate in equilibrium. An aldolase-hydratase enzyme metabolizes the tHBPA formed, producing salicylaldehyde and pyruvate. Salicylaldehyde is dehydrogenated to form salicylate, the final end product of the upper pathway and substrate of the lower pathway. In *Pseudomonas* spp. salicylate is a central intermediate of naphthalene and higher molecular weight PAHs degradation and also acts as the inducer of both the upper and lower naphthalene pathways (Yen and Gunsalus, 1985).

The higher molecular weight PAHs are more recalcitrant than naphthalene and the pathways leading to their degradation are less well understood; however bacterial species have been isolated that have the ability to degrade such compounds (Cerniglia, 1992). Species of *Burkholderia* (Juhasz *et al.*, 1997), *Stenotrophomonas* (Boonchan *et al.*, 1998), *Mycobacterium* (Schneider *et al.*, 1996; Vila *et al.*, 2001) and *Sphingomonas* (Nadalig *et al.*, 2002) have been shown to degrade a wide variety of high molecular weight PAHs. Recently identified pathways in sphingomonads for the degradation of mono- and polycyclic aromatic hydrocarbons are very similar to each other but dissimilar

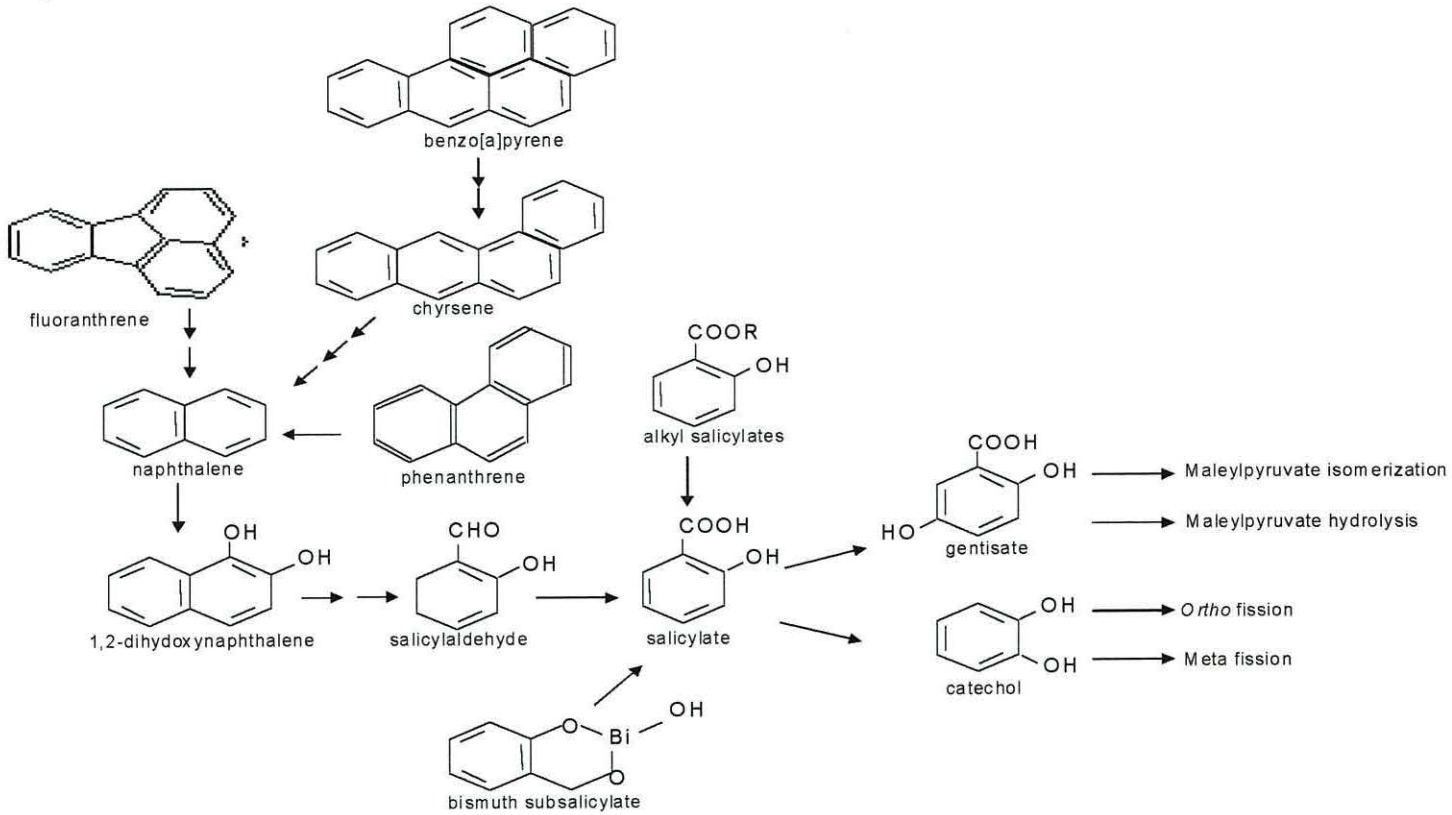


Fig.1.1.8. Some examples of aromatic compounds that involve the formation of salicylate during the course of their degradation. Salicylate itself may be further catabolized via a catechol or gentisate intermediate ultimately forming TCA cycle compounds.

to those described for the pseudomonads (Pinyakong *et al.*, 2003).

1.1.8 Salicylate degradation

Salicylate is a central intermediate of naphthalene and higher molecular weight PAH degradation, such as benzo[a]pyrene and chrysene, in some bacterial species (Fig. 1.1.8). Salicylate itself is a compound widely distributed within soils and many bacteria mineralize it via one of the three catecholic intermediates. It may be assimilated by either *ortho* or *meta* fission as catalyzed by salicylate 1-hydroxylase or by *para* cleavage catalyzed by salicylate 5-hydroxylase converting salicylate to gentisate. Although the catechol route is most commonly observed, identification and analysis of gentisate routes are being increasingly reported (Grund *et al.*, 1992; Allen *et al.*, 1997; Fuenmayor *et al.*, 1999; Hickey *et al.*, 2001; Ishiyama *et al.*, 2004).

1.1.8.1 Salicylate 1-hydroxylase

Salicylate 1-hydroxylase is a FAD-dependent external monooxygenase and catalyzes the conversion of salicylate to catechol and CO_2 with stoichiometric consumption of salicylate, NADH and oxygen. It was first isolated from *P. putida* S-1 (Yamamoto *et al.*, 1965) and from a then unclassified soil bacterium now identified as *Burkholderia cepacia* ATCC 29352 and ATCC 29351 (White-Stevens and Kamin, 1972; Wang and Tu, 1984). As the first purified flavoprotein it has been studied in great detail, in terms of substrate specificity, co-factor requirements, kinetic properties and enzyme mechanisms (White-Stevens and Kamin, 1972a; Katagiri *et al.*, 1966; Takemori *et al.*, 1970; Wang and Tu, 1984). The *P. putida* salicylate hydroxylase is monomeric with a molecular weight of 57 kDa with one mole of bound FAD per unit (Yamamoto

et al., 1965). The *B. cepacia* enzyme is dimeric with two identical subunits and two moles of FAD per unit with a molecular weight of 91 kDa (White-Stevens and Kamin, 1972). This enzyme was observed to undergo NADH oxidation reactions with certain substrate analogues termed 'pseudosubstrates' such as benzoate where the substrate remains unhydroxylated but O_2 is reduced to H_2O_2. This phenomenon was termed 'uncoupling' of oxygen reduction from hydroxylation and was also observed in studies performed with other flavoproteins (White-Stevens and Kamin, 1972). Both salicylate hydroxylases were active against a broad range of substituted salicylates in addition to salicylate. The *P. putida* S-1 enzyme binds salicylate and an external reductant (NAD(P)H) in a substituted ordered sequence to form a reduced enzyme-substrate complex. Oxidized $NAD(P)^+$ is released, facilitating the substituted binding of molecular oxygen to the reduced enzyme-substrate complex for production and ordered release of the products catechol, CO_2 and H_2O (Wang and Tu, 1984). The *B. cepacia* enzyme follows essentially the same reaction mechanism but the initial formation of reduced enzyme-substrate complex occurs with salicylate and NAD(P)H binding in a random order (White-Stevens and Kamin, 1972).

Two isofunctional salicylate1-hydroxylases NahG and NahW were identified in naphthalene-degrading *Pseudomonas stutzeri* strain AN10 (Bosch *et al.*, 1999). The chromosomally-located naphthalene genes of this strain are organized into two transcriptional units forming the upper and lower naphthalene pathways as in the classical NAH7 pathway. *nahG* is located in the same transcriptional unit as genes encoding the *meta* pathway. The *nahW* gene was identified in close proximity to but outside of this operon, upstream of NahR, the LysR type transcriptional regulator. Both salicylate hydroxylases were inducible in the presence of salicylate and were demonstrated to have broad substrate

specificities, catalysing the transformation of salicylate, methylsalicylates and chlorosalicylates. However, the relative activities with substituted salicylates differed between the two hydroxylases. NahG showed better activity with methylated salicylates than NahW, whereas the latter enzyme showed increased activity with chlorosalicylates. Furthermore, Southern hybridization studies demonstrated that this gene duplication event was present in all naphthalene-degrading strains of *P. stutzeri*.

Duplication of salicylate 1-hydroxylase has also been identified in the naphthalene catabolic plasmid pND6-1 (Li *et al.*, 2004). The complete nucleotide sequence of this plasmid revealed a high identity (99%) between the naphthalene genes, including *nahG*, and those of pDTG1. A second *nahG* located outside of the naphthalene lower pathway transcriptional unit was identified with closest amino acid identity to *nahG* of *P. stutzeri*. Gene amplification and duplication events may be advantageous to bacterial cells under extreme conditions and may occur as a response to environmental stress, providing a means of upregulating specific gene expression. Gene duplication may also provide the cell with a dispensable gene copy, the mutability of which provides the cells with enhanced catabolic ability.

In a recent study a three-component ring hydroxylating dioxygenase has been identified (PhnII) in a chrysene-degrading *Sphingomonas* strain. PhnII was demonstrated to use salicylate, methylsalicylates and anthranilate as substrates and functioned as a salicylate 1-hydroxylase catalysing the stoichiometric conversion of salicylate to catechol (Demaneche *et al.*, 2004). This enzyme is unrelated to known salicylate hydroxylases. PhnII is encoded by two structural genes *phnA2b* and *phnA1b* and requires coexpression of *phnA3*, encoding a reductase and *phnA4*, encoding a ferredoxin, for full activity. Amino acid

sequence comparisons of the α and β-subunits showed identity of 52 and 38% respectively with anthranilate dioxygenase from *Burkholderia cepacia*. In contrast to the *B. cepacia* enzyme that converts anthranilate to catechol, PhnII converted anthranilate to 2-aminophenol. PhnII showed broad substrate specificity for methyl substituted salicylates and functioned essentially as a monooxygenase.

1.1.8.2 Salicylate 5-hydroxylase

The naphthalene degradation pathway of *Ralstonia* sp. strain U2 (formerly *Pseudomonas* sp.) mediates the mineralization of naphthalene through gentisate to pyruvate and fumarate. The genes encoding the enzymes for conversion of naphthalene to salicylate are homologous with and in the same order as the classical NAH upper pathway (*nagAaAbAcAdBFCQED*) except for the insertion of two genes termed *nagGH* between *nagAa* and *nagAb* (Fuenmayor *et al.*, 1999; Zhou *et al.*, 2002). The genes for the conversion of gentisate to central metabolites (*nagJIKLMN*) are directly downstream. Unlike the classical naphthalene pathway the genes for the complete conversion of naphthalene are transcribed as a single unit under the control of NagR, a divergently transcribed LysR type transcriptional regulator (Jones *et al.*, 2003). NagGH is a unique monooxygenase that has significant similarity to multicomponent dihydroxylating dioxygenases and functions as a salicylate 5-hydroxylase. NagGH shares the electron transport chain of the naphthalene dioxygenase (Zhou *et al.*, 2001) and represents a novel method in salicylate catabolism.

A novel mechanism of salicylate catabolism has been identified in some *Rhodococcus* spp. (Grund *et al.*, 1992; Allen *et al.*, 1997) and *Streptomyces* sp. strain WA46 (Ishiyama *et al.*, 2004) where salicylate is converted to salicylyl-

CoA via a CoA ligase step, dependent on the presence of CoA and ATP. Salicylyl-CoA is the substrate for a salicylyl 5-hydroxylase, catalyzing conversion to gentisyl-CoA. Gentisyl-CoA appears to undergo spontaneous cleavage to form gentisate that is further catabolized by gentisate dioxygenase. This CoA addition is analogous to the process of anaerobic aromatic acid catabolism and is rarely observed in aerobic aromatic acid degradation (Ishiyama *et al.*, 2004).

1.2 The genus *Acinetobacter*

Acinetobacter are Gram-negative, strictly aerobic, non-motile, coccobacilli with diverse nutritional capabilities. Classification of *Acinetobacter* spp. has met with difficulties due to the lack of sufficiently distinctive phenotypic characteristics. They have previously been misclassified as strains of *Bacterium anitratum, Herellea vaginicola, Mima polymorpha, Moraxella lwoffi, Micrococcus* and *Achromobacter* (Juni, 1978). In a study of 106 *Acinetobacter* strains Baumann (1968) concluded that enough taxonomic evidence existed to classify these strains within a single genus, despite the variation in phenotypic properties. Further confirmation of the single genus status of *Acinetobacter* was put forward by Johnson *et al* (1970) through the identification of strong homologies in DNA-DNA hybridisation studies. A definitive classification assay was developed after the discovery that an unencapsulated mutant strain of *Acinetobacter* sp. strain BD4 termed BD413 (also ADP1) is highly competent for natural genetic transformation (Juni and Janik, 1969; Juni 1972). In Juni's genetic transformation assay method (1972), 265 *Acinetobacter* isolates were identified by the ability to transform an auxotrophic strain of BD413 to prototrophy. Similar but unrelated Gram-negative organisms produced no transformant colonies. Weak genetic interactions were observed between the

oxidase negative *Acinetobacter* and the oxidase positive *Moraxella osloensis* but were concluded to be induced by the highly conserved nature of ribosomal genes.

Recently seven novel species of *Acinetobacter* were isolated from activated sludge and assigned species names by virtue of the 16S-23S rRNA intergenic spacer region fingerprint (Carr *et al.*, 2003). This has increased the number of currently recognised species to 34. Studies with isolates from activated sludge have shown that many of these do not fit into the previously described DNA groups, indicating the genus is extremely taxonomically diverse. The wild type transformable BD4 strain and its derivatives (BD413 and ADP1) have been shown to form a separate entity within the genus *Acinetobacter*. Determination of the 16S sequence of strain BD413 (ADP1) revealed that BD413 clustered together with one of the newly identified species termed *Acinetobacter baylyi*. As a result BD4 and its derivative strains BD413 and ADP1 are reclassified as *A. baylyi* (De Baere *et al.*, 2004).

1.2.1 *Acinetobacter baylyi* strain ADP1

The ADP1 genome sequencing project has recently been completed (Barbe *et al.*, 2004). ADP1 possesses a single circular chromosome of 3,598,621 bp in length with an average G+C content of 40.3 mole % and 3325 coding sequences averaging 930 bp in size that cover 88.8% of the chromosome, distributed 60/40% between the leading and lagging strand respectively. About 20% of the genes in ADP1 are associated with catabolic functions. The diverse catabolic ability of ADP1 and the ease with which this strain undergoes natural transformation have made ADP1 a very useful and much used tool in many areas of study. The simplicity of genetic manipulation in ADP1 via natural

transformation and homologous recombination is described by Metzgar *et al.*, (2004). This laboratory has treated ADP1 as a model organism to create insertion deletions, marked chromosomal genes and unmarked gene deletions by simple PCR splicing techniques and selective plating with a view to automation of the process.

Two systems have been designed facilitating functional analysis of unknown, heterologous or mutated genes by recombinational DNA capture and marker rescue. PCR amplified/mutagenized DNA fragments are internalized by cells and directly incorporated by homologous recombination within the ADP1 host, directly into the genome (Kok *et al.*, 1999) or a resident plasmid (Melnikov and Youngman, 1999). Both systems require the use of a plasmid carrying either a flanking chromosomal ADP1 segment or regions homologous to the resident plasmid next to a kanamycin resistance marker gene. PCR amplification or mutagenesis is performed on the insert and marker gene followed by transformation of the product into an engineered host strain containing a partial kanamycin gene next to the homologous regions. Integration of the PCR fragment into this site can be selected for by restoration of marker gene function and phenotypic screening of transformants. This system allows repeated rounds of PCR mutagenesis whilst eliminating the need for repeated sub-cloning of PCR fragments and provides a mechanism to study heterologous genes expressed within an adaptable host strain other than *Escherichia coli.*

In genetic studies of ADP1 the function, organisation and regulation of many catabolic genes have been elucidated. Natural transformation has been used to assist in the cloning of DNA that complements mutations (Doten *et al.,* 1987b), to accurately map mutations within a gene (Doten *et al.,* 1987a; Neidle *et al.,* 1987), to study the effects of random mutations in essential catabolic pathway

genes (D'Argenio *et al.*, 1999) and in recovery of DNA from the chromosome by gap repair (Gregg-Jolly and Ornston, 1990). An example of the gene mapping precision facilitated by natural transformation is given by Gralton *et al.* (1997). In this study they used modified plasmid DNA to transform *Acinetobacter* recipient strains introducing restriction endonuclease sites into the chromosome. The analysis of artificial restriction site digested genomic DNA by Southern hybridization produced the first physical and genetic map of the ADP1 chromosome. The study revealed the relative positions of two supraoperonic clusters involved in the dissimilation of aromatic compounds.

The ability of ADP1 to take up DNA from any source and, providing a sufficient region of flanking homology exists or is provided within the cell, subsequent incorporation and expression of heterologous genes has facilitated research into the mechanisms of natural transformation and horizontal gene transfer within the environment. This is an area of study that has expanded due to environmental and health concerns raised as a result of biotechnological advances in the production and usage of genetically modified organisms (GMOs).

Kloos *et al.* (1994) developed a novel broad host range cell lysis system to study natural transformation and the environmental fate of DNA released by cell death. A plasmid, pDKL02, was constructed containing bacteriophage λ lysis genes under the control of the *Ptac* promoter. The addition of inducer to *E. coli*, *A. calcoaceticus* BD413 (now known as *A. baylyi* strain ADP1) and *P. stutzeri* cells resulted in cell lysis coincident with the release of high amounts of DNA into the surrounding medium. DNA released by this system was observed to persist in the culture medium for days. Natural transformation using the pDKL02 lysis system was monitored using isogenic but differentially marked

donor-recipient pairs of *A. baylyi* strain ADP1 (auxotrophy) and *P. stutzeri* (antibiotic resistance). In comparison with purified DNA, cell lysis generated free DNA was able to transform the recipient cells with the same frequency. This system has also been used in a modified format to identify antibiotic hypersusceptible mutants of *Acinetobacter* spp. (Lee *et al.*, 2003).

A. baylyi strain ADP1 has also been utilised to study natural transformation events that occur within biofilms (Hendrickx *et al.*, 2003). Biofilms are important natural components of bioremediation and phytoremediation systems. The degradation of xenobiotic compounds is reliant on the presence of specific biocatalysts within the system. Bioaugmentation of the system is required should a particular biodegradative capability be absent. Genetic transfer is a possible route for *in situ* bioaugmentation of the resident biofilm community. Hendrickx *et al.* (2003) investigated the potential feasibility of such augmentation utilising *A. baylyi* strain ADP1 in a monoculture biofilm model system.

1.2.2 ADP1 supraoperonic clustering

Supraoperonic clustering of biodegradative genes has been observed in many bacteria. The collective organisation of genes with related metabolic function may have resulted from ancestral gene amplification and duplication events. Gene duplication provides a mutable copy of a given gene providing an opportunity for increased substrate specificities and possible ecological niche advantage to the cell. Collective gene organisation allows tight regulation, sequential gene induction and may provide economy of gene expression.

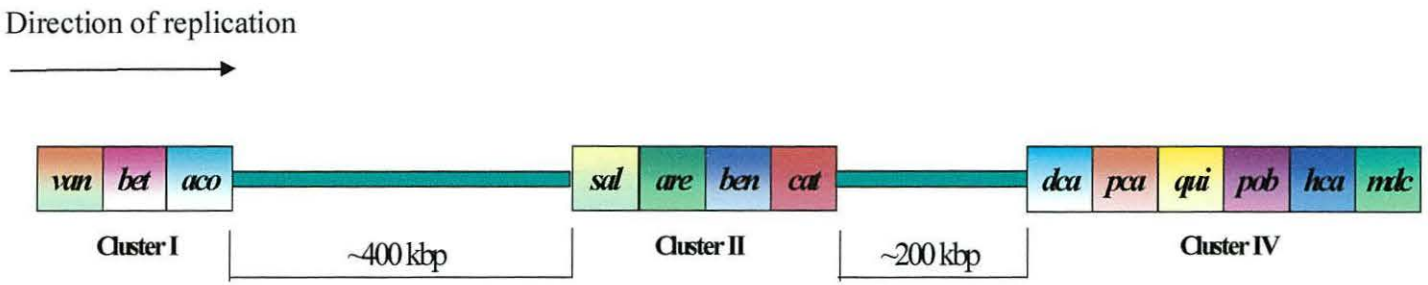


Fig.1.2. Organisation of ADP1 supraoperonic clusters of genes involved in aromatic compound degradation located clockwise within the ADP1 chromosome (Gralton *et al.*, 1997; Segura *et al.*, 1999; Barbe *et al.*, 2004). Gene designations: Cluster I - *van* - vanillate genes, *bet* and *aco* genes for 2,3-butanediol degradation. Cluster II - *sal* - salicylate genes, *are* - Aryl esters, *ben* - benzoate, *cat* - catechol. Cluster IV - *dca* - straight chain dicarboxylic acids, *pca* - protocatechuate, *qui* - quinate, *pob* - p-hydroxybenzoate, *hca* - hydroxycinnamate, *mdc* - malonate degradation.

In ADP1, enzymes for the degradation of aromatic compounds are organised in supraoperonic clusters and are channelled through the β-ketoadipate pathway via protocatechuate and catechol intermediates. The complete genome data has reported five catabolic gene clusters, three of which are involved in aromatic hydrocarbon degradation (Fig.1.2). Contained in the first island are the genes for vanillate degradation. Gene cluster IV encodes enzymes responsible for the degradation of plant-derived compounds such as coumarate, chlorogenate, quinate, shikimate and 4-hydroxybenzoate, which are degraded via the protocatechuate branch of the β-ketoadipate pathway (Fig.1.2.1). Identified as being associated with this gene cluster are the malonate catabolic genes (*mdc*) (Barbe *et al.*, 2004). Between catabolic clusters I and IV lie the supraoperonic *sal-are-ben-cat* genes. Benzoate, salicylate, anthranilate are channelled through the catechol branch forming the TCA cycle intermediates succinate and acetyl CoA.

1.2.2.1 Protocatechuate supraoperonic cluster

In ADP1 the protocatechuate branch of the β-ketoadipate pathway is encoded by *pcaUIJFBDKCHG* arranged as a single transcriptional unit (Gerischer *et al.*, 1998; Parke *et al.*, 2000). *PcaHG* and *pcaIJ* encode the separate subunits of single enzymes, the α and β subunits of protocatechuate 3,4-dioxygenase (P3,4O) and the two subunits of succinyl Coenzyme-A transferase, respectively. The genes for succinyl Coenzyme-A transferase in the catechol degradative branch, *catIJ* are 99% identical to those of *pcaIJ* of the protocatechuate branch (Kowalchuk *et al.*, 1994). P3,4O and catechol 1,2 dioxygenase (C1,2O) occupy equivalent positions in the β-ketoadipate pathway with their substrates differing by only a carboxyl group.

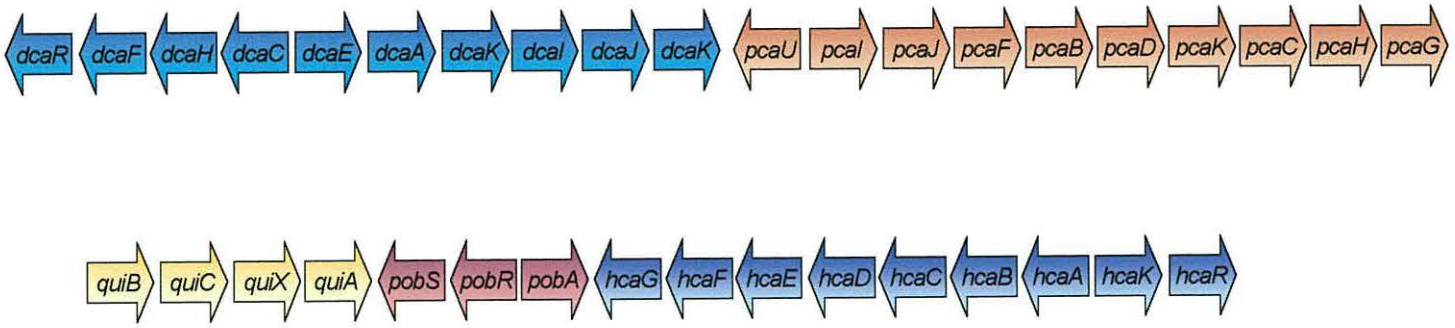


Fig. 1.2.1 Organization of the *dca-pca-qui-pob-hca* supraoperonic cluster. The direction of the gene arrows indicates the direction of transcription.

Comparisons of the amino acid sequence of these dioxygenases reveal a common ancestry. In a strain where C1,2O was inactivated it was shown that a single amino acid substitution within PcaG could confer the ability of this strain to grow with benzoate as the sole carbon source replacing C1,2O function (D'Argenio *et al.*, 1999).

The other structural genes of the protocatechuate branch of the β-ketoadipate pathway are encoded by *pcaBCDF* (Fig.1.1.6.2). The *pcaU* gene encoding the regulatory gene for this operon is divergently transcribed upstream of the *pca* cluster. PcaU is a bifunctional regulator, repressing and activating transcription at the *pcaI* promoter in the presence and absence respectively of inducer and acting as a repressor of its own expression (Trautwein and Gerischer, 2001). Transcriptional activation of the *pca* operon occurs through the synergistic association of two transcriptional regulators PobR and PcaU (Gerischer *et al.*, 1998). In response to the presence of its coinducer p-hydroxybenzoate, PobR activates the transcription of *pobA*, the structural gene encoding *p*-hydroxybenzoate hydroxylase catalyzing the conversion of p-hydroxybenzoate to protocatechuate (DiMarco *et al.*, 1993). Protocatechuate in turn is the co-inducer for PcaU. The level of expression of the pca genes was shown to vary depending on the individual inducing growth substrates. p-hydroxybenzoate as an inducer produced the highest levels of expression and ferulate the lowest (Trautwein and Gerischer, 2001).

PcaU has 50% identity with that of PobR and both proteins are classified as members of the IclR DNA-binding protein family (Gerischer *et al.*, 1998; Popp *et al.*, 2002). PCR mutagenesis of PobR produced an altered protein that could activate *pca* gene expression whilst still requiring p-hydroxybenzoate as an

inducer. Similarily PCR mutagenesis of *pcaU* created nucleotide substitutions allowing this protein to mimic PobR function (Kok *et al.*, 1998).

The *quiBCXA* and *pobSRA* operons are clustered directly downstream of the *pca* operon (Fig.1.2.1). The catabolism of quinate occurs in the periplasm yielding protocatechuate and requires the action of quinate dehydrogenase (*quiA*), dehydroquinate dehydratase (*quiB*) and dehydroshikimate dehydratase (*quiC*). Sequence comparisons suggest that *quiX* encodes a porin although its precise function is not known. QuiA is a member of a family of membrane associated pyrrolo-quinolone quinone-dependent dehydrogenases and is quite unique in that other microorganisms utilise NAD^+-dependent dehydrogenases (Elsemore and Ornston, 1995).

Many studies of protocatechuate catabolism have exploited the occurrence of spontaneous secondary mutations in strains unable to metabolize the toxic metabolite carboxymuconate. Deletions in a region upstream of the *pca* genes incurred through this selection procedure led to the identification of a strain unable to grow at the expense of the six-carbon dicarboxylic acid adipate and other medium chain length dicarboxylic acids (Parke *et al.*, 2001). Preliminary characterizations of this upstream region led to the identification of a cluster of ORFs linked to the *pca* genes involved in the dissimilation of straight chain dicarboxylic acids termed the *dca* genes. The *dca-pca-qui-pob* cluster was further extended upstream with the identification of linked genes involved in hydroxycinnamate dissimilation (Smith *et al.*, 2003).

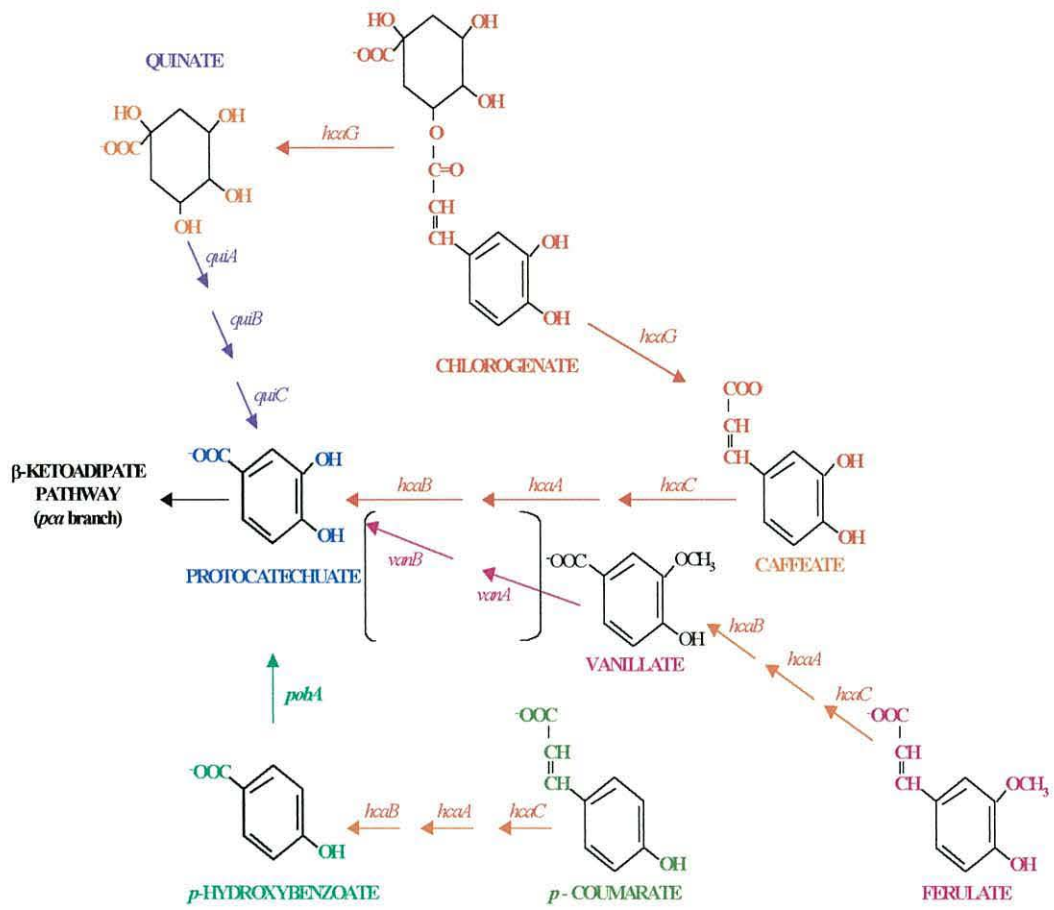


Fig. 1.2.2. Reactions catalyzed by the *pob-qui-pca-hca* genes. Funnelling of hyroxycinnamates such as chlorogenate, ferulate, caffeate and *p*-coumarate through protocatechuate and the *pca* branch of the β-ketoadipate pathway. The *vanAB* gees are bracketed as they do not occur as part of the *dca-pob-qui-pca-hca* supraoperonic cluster but are active in the dissimilation of vanillate to protocatechuate (figure adapted from Young *et al*., 2005).

Hydroxycinnamates are abundant in plants and are used in structural and chemical plant defence mechanisms (Young *et al.*, 2005). Chlorogenate, a highly significant and abundant hydroxycinnamate, is an ester commonly occurring as 5-*O*-caffeoylquinic acid (5-*O*CQ). Hydrolysis of 5-*O*CQ yields quinate and caffeate, compounds funnelled through the protocatechuate branch of the β-ketoadipate pathway. Downstream of *pobA* are five genes of the order *hcaCDEFG*. A mutation in *hcaC*, encoding a CoA-ligase, eliminated the ability to grow with the hydroxycinnamates caffeate, p-coumarate, ferulate and 3,4-dihydroxyphenylpropionate. A gene knockout in *hcaG* indicated this gene encodes a membrane-associated esterase that hydrolyzes chlorogenate to quinate and caffeate (Smith *et al.*, 2003). Figure 1.2.2 shows reactions catalysed by the *qui-pob-hca* genes in the channelling of hydroxycinnamate compounds through protocatechuate. The *dca-pca-qui-pob-hca* supraoperonic cluster is dedicated to metabolic channelling and degradation of a plethora of compounds originating from plants.

Exceptions to the rule of supraoperonic clustering are provided by the *van* genes, *vanAB*, which catalyze the degradation of vanillate, an intermediary in the catabolism of ferulate and the *ant* genes encoding the enzyme that converts anthranilate to catechol (Segura *et al.,* 1999; Bundy *et al.*, 1998). The *van* genes are located within catabolic cluster I, ~700 kbp upstream from the *dca-pca-qui-pob-hca* cluster containing the other genes necessary for the completion of the catabolism of ferulate and its structural analogues caffeate and coumarate (Fig.1.2). The *van* genes occupy a genetically unstable region of the chromosome where spontaneous mutations were shown to be mediated by IS*1236* including insertion of the composite trasposon Tn*5613* (Gerischer *et al.*, 1998; Segura *et al.*, 1999; Barbe *et al.*, 2004).

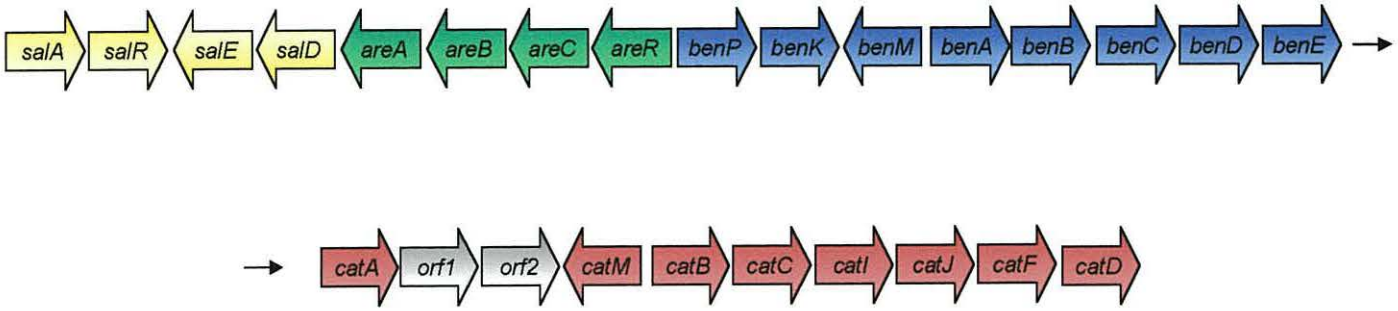


Fig.1.2.3. Gene organization of the *sal-are-ben-cat* supraoperonic cluster. Gene arrows indicate the direction of transcription. Small black arrows indicate continuation of the supraoperonic cluster.

1.2.2.2 Catechol supraoperonic cluster

The *sal-are-ben-cat* cluster (Fig.1.2.3) is part of a 30 kbp supraoperonic cluster which includes the genes of the *areABC* and *salARED* operons. The *sal* genes were demonstrated to catalyze the conversion of salicylate esters through salicylate to catechol (Fig.1.2.4) (Jones *et al.,* 1999; Jones *et al.,* 2000). Located in a 5 kbp region upstream of the *are-ben-cat* cluster the *sal* genes are organised into two transcriptional units, *salAR* and *salDE*. Salicylate 1-hydroxylase encoded by *salA* converts salicylate to catechol. Transcribed with *sal*A is its regulatory protein, *sal*R, a LysR-type transcriptional regulator. Amino acid alignments of SalA demonstrated 48% and 47% identity to salicylate 1-hydroxylases (*nahG*) of *P. putida* and *P. putida* Nah7 respectively. *salD* encodes a putative membrane porin resembling the FadL membrane protein of *Escherichia coli* involved in fatty acid transport and is located immediately upstream of *areA*. Convergently transcribed with *salD* is *salE*, an esterase with activity against short-chain alkyl esters of 4-nitrophenol. Hydrolysis of ethyl and methyl salicylate to form ethanol and salicylate enabled growth with these compounds as sole carbon sources channelled through the catechol branch of the β-ketoadipate pathway. Induction of *salE* occurred with ethyl salicylate but not with salicylate alone.

The *are* genes are located upstream from the *ben-cat* cluster (Fig.1.2.3). The *are* genes are responsible for the conversion of benzyl alkanoates to benzoate, salicylate or 4-hydroxybenzoate (Jones *et al.,* 1999; Jones *et al.,* 2001). The *are* genes, transcribed in the order *areCBA* are under the control of *areR*, a σ^{54} dependent regulator. a*reA* encodes a benzyl esterase, a*reB* encodes a benzyl alcohol

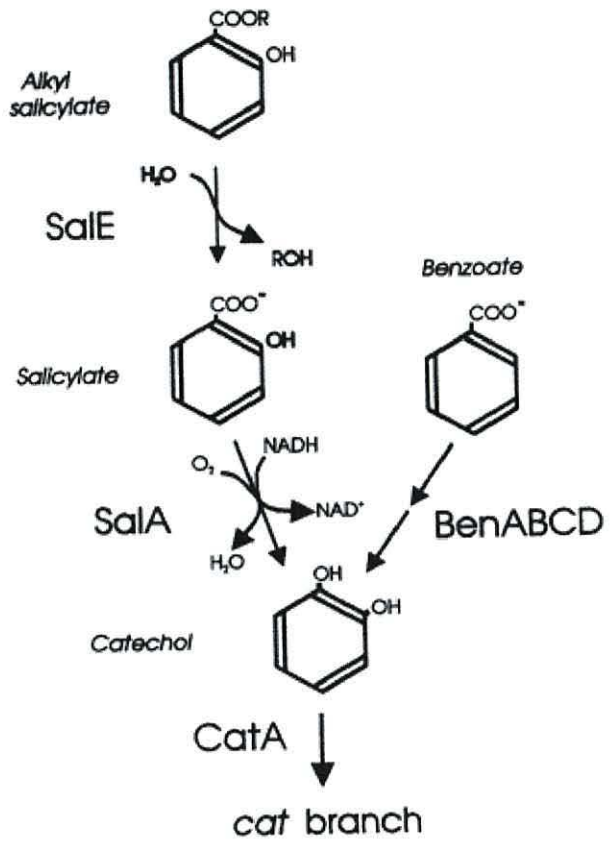


Fig.1.2.4 The reaction series of salicylate esterase (*salE*) and salicylate hydroxylase in *Acinetobacter* sp. strain ADP1 (Jones *et al.,* 1999).

dehydrogenase and *areC* encodes a benzaldehyde dehydrogenase (Fig.1.2.5) (Jones *et al.,* 1999).

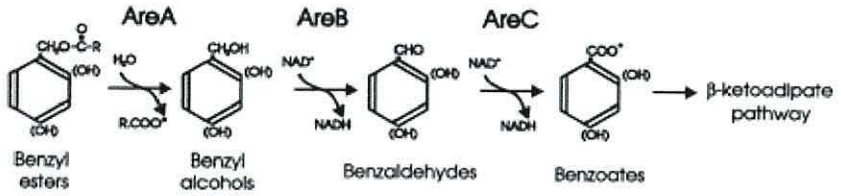


Fig.1.2.5. The proposed reaction series catalysed by the *areCBA* gene products serving to funnel benzyl alkanoates through the catechol branch of the β-ketoadipate pathway (Jones *et al.*, 1999).

The *benABC* genes encode a low specificity benzoate 1,2-dioxygenase system composed of NADH-cytochrome *c* reductase and terminal oxygenase components which convert benzoate to 2-hydro-1,2-dihydroxybenzoate. Consecutive in the sequence is *benD* encoding a *cis*-diol dehydrogenase that catalyzes the formation of catechol (Neidle *et al.,* 1987). In amino acid sequence comparisons, the *benABC* genes have been shown to have significant sequence identity with the broad specificity TOL plasmid pWW0 genes *xylXYZ* acting on methylbenzoates. This identity suggests a common ancestry for the two dioxygenases but that evolutionary divergence has occurred through acquisition of sequence repetitions via DNA strand slippage events that have been maintained by the mismatch repair system (Harayama *et al.*, 1989; 1991). Upstream of *benABCDE* (Fig.1.2.3) are the divergently transcribed LysR-type regulator gene *benM* and the *benPK* genes which are co-transcribed as an independent unit, encoding respectively, a putative porin and a benzoate permease transport protein belonging to the Major Facilitator Super family (MFS) (Collier *et al.*, 1997). The BenM LysR-type transcriptional activator

responds to both benzoate and *cis-cis*-muconate as inducing compounds which synergistically effect *ben* gene transcription (Romero-Arroyo *et al.,* 1995; Clark *et al.*, 2004).

The *cat* genes, *catAMBCDIJFD*, (Fig. 1.2.3) for catechol dissimilation through the β-ketoadipate pathway (Fig.1.1.6.2). The structural genes are transcribed in the same direction and are located downstream of the *ben* genes. The *catBCDIJF* genes are transcribed as a single unit (Neidle & Ornston, 1986). *catA* encoding catechol 1,2-dioxygenase is separated from the other structural *cat* genes by two open reading frames (ORFs) of unknown function comprising ~ 3 kbp (Shanley *et al.,*1986). The *cat* regulator gene, *catM*, is divergently transcribed from the structural genes and negatively regulates its own expression.

1.2.3 Regulation of the *ben-cat* gene cluster

1.2.3.1 LysR-type transcriptional activators (LTTRs)

The LTTR family was reported by Henikoff *et al.* (1988) to contain at least 9 homologous bacterial transcriptional activators. Since then the identity of many more LTTR family members has been established (Schell, 1993). The LTTRs are similar sized, auto-regulatory DNA binding proteins activating divergently transcribed linked-target genes and unlinked regulons possessing diverse functions. Common features identifying the family were elucidated in mutational analyses and amino acid sequence similarity studies and were shown to include a conserved helix-turn-helix motif (HTH) involved in DNA binding identified in many DNA binding proteins.

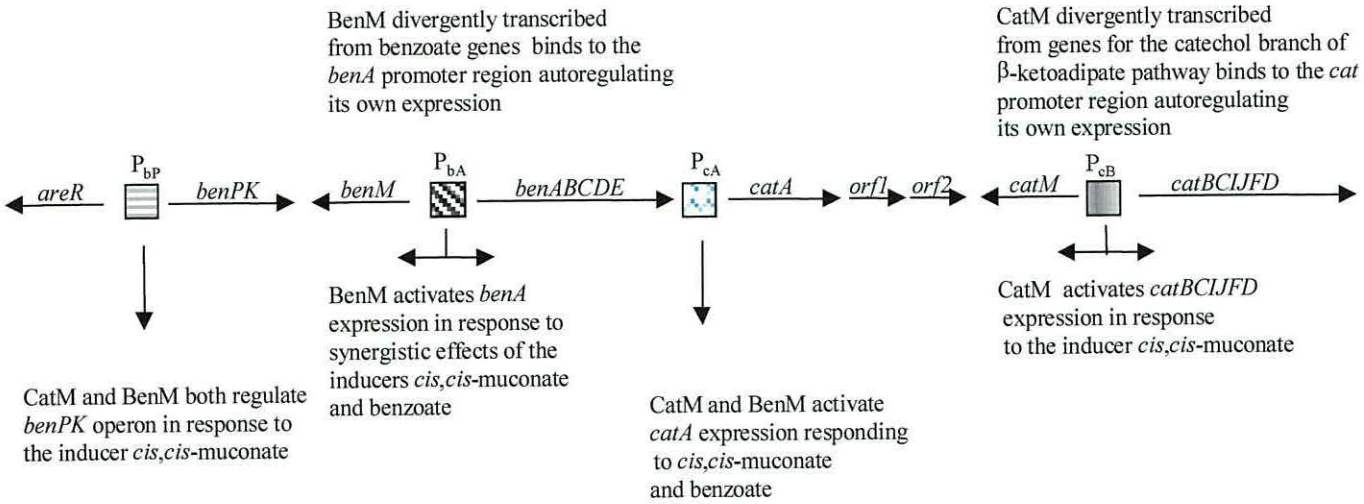


Fig.1.2.6. The roles of BenM and CatM in activating transcription and gene expression of the *ben-cat* genes in response to the inducing metabolites benzoate and *cis,cis*-muconate (figure adapted from Clark *et al.*, 2002).

The HTH motif is positioned between residues 1 - 65. Domains involved in co-inducer recognition and/or response (residues 100-173 and 196-206) are highly conserved as is a domain which is required for both DNA binding and co-inducer response (residues 227-253) (Schell, 1993). They have highly conserved sequences at the N-terminus of the proteins and significant sequence identity. In the absence of co-inducer many LTTRs bind to regulated promoters via a 15 bp dyadic sequence with a common structure and position. The co-inducer, when present in the cell, causes additional interactions of the LTTRs with sequences near the -35 RNA polymerase binding site and possibly causes DNA bending resulting in transcriptional activation (Schell, 1993).

BenM and CatM are LTTRs in ADP1. In concert with benzoate its coinducer, BenM, activates the transcription of the *benABCD* operon and also exerts transcriptional activation over *catA* (Harwood and Parales, 1996). CatM exerts transcriptional activation over *catA* and *catBCIJFD* in the presence of *cis,cis*-muconate, the product of the *catA*-encoded C1,20 reaction with catechol (Fig.1.2.6). Benzoate and *cis,cis*-muconate are inducing metabolites active in the expression of the *benABCDE, catA* and *catBCIJFD* genes (Neidle and Ornston, 1987; Collier *et al.,* 1998; Romero-Arroyo *et al.,* 1995). Strains lacking a functional BenM fail to support growth with benzoate as the sole carbon source. Cosper *et al.* (2000) demonstrated that in strains where *benM* was disrupted, spontaneous mutations in the *catB*-encoded muconate-cycloisomerase responsible for converting *cis,cis*-muconate to muconolactone allowed BenM-independent growth on benzoate. The mutations reduced muconate cycloisomerase activity during growth on benzoate, generating increased intracellular levels of *cis,cis*-muconate, stimulating CatM to activate *ben* gene expression in the absence of BenM. A complex regulatory circuit exists within the *ben-cat* supraoperonic operons where the regulatory proteins

BenM and CatM operate in a sequential cascade and have overlapping functions. Bundy *et al.* (2002) have shown synergistic effects on BenM activated gene expression exerted by its dual-inducer metabolites benzoate and *cis,cis*-muconate (Fig.1.2.6).

1.3 Genetic exchange in soil microbial populations

Comparative analysis of bacterial, archaeal and eukaryotic genome sequences suggest that a significant number of prokaryotic genomes have been acquired by horizontal gene transfer (HGT) (Koonin *et al.*, 2001). Bacterial evolution and adaptation to environmental constraints and niches may occur via point mutations, gene duplication events and gene re-arrangements (Reams and Neidle, 2003; Aber, 2000). However, in many instances the distribution of both chromosomal and plasmid genes is inconsistent with the concept of clonal inheritance. The availability of complete genome sequences provides sufficient unbiased coding sequences to compare open reading frames whose sequence characteristics differ significantly from the prevalent features of the resident genome, e.g. atypical nucleotide compositions and patterns of codon usage bias. This type of phylogenetic conflict infers HGT but does not necessarily provide unequivocal evidence for it and new analytical computer-based techniques such as phylogenetic reconstructions may better detect lateral acquisitions (Gogarten and Townsend, 2005). A review of lateral gene transfer in bacterial species by Ochman *et al.* (2000) estimated that in comparison with other known genome sequences available at that time approximately 16% of the *E. coli* protein encoding DNA had been acquired during evolutionary HGT events.

Bacteria possess highly efficient mechanisms of genetic exchange such as conjugation, transformation, transduction and transposition (Droge *et al.*, 1998;

Saye *et al.,* 1990). Conjugation requires cell to cell contact and is a process whereby plasmids or transposons transfer from donor to recipient cells. Many conjugative elements exhibit a broad host range for transfer and autonomous replication. Conjugative gene transfer is considered an important vehicle for the lateral transmission of acquired genetic ability and gene flux within bacterial populations (Davidson, 1999). The transfer of plasmids often provides the bacterial host with additional advantageous properties such as catabolic ability, heavy metal resistance and antibiotic resistance.

Transposition events mediated by transposons and insertion elements have played an important role in the evolution of catabolic plasmids (van der Meer *et al.,* 1992; Williams *et al.*, 2002). Integron-associated antibiotic resistance gene cassettes within plasmids have also been demonstrated to be widely distributed within activated sludge and the final effluents of a wastewater treatment plant (Tennstedt *et al.*, 2003). Bacterial transduction is the transfer of genetic material mediated by bacteriophages that have acquired host DNA during lysogenic infection. Bacteriophages exhibit a narrow host range of infection and this HGT mechanism is not considered to be significant in disseminating genetic material across species boundaries.

Conjugative plasmid transfer has been demonstrated *in situ* in different natural environments using natural isolates of both indigenous bacteria and plasmids (Lilley *et al.*, 1994; Bale *et al.*, 1988; Bjorklof *et al.*, 1995; Paul *et al.,* 1991; Richaume *et al.,* 1992; Nielsen *et al.,* 1994; Sandaa and Enger, 1994). A potentially recent HGT event (within 35 years) was detected in a coal-tar contaminated field site (Herrick *et al.*, 1997). PCR amplification identified homologues of the *P. putida* PpG7 naphthalene dioxygenase (*nahAc*) gene within divergent naphthalene degrading isolates. Sequencing of eight of these

isolates revealed complete sequence conservation of *nahAc* in six of the divergent strains indicating plasmid modification occurred after transfer that may have been effected by transposons.

Many plasmid gene transfer and transposition studies have been conducted in model soil or aqueous microcosm systems (Davidson, 1999; Droge *et al.*, 1998). In one such study (Daane *et al.*, 1996) the influence of earthworm activity on gene transfer was investigated. Depending on their ecological classification, earthworms were consistently found to positively influence the total numbers and diversity of indigenous soil transconjugants and to enhance the distribution of donor and recipient bacterial cells in soil. Transconjugant and donor bacteria were isolated from the earthworm soil casts and cocoons (worm embryo) suggesting these features potentially protect and sequester certain bacteria from the competitive soil environment. Soil casts and cocoons may also provide the resident bacteria with an environmental hotspot of gene transfer.

The ability of bacteria to undergo natural transformation i.e. the uptake and integration into the chromosome of naked DNA via homologous recombination and conjugative plasmid transfer are considered to be the primary sources of horizontal gene transfer within and across bacterial species boundaries. Natural transformation is also considered to be a potentially active mechanism for interdomain gene transfer events.

1.3.1 Natural transformation

Natural transformation is a widespread phenomenon among bacterial species and is defined as a genetically programmed physiological state allowing the

efficient uptake of extraneous macromolecular DNA and the incorporation of this DNA into the genome of the transforming bacteria through a genetic recombination event (Dubnau, 1999). The term "bacterial competence" is used to describe the process. Both Gram-negative and Gram-positive species undergo natural transformation although the process differs among the different species. In laboratory experiments more than 50 bacterial species from different environments are known to be naturally transformable (Lorenz and Wackernagel, 1994) and a recent article by de Vries and Wackernagel (2005) lists 87 species capable of natural transformation. However, the true number of naturally transformable species is unknown and many may have been overlooked as competence induction may have specific requirements not provided in *in vitro* studies (Demaneche *et al.*, 2001b). The pathogenic *Leigionella pneumophilus* was considered to be non-competent for natural transformation for >20 years. The discovery of unusual requirements for competence induction led to the identification of two strains highly competent for natural transformation (Stone and Kwaik, 1999; Sexton and Vogel, 2004). Laboratory strains of *E. coli* have also been demonstrated to be competent for natural transformation under environmental conditions when in contact with surface water originating from calcareous regions. Concentrations of calcium > 1mM often found in springwater and river water are sufficient to induce natural competence in *E. coli* (Baur *et al.*, 1996). *Bacillus subtilis, Streptococcus pneumoniae, Mycobacterium* and *Streptomyces* species are naturally competent organisms representative of Gram-positive species. *Haemophilus influenzae*, *A. baylyi* strain ADP1, *Neisseria gonorrhoeae, Azotobacter, Helicobacter pylori, Ralstonia solanacearum*, *Pseudomonas stutzeri* and the thermophilic *Thermus* spp. are representatives of competent Gram-negative species (Lorenz and Wackernagel, 1994; Bertolla *et al.*, 1997; Koyama *et al.*, 1986). The process of natural transformation can be divided into four distinct phases: competence

induction, DNA binding, DNA uptake, and integration of the internalized DNA into the resident chromosome.

1.3.1.1 Competence induction

The first stage of natural transformation demands induction of a competent physiological state. Depending on the bacterial species this competent state may affect all or a subset of cells within a population, may be growth phase specific, under nutitional control or both or may be constitutively expressed. In the Gram-positive species *Streptococcus pneumoniae* and *Bacillus subtilis* natural transformation occurs at entry to and during stationary phase, regulated by a soluble competence stimulating peptide (CSP), which accumulates extracellularly, in a quorum sensing-dependent manner. CSP accumulation to a critical level, triggers development of competence and induction of a set of new proteins, several of which are membrane localised (Dubnau, 1991; Clavery *et al.,* 2000). *H. influenzae*, *Pseudomonas stutzeri* and *Azotobacter vinelandii* also develop competence at the entry to stationary phase although no CSP has been identified and growth conditions appear to trigger competence induction (Dubnau, 1991; Khan and Smith, 1984). *A. baylyi* strain ADP1, *Thermus thermophilus*, *L. pneumophila* and *Staphylococcus aureus* enter a maximally competent state between the early to late log phase of growth. *Neisseria gonorrhoeae* and *Helicobacter pylori* are constituitively competent for natural transformation (Biswas, 1977; Israel *et al.*, 2000).

1.3.1.2 DNA binding and uptake

Two models have been devised to interpret the mechanism of DNA binding and uptake in Gram-positive and Gram-negative bacterial systems. In Gram-positive bacteria the DNA must traverse the cell wall and the cytoplasmic membrane. In the Gram-negative system the DNA must also be transported through the outer membrane and the periplasmic space. The differences in cell envelope require differences in the DNA binding and uptake machineries, although this is not absolute, as Gram-negative systems have been identified that more closely resemble the Gram-positive uptake process than the generalized Gram-negative one (Dubnau, 1999; Averhoff, 2004). In the Gram-positive system double stranded DNA binds to the cell surface and undergoes fragmentation. In *S. pneumoniae* single-strand nicks occur first followed by double-strand breaks facilitating the linear uptake of single-stranded (ss) DNA from newly formed ends. The DNA becomes DNase resistant and is transported into the cytosol while the non-transforming strand is degraded. The donor DNA is integrated by forming a heteroduplex complex with a homologous sequence of the recipient chromosome (Dubnau, 1999).

In the Gram-negative model, DNA uptake proceeds linearly after double strand cleavage. Acquisition of DNase resistance does not occur concomitantly with the formation of ss but probably corresponds to entry into the periplasmic compartment or a specialised structure. In *H. influenzae*, double-stranded DNA appears to associate with discrete membrane vesicles termed transformasomes, an association which protects the exogenous DNA from DNase action (Khan *et al.*, 1982). The DNA is then transported into the cytoplasm, at which point it undergoes conversion to the ss DNA form and is integrated via homologous

recombination whilst the non-transforming strand is degraded (Dubnau, 1999). DNA uptake in *H. influenzae* and *N. gonorrhoeae* requires the presence of specific recognition sequences of 9 and 10 bp respectively, often found as inverted repeats between genes acting as transcriptional terminator sequences (Goodgal, 1982; Goodman and Socca, 1988).

1.3.1.3 DNA integration

The integration of the incoming ss DNA into the resident chromosome is a RecA-dependent process and requires the presence of a functional recombination system. *recA* is inducibly co-expressed with the competence genes of Gram-positive species (Palmen and Hellingwerf, 1997). In ADP1 and *P. stutzeri* RecA expresion is not affected by competence induction but is apparently constitutively expressed at levels high enough to allow DNA integration (Rauch *et al.*, 1996; Vosman *et al.*, 1993).

1.3.1.4 Transformation in *A. balyii* strain ADP1

The natural transformation system in ADP1 occurs maximally in exponential growth phase after dilution of an overnight stationary phase culture, and is growth phase-dependent, (Palmen *et al.,* 1993; Nielsen *et al.,* 1997; Friedrich *et al.,* 2001). The uptake of DNA is an energy-requiring process with a requirement for divalent cations, such as Mg^{2+}, Ca^{2+} and Mn^{2+}, as is the case for *B. subtilis*, *S. pneumoniae* and *Thermus thermophilus* (Lorenz and Wackernagel 1994; Koyama *et al.*, 1986). The process of DNA uptake and/or integration in ADP1 is inhibited at a pH value below 6.5. In comparison with other competent Gram-negative bacteria that recognize specific DNA binding sequences, ADP1

does not discriminate between heterologous and homologous DNA and takes up DNA from any source (Lorenz *et al.,* 1992: Palmen *et al.,* 1993). In this respect, the ADP1 transformation uptake system resembles that of Gram-positive bacteria (Dubnau, 1999; Claverys *et al.,* 2000). The transforming ss DNA is integrated into the resident chromosome via homologous recombination (Palmen *et al.*, 1993; Porstendorfer *et al.,* 1997).

1.3.1.5 Competence proteins in *A. baylyi* strain ADP1

Several conserved competence genes and proteins have been identified (Hahn *et al.,* 1987; Inamine and Dubnau, 1995; Karudapuram and Barcak, 1997; Fusseneger *et al.,* 1997; 1996). Many of these proteins resemble a widespread group required in Gram-negative bacteria for the assembly of type IV pili, the type II secretion pathway and twitching motility (Hobbs and Mattick, 1993; Dubnau, 1999; Averhoff, 2004). In ADP1, six putative competence factors have been identified, encoded by *comP*, *comC*, *comB*, *comE*, *comF* and *comA*. ComB,-E,-F and P resemble prepilins of type IV pilin-like components of protein translocation systems and are membrane-associated, although studies show they are not essential for pilus production or twitching motility. ComC resembles type IV pilus biogenesis proteins and has large regions of hydrophilic amino acids. ComC and ComP are required for DNA binding and uptake. ComP has a peptide leader sequence nearly identical to the those of pilins and pilin-like proteins involved in secretion (Porstendorfer *et al.,* 1997; Link *et al.,* 1998). *comE* and *comF*, are located 61 bp downstream from *comC* and are specific for natural transformation, since mutations in these genes affected neither pilation nor lipase secretion but resulted in decreased transformation frequencies (Busch *et al.,* 1999). ComB exhibits conserved

endopeptide cleavage motifs of prepilin proteins and has elevated expression levels during prolonged exponential and stationary growth-phases (Herzberg *et al.*, 2000). ComA displays significant similarities to the competence proteins ComA and ComEC in *Neisseria* and *Bacillus* spp. involved in the transport of transforming DNA. That ComA is part of a competence regulon is suggested by correlation of *comA* transcription with growth phase-dependent transcriptional regulation of the pilin-like factors of the transformation machinery (Friedrich *et al.*, 2001). Homologues of type IV pilus subunits are involved in natural transformation in a wide variety of bacteria and generally function synergistically to facilitate DNA uptake.

1.3.2 Evolution of natural transformation

Natural transformation has been suggested to have evolved as a food scavenging mechanism, a system for DNA repair to counterbalance the effects of nutritional deprivation, oxidative or heat shock damage, UV radiation, or as a mechanism to ensure genome plasticity and diversity in highly competitive ecological niches (Dubnau, 1999; Finkel and Kolter, 2001). The acquisition of new traits for niche colonization and species fitness is provided by the ability of naturally transformable species to sample the available gene pool, successfully integrating, expressing and maintaining beneficial DNA acquisitions. Competence may have evolved in different species to serve different requirements and purposes. Diversity of function could then explain the diversity of competence genes, regulation and induction mechanisms identified within the broad range of naturally transformable species.

1.3.3 Natural transformation in the environment

Natural transformation in the environment has been demonstrated in soil extracts, marine and freshwater habitats and aquifer material (Chamier *et al.*, 1993; Lorenz *et al.*, 1988; 1992; Lorenz and Wackernagel, 1990; Romanowski *et al.*, 1993). Large quantities of high molecular weight exogenous DNA is present in soils and is an important component of organic phosphorus pools. The DNA may originate from cell lysis of microorganisms, plants and animals, from root exudates and DNA excreted from microorganisms themselves (Lorenz and Wackernagel, 1994). DNA persists in the environment through adsorption with the particulate constituents of soils and sediments (Lorenz and Wackernagel, 1994; Ogram *et al.*, 1988). Chamier *et al.*, (1993) showed the viability of bound DNA in a soil or sedimentary environment for natural transformation of ADP1 and that both chromosomal and plasmid DNA adsorbed to non-sterile aquifer material and sterile sea sand. Transformation efficiencies of chromosomal DNA were similar in sterile and non-sterile sand; however, plasmid DNA was not transformed until loaded onto the system with the chromosomal DNA. They also provided evidence of natural transformation occurring at the liquid-solid interface. Demaneche *et al.* (2001a) demonstrated that plasmid DNA adsorbed onto clay particles was physically protected from degradation by nucleases. The biological potential of the DNA was shown to be only partially available for transformation, while the remaining fraction of DNA was sequestered by clay adsorption, rendering it inaccessible to both nucleases and competent *Acinetobacter* cells. DNA transport and transformability in water-saturated soil columns showed a potential for biologically active DNA to be transported over a considerable distance without nuclease degradation (Pote *et al.*, 2003). Dead cells are also considered to protect the DNA against DNases in soil. Cell lysates of *P. fluorescens,*

Acinetobacter sp. and *Burkholderia cepacia* were able to transform *Acinetobacter* sp. populations in non-sterile soils for a period of days and cell debris prevented DNA degradation (Nielsen *et al.,* 2000a). The long-term persistence of DNA released from senescent or rotting transgenic plant material under field conditions has also been observed (Gebhard and Smalla, 1999).

1.3.4 Environmental factors affecting ADP1 transformation

Habitat, soil type, environmental conditions such as temperature, pH, nutrient level and flux are all variable parameters affecting transformation in the environment. Nielsen *et al.* (1997) investigated the factors affecting natural transformation of ADP1 using homologous chromosomal DNA in a silt loam microcosm. With the addition of a simple carbon source, salts and freshly added DNA, inducible transformation was detected for 6 days in declining populations of non-competent ADP1 cells. Environmental parameters such as variable ammonium levels and divalent cations had no real effect on transformation efficiencies, whereas increasing the levels of phosphate salts and the moisture content of the soil significantly enhanced the transformation efficiency. Soil moisture content was optimal for transformation at ~ 35%. Abiotic factors and access to nutrients were critical to ADP1 transformation in a silt loam microcosm. Organic compounds, such as organic acids, amino acids and sugars, naturally occurring in crop plant rhizospheres were reported to have a stimulatory effect on *Acinetobacter* sp. strain BD413 (pFG4) transformation. In sterile soil, most of the compounds induced competence, whereas in non-sterile soil added alanine, lactate, acetate and glucose in the presence of high supplements of phosphate salts produced transformants. Mixtures of these compounds added to the non-sterile soil produced the highest transformation

frequencies, providing evidence of competence induction by added nutrients (Nielsen and van Elsas, 2001).

1.3.5 HGT - transgenic plants to bacteria

The ability of bacterial species to undergo natural transformation in the environment has raised concerns about HGT occurring between GMOs and the natural microbial population. The issue of natural selection and continuation of transgenes in the bacterial genome is not certain (Smalla *et al.,* 2000). HGT from transgenic plants to indigenous soil microbial populations via natural transformation leading to genetic integration, expression and selection of transgenes and the potential environmental and health impacts of such an event are under debate. The occurrence of multiple antibiotic resistance "superbugs" is a severe problem in the hospital environment in relation to nosocomial infection. HGT of antibiotic resistance transgenes between plants and bacteria in soil or the commensal bacteria of animals and humans is a potential hazard associated with genetically modified plants using antibiotic resistance marker genes as an aid to transformant selection (Bergmans, 1993; Nielsen *et al.,* 1998; Duggan *et al.*, 2000; Smalla *et al.*, 2000). However, it is noted that in the soil environment many antibiotic resistance determinants employed as markers are widely distributed and may not present a new and unique addition to the current gene pool in this particular environment.

Early studies to detect HGT via natural transformation from transgenic plants to phytosphere bacteria such as *Agrobacterium tumefaciens* (Broer *et al.*, 1996) *Erwinia chrysanthemi* (Schluter *et al.*, 1995) and indigenous soil bacteria (Smalla *et al.*, 1994; Paget and Simonet, 1994) failed to detect transformant bacteria in all but one report. Transgenic *Brassicaceae* spp. containing a pUC18

derivative conferring hygromycin B resistance, was co-cultivated with *Aspergillus niger* in soil and one fungal transformant was obtained containing pUC18 sequences (Hoffman *et al.*, 1994). This general failure to detect HGT events may have been due to the absence of homologous sequences present in donor and recipient DNA or use of less efficiently transformable bacteria. The first occurrence of HGT from transgenic plants to bacteria under optimized laboratory conditions was reported by Gebhard and Smalla (1998) using transgenic sugar beet DNA containing a complete *nptII* gene conferring kanamycin resistance as the DNA donor. A strain of ADP1 containing a plasmid harbouring a null *nptII* engineered to contain a 317 bp deletion was used as the recipient. Restoration of the antibiotic resistance determinant detected HGT to the recipient bacterium. Artificial homology created between donor and recipient DNA demonstrated that recombination, integration and expression of transgenic plant DNA could, under optimized transformation conditions, transform bacteria.

DNA transfer from transgenic plants to bacteria has been studied using *nptII* assays involving recombinational repair of an internal 10 bp deletion within recipient *Acinetobacter* sp. (de Vries and Wackernagel, 1998; Tepfer *et al.*, 2003). The system provides a highly specific mechansim of monitoring the spread of recombinant DNA in the soil environment. Widespread dissemination of DNA by transgenic potato plant roots and pollen in field experiments was detected when *Acinetobacter* cells with a specific *nptII'*::tg4 fusion (truncated *nptII* fused to a eukaryotic terminator) plasmid were used as the homologous screen. Restoration of a deletion within this fusion restored kanamycin resistance. The presence of the downstream tg4 sequence ensured recombination with specific donor DNA that contained the unadulterated

fusion. No transformants were detected with transgenic donor DNA containing only the *nptII* allele in different genomic contexts (de Vries *et al.*, 2003).

A marker rescue assay procedure was utilized by Ceccherini *et al.* (2000) to monitor the degradation and transformability of DNA released from transplastomic leaves (where the chloroplast contains the transgenes). Transplastomic leaves were exposed to degradation conditions simulating those encountered in nature to study the fate of plant DNA after plant decay but prior to DNA entering the soil. Recipient bacteria carried a plasmid containing plastid sequences to facilitate homologous recombination, transformants were identified by selective plating and acquisition of *aadA* encoded streptomycin/spectinomycin antibiotic resistance. Plant leaves subjected to grinding or treatment with cellulytic enzymes showed a loss in transformation frequency of approx. 99.8% after 3 days corresponding to an observed decrease in *aadA* sequences as a result of degradation.

Conditions favouring HGT could occur *in planta* in bacteria that have developed symbiotic or pathogenic associations with transgenic plants. To investigate this, transplastomic tobacco plants were co-colonized with the plant pathogen *Ralstonia solanacearum* and *Acinetobacter* sp. strain BD413 (ADP1). *Acinetobacter* sp acts as an opportunistic organism co-colonizing plants infected by *R. solanacearum* and is able to develop a competent state *in planta*. HGT from transplastomic DNA *in planta* to *Acinetobacter* was observed in the presence of plasmid-borne sequences homologous to chloroplastic DNA. The higher transgene copy number present in transplastomic DNA (approx. 10,000 copies per cell) facilitated higher transformation frequencies than previously observed using nuclear-modified plants which contain approx 10 copies per cell (Kay *et al.*, 2002). In all the above reports no HGT transformants were

detected in the absence of the homologous sequences provided by the transformation assay. de Vries *et al.* (2001) reported that integration by natural transformation of transgenic plant DNA in *Acinetobacter* sp and *P. stutzeri* depends on homologous sequences in the recipient cells, suggesting that quite large sequences of DNA homology are required for bacteria to incorporate transgenes within the resident genome. However, illegitimate recombination of non-homologous DNA has been observed with plasmid DNA and plastid DNA from transgenic plants. Integration occurs when it is flanked by homologous sequences at hotspots for illegitimate recombination, consisting of GC-rich microhomologies of 3-6 bp in size (de Vries and Wackernagel, 2002; Meier and Wackernagel, 2003; de Vries *et al.*, 2004).

1.4 Aims of study

Previous studies of HGT from transgenic plants to indigenous soil bacteria, as outlined above, have utilized antibiotic resistance determinants to identify and quantify positive transformation events. The purpose of this PhD thesis was to quantitatively study the occurrence of a specific gene transfer event between *A. baylyi* strain ADP1 and closely related species using a non-antibiotic screening method. Detection of gene transfer was demonstrated by regain of gene function i.e. the ability to catabolize salicylate and utilize this compound as a sole carbon and energy source. This study was performed in association with a microcosm study conducted by a UW Aberystwyth group investigating potential environmental hotspots for gene transfer in soil. The worm gut and/or the nutritionally rich soil casts of worms are considered to serve as environmental hotspots for horizontal gene transfer. Through the natural feeding process of worms, bacteria are thought to be sequestered from the bulk

soil in the presence of plant debris. The close proximity of internalized bacteria and DNA released from transgenic plants within the nutritionally rich microhabitat of the worm gut or soil casts may provide a permissive environment wherein the induction of natural competence for genetic transformation could occur. DNA released from digested plants could be taken up and incorporated into the bacterial genome. Transgenic tobacco plants containing the ADP1 *salA* gene were fed to worms in a soil microcosm and the ability of the transgenic DNA to transform an ADP1 strain deficient in salicylate catabolic function to prototrophy was examined. My role in this study was to identify and isolate homologous *salA* alleles from environmental isolates of the diverse genus *Acinetobacter* able to grow with salicylate and quantify the frequency of transformation between these and *salA* mutated strains of *A. baylyi* strain ADP1. These transformation assays were performed under laboratory conditions and sought to provide information on the possibility of background transformation that may occur within soil or during/after worm digest via gene exchange between *salA* homologues of closely related species.

Five *salA*$^-$ knock-out strains were created incorporating small 4 bp base pair deletions within separate regions of the gene. Restoration of salicylate hydroxylase activity detects natural transformation with homologous genes from related donor *Acinetobacter* spp. creating a non-antibiotic based system to detect gene transfer events. Thirty-four *Acinetobacter* environmental isolates were screened for salicylate 1-hydroxylase activity. Strains with a *salA*$^+$ phenotype were selected and four *salA* genes were isolated using degenerate PCR or shotgun cloning techniques. Shotgun cloning isolated two DNA fragments from two different *Acinetobacter* strains, both of which contained highly similar putative *sal* operons. One such clone, isolated from *Acinetobacter* sp. strain AD3-1, underwent high quality sequencing and was

characterized in detail. Transformation frequency assays were performed using this as strain as the donor DNA.

Chapter 2 section 1

2.1 Chemicals and media

2.1.1 Chemicals and reagents

Standard laboratory chemicals were purchased from Sigma Aldrich, Lancaster Synthesis Ltd., and VWR International Ltd. Microbiological growth media, LabM™ was purchased from International diagnostics group plc (idg™). DNA manipulation materials were purchased from Promega, Qiagen and GibcoBRL Life technologies. The 1 kb Generuler™ DNA molecular weight marker used in all DNA electrophoresis was purchased from Fermentas. Hi-Trap metal chelating affinity columns and nucleic acid labelling and detection kits were purchased from Amersham Pharmacia Biotech. The protein molecular weight marker - Precision plus protein™standard was purchased from Bio-Rad, as were 30% acrylamide/Bis solution and Zeta-Probe® GT blotting membranes.

2.1.2 Bacterial growth media

Bacteria were grown using liquid or solid media for the desired growth conditions. Luria Bertani (LB) media was used (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl) for growth of *Escherichia coli* in liquid or solid media. When required for the selection and identification of salicylate hydroxylase-bearing clones, the growth of *Escherichia coli* DH5α in LB medium included salicylate, at a concentration of 1 mM. For rich or minimal solid media, Agar No.1 was added (15 g L^{-1}). For selection procedures involving the

chromosomal *sacB*-Km cassette in strains of ADP1 carrying this marker, sucrose 5% (w/v) was added to LB agar.

For bacterial growth in liquid mineral salts medium (Bauchop & Elden, 1960) defined carbon sources were added at the following concentrations: succinate 10 mM, salicylate 2.5 mM, benzoate and *p*-hydroxybenzoate 2.5 mM. For solid minimal growth medium, non-volatile carbon sources were added to molten agar after heat sterilization at the above concentrations. For the identification of salicylate hydroxylase activity *Acinetobacter* spp. were streaked onto 2.5 mM and 1.5 mM salicylate minimal medium. Growth on ethyl salicylate was achieved by plating onto minimal agar medium with the volatile carbon source provided in small tubes in the inverted lid of petri dishes. When appropriate, antibiotics were included at the following concentrations, ampicillin (Ap) 100 $\mu g\ ml^{-1}$; kanamycin (km) 50 $\mu g\ ml^{-1}$ for *E. coli* or 10 $\mu g\ ml^{-1}$ for *Acinetobacter baylyi* strains.

2.1.3 Bacteria and growth conditions

All bacterial strains and plasmids utilized and constructed are listed in the individual chapters of this thesis.

2.1.3.1 *Escherichia coli*

For the routine growth of *E.coli* DH5α and *E.coli* Rosetta (DE3 lysogen), cultures were incubated at 37°C with shaking for liquid culture at approx. 160 rpm. *E.coli* Rosetta (DE3) carrying the expression vector p196EXA was grown under optimized conditions for protein expression, (see protein purification section).

2.1.3.2 *Acinetobacter* species

For the routine growth of *Acinetobacter baylyi* strain ADP1 cultures were incubated at 37°C, with shaking for liquid cultures at approx. 160 rpm. *Acinetobacter baylyi* strains ADPW251 and ADPW249 containing the *sacB*-Km cassette were grown in LB-km medium (10 μg ml^{-1}) at room temperature to allow expression of the *sacB* phenotype as the sucrose-sensitive phenotype is temperature-dependent in ADP1. All other *Acinetobacter* spp were incubated at 30°C under the same aeration conditions.

Section 2

2.2 DNA manipulation techniques

2.2.1 Competent cell preparation

Competent cells of *E. coli* strains DH5α and Rosetta were prepared according to the method of Sambrook *et al.* (1989). A single colony was selected from a fresh LB plate and inoculated into 10 ml LB liquid medium, grown overnight at 37°C and sub-cultured (1:100) into a final volume. The sub-culture was incubated at 37°C to an $OD_{550\ nm}$ of approx. 0.35. The cells were chilled on ice for 5 mins then centrifuged at 8 000 x g for 5 mins at 4°C. The supernatant was removed and the cell pellet was resuspended in 2/5 volumes of competent cell Buffer 1: potassium acetate, (30 mM), rubidium chloride, (100 mM), calcium chloride, (10 mM), manganese chloride, (50 mM) and glycerol, 15% (v/v). The resuspended cells were chilled on ice for 5 min then centrifuged as before. The

supernatant was removed and the cell pellet was resuspended in competent cell Buffer 2: MOPS (10 mM), calcium chloride, (75 mM), rubidium chloride (10 mM) and glycerol 15% (v/v). After a 15 min period on ice the cells were aliquoted (200 µl) into pre-chilled microcentrifuge tubes and stored at -70°C.

2.2.2 Plasmid transformation

Plasmids were transformed into *E.coli* DH5α or Rosetta strains as part of cloning experiments and/or routine plasmid stock maintenance as follows: either 1 µl (for routine plasmid maintenance) or 10 µl (for cloning experiments) was added to 200 µl of competent cells and kept on ice for 30 min. Heat shock treatment at 42°C for 50 s was followed by a further 2 min on ice. LB broth (400 µl) was added and the cells were incubated at 37°C for 45-60 min to induce cell recovery and antibiotic resistance gene expression. The culture was then plated in suitable volumes onto selective media and incubated at 37°C overnight.

2.2.3 Plasmid isolation and screening

For plasmid isolation, single colonies were selected and inoculated into 5 ml LB broth with the appropriate antibiotics and grown overnight at 37°C. Plasmids were isolated using the alkaline lysis Qiaquick spin miniprep kit or the GibcoBRL Concert™ High purity plasmid miniprep kit according to the manufacturer's instructions. Plasmids were stored in UP H_2O at -20°C. Strains containing plasmids were stored at -70°C in glycerol 70% (v/v). The isolated plasmids were screened using appropriate restriction endonucleases and visualized by ethidium-bromide agarose gel electrophoresis (AGE).

2.2.4 Restriction enzyme DNA digestion

Restriction endonuclease digestion of plasmid DNA was performed using approx. 0.2 - 1.0 µg of DNA, according to the manufacturer's instructions and incubated for 45-60 min at 37°C. Restriction digests of genomic DNA were incubated for approx. 18 h at 37°C.

2.2.5 Agarose gel electrophoresis (AGE)

Plasmid DNA analysis was performed using ethidium bromide AGE. Agarose gels 0.7% (w/v) were prepared with 20% (v/v) 5x TBE buffer. Running buffer contained 10% (v/v) 5x TBE buffer. For the analysis of genomic DNA, gels were prepared with the same buffer constituents at 1% (w/v) agarose.

2.2.6 Genomic DNA isolation

Genomic DNA was isolated by the hexadecyl trimethyl ammonium bromide method of Ausubel *et al.* (1989) or by the use of either the Qiagen DNeasy® Tissue Kit or the Promega Wizard® Genomic DNA purification kit. Purified genomic DNA was stored in TE buffer (pH 8.0) at 4°C. DNA quantification was performed in quartz cuvettes and measured spectrophotometrically using the absorbance A260/A280 ratio on a Pharmacia Biotech Ultrospec® 2000.

2.2.7 Polymerase chain reaction (PCR)

PCR procedures were performed using a Techne® PCR machine with PCR kits from various manufacturers. These comprised Qiagen: Hotstar Taq™

polymerase, Hotstar master mix® and Proofstart ™ polymerase, Roche:- Expand High Fidelity PCR system and Promega: *Pfu* DNA polymerase and PCR master mix. Procedures were carried out according to the manufacturer's instructions and at optimized PCR conditions. All PCR reactions contained a no template control using the same solutions as the experimental reactions to ensure product integrity.

PCR mutagenesis techniques were utilised for the creation of specific restriction sites within DNA sequences and the creation of engineered mutations. For the creation of restriction sites, forward and reverse primers were designed to allow wild type annealing at the 5' and 3' terminal primer ends with the mutation centrally located within the primer sequence.

Overlap extension PCR mutagenesis (OEPM) incorporates engineered mutations within a gene sequence. Two sets of forward and reverse primers, overlapping in sequence at the mutation site are required for the first step of mutagenesis. One primer of each pair codes for the desired mutation, e.g. the reverse primer of set 1 and the forward primer of set 2. PCR reactions are performed separately for each pair. The products from these reactions were mixed and a PCR reaction was performed using the two products as template, with the forward primer from set 1 and the reverse primer from set 2 as external primers. The two products self-annealled at the overlapping region and were amplified to produce a single PCR product carrying the desired mutation.

2.2.8 Primer synthesis, DNA sequencing and analysis

Primers were designed using the Lasergene software (DNAstar), and synthesized by MWG Biotech Ltd, Ebersberg, Germany. Sequencing was

performed by MWG Biotech and the results were analysed using Lasergene software (DNAstar) and the NCBI database, blastx. DNA and amino acid alignments were analysed using the Lasergene software (DNAstar - Megalign package), the NCBI clustalx package and the Genedoc multiple alignment software.

2.2.9 Cloning procedures

DNA fragments required for cloning were either excised from agarose gels by band extraction and purified using the Qiaquick Gel extraction kit or purified directly from a PCR reaction mixture using Qiaquick PCR purification kit according to the manufacturer's instructions.

The thermostable *Taq* DNA polymerase operating in Hotstar™ *Taq* polymerase, Hotstar Master mix, the Expand High Fidelity PCR system and Promega Master mix kits has terminal transferase activity and preferentially adds extra A nucleotides to the 3' ends of PCR products. This activity allows the cloning into TA vectors such as pGem®-T-easy system I (Promega) which contain complementary 3' overhangs. The insertion of the PCR product interrupts the vector encoded β-galactosidase activity, therefore in the presence of X-Gal, a chromogenic blue-white selection process can be followed. The proofreading DNA polymerases, (DNAP), Proofstart polymerase and the *Pfu* DNAP, lack this terminal transferase activity and produce blunt-ended PCR products which require poly-A-tailing procedures for cloning into TA vectors. Poly-A-Tail addition can be achieved by a 5-15 min incubation at 72°C of the PCR product with *Taq* DNAP, dATP (2 mM), and $MgCl_2$. Alternatively, blunt-ended PCR products can be engineered to contain terminal restriction sites

allowing direct digestion of the PCR product and ligation into the multi-cloning site of a similarily prepared vector.

2.2.10 Ligations

The ligation of DNA fragments into plasmid vectors was performed using T4 DNA ligase. Cohesive-end TA cloning into pGem®-T-easy proceeded according to the manufacturer's instructions. The cloning of DNA fragments into the pUC19 or pET28a vectors involved ligation of lysophilized DNA fragments and restriction-digested vector DNA. The DNA, resuspended in UP H_2O and T4 DNA ligase plus T4 DNA ligase buffer to a total volume of 10 µl was incubated at 16°C overnight (approx.18 h) for cohesive-end ligations. Blunt-ended ligations, where necessary, were treated in the same way but incubated overnight at 22°C.

2.2.11 Dephosphorylation and frameshift deletions

The dephosphorylation of 3' -OH DNA ends using calf intestinal alkaline phosphatase (CIAP) was performed to minimize the self-ligation of plasmids during cloning experiments, according to the manufacturer's instructions and standard procedures (Sambrook *et al.*, 1989).

Frameshift deletions were introduced into gene sequences with unique restriction sites by two methods: a 'fill in' deletion and a 'chew back' deletion. When a restriction digest leaves complementary overhanging ends it is possible to frameshift the sequence by the addition of dNTPs and T4 DNA polymerase. The dNTPs form complementary base-pairs with the overhanging ends and are joined to the sequence by the polymerase reaction, thus 'filling in' the gap and

deleting the restriction site. Restriction digests that produce blunt-ended DNA fragments can also undergo frameshift mutations, which can be introduced using exonuclease III and S1 nuclease. Exonuclease III is a double stranded 3' exonuclease used in the stepwise removal of nucleotides from blunt ended DNA leaving overhanging termini. The rate of nucleotide removal is temperature dependent, approx. 60 bp/min at 22°C. Exonuclease III activity followed by S1 nuclease treatment forms blunt-ended DNA fragments which can be ligated together, forming the frameshift deletion (Promega, 1996).

2.2.12 Shotgun cloning

Genomic DNA isolated from *Acinetobacter spp.* with a salicylate positive (*salA*$^+$) phenotype was digested with *SacI* or *XbaI-SacI* restriction enzymes. The genomic digest reaction was purified using the PCR purification kit (Qiagen) and mixed with similarily digested and purified dephosphorylated pUC19 DNA and ligated overnight at 16°C. The ligation mixture was transformed into *E.coli* DH5α and plated onto LBA-salicylate (1 mM), (Ap 100 μg ml^{-1}), selective media. After incubation at 37°C the transformant plates were placed at room temperature for up to 2 weeks and monitored for the appearance of a brown discolouration of the media indicating catechol accumulation and the presence of a cloned *salA* gene.

2.2.13 Southern hybridization

Southern hybridizations were performed using ECL direct nucleic acid labelling and detection system RPN 3000™ and Chemiluminescent CDP-*Star* detection kit™, (Amersham Pharmacia Biotech). Digested genomic (~10 μg) or plasmid DNA (~ 0.002 - 0.01 μg) were loaded onto 1% (w/v) agarose gels and

were subjected to overnight electrophoresis. The gels were processed prior to blotting with washes in depurination, denaturation and neutralization solutions according to the ECL protocol. Zeta-Probe® GT blotting membrane was cut to size and used as the DNA matrix. Capillary blotting of the gel was performed overnight in 20X SSC with a Whatman 3MM paper wick. The blots were processed according to the protocols and hybridized in a rotary tube at either 42°C for the ECL kit or at 55°C for the CDP-*Star* kit. The probe (~ 10 ng/µl) was prepared using band excised digested plasmid DNA fragments or purified PCR products and labelled according to the specific protocol. After overnight hybridization the blots were washed in primary and secondary wash buffer and chemiluminescence was induced by incubation with the signal detection reagent. The blot was exposed onto autoradiography film for the recommended length of time and developed. For some blots it was necessary to increase the exposure time to 24 h achieve any signal detection.

Section 3

2.3 Natural transformation assays

2.3.1 Natural transformation of *Acinetobacter baylyi* strain ADP1

For initial transformations of ADP1, overnight 1 ml LB cultures were diluted 1/20 into fresh LB media containing ~1 µg of digested plasmid DNA, grown overnight at 37°C and plated onto selective media. Single colonies were selected and screened as appropriate for the specific mutation.

2.3.2 Chromosomal gene knockout and genetic marker eviction

The natural transformation system of *Acinetobacter baylyi* strain ADP1 facilitates the creation of specific mutations in ADP1 chromosomal genes. Using PCR techniques, the required gene can be amplified from the chromosome and labelled with the *sacB*-Km cassette genetic marker (Jones and Williams, 2003a). This construct when taken up by the ADP1 natural transformation machinery, replaces the original chromosomal gene via homologous recombination creating a gene knockout marker strain with a positive transformant selection process.

The *sacB* gene encodes the *B. subtilis* extracellular levansucrase (sucrose: 2,6-β-D-fructan 6-β-D-fructosyltransferase) protein which catalyses the hydrolysis of sucrose and the synthesis of levans, high molecular weight polymers of fructose from sucrose (Dodender, 1966). When expressed by Gram-negative cells these polymers are highly toxic causing cell lysis or inhibition of growth (Gay *et al.*, 1985). This sucrose intolerance combined with the Km resistance cassette provides a triple phenotypic selection assay when inserted into a specific gene, i.e. sucrose sensitivity, kanamycin resistance and loss of activity from the chosen gene. The use of the *sacB*-Km cassette allows the creation of unmarked gene deletion mutants. Thus, specific mutations can be inserted into the target gene and transformed into the chromosome replacing the cassette via homologous recombination. This marker eviction strategy allows colony selection by phenotype reversal, i.e. sucrose resistance, kanamycin sensitivity but still negative for the target gene activity. In ADP1, the sucrose sensitivity phenotype is temperature-dependent and observed only in cultures incubated at room temperature therefore both the cassette insertion selection procedure and the marker replacement transformations were performed at this permissive

temperature. This marker-replacement technique was used to create five separate *salA*$^-$ strains of ADP1 with 4 bp deletions located within different regions of the *salA* gene.

2.3.3 Transformation frequency assays

The donor DNA was provided as homologous ADP1 DNA and heterologous *Acinetobacter* sp. DNA. The transformation assay, was similar to that utilized by Young and Ornston (2001). Overnight liquid cultures (30°C, 160 rpm) were diluted 1/10 into fresh succinate media and incubated for a further 2 h to induce competence for natural transformation. From this competent culture 0.4 ml aliquots were transferred to microcentrifuge tubes and incubated under the same conditions with either 2.5 μg ml^{-1} of linearized plasmid donor DNA or *Sal*I digested genomic DNA for 1 h. The transformation cultures were then placed directly on ice, serially diluted in phosphate-buffered saline and plated onto selective minimal media supplemented with salicylate (2.5 mM).

Each transformation culture was plated in duplicate and performed in parallel for a total of three replicates. For homologous donor DNA 0.1 ml of the serial dilution was plated, for heterologous donor DNA 0.2 ml was plated. Controls consisted of cultures incubated in the absence of donor DNA. Viable cells counts plated on succinate (10 mM) minimal media were also performed in parallel and replicated three times. The transformation plates were incubated at 30°C for 48 h. The transformation frequency was calculated as number of transformants / number of viable cells.

Section 4

2.4 Protein purification and enzyme assays

2.4.1 Affinity chromatography

Affinity chromatography separates proteins on the basis of a reversible interaction between a protein and a specific ligand coupled to a biologically inert matrix. Biological interactions between ligand and target molecule can occur as a result of electrostatic or hydrophobic interactions, van der Waal's forces and/or hydrogen bonding. To elute the target protein from the affinity media the interaction can be reversed, either specifically using a competitive ligand or non-specifically, by changing the pH, ionic strength or polarity. The $(His)_6$ tag is one of the most common fusion tags to facilitate purification and detection of recombinant proteins. Chelating sepharose when charged with Ni^{2+} ions selectively binds proteins if complex-forming amino acid residues, in particular histidine, are exposed on the protein surface. Bound $(His)_6$ fusion proteins can be easily eluted by displacement with buffers containing imidazole.

2.4.2 Cell harvesting and cell free extract preparation

Cultures grown for the production of cell extracts were centrifuged using a Beckman model J2-21 centrifuge and JA-14 rotor at 10 000 x g for 10 min to pellet the cells. The cells were washed once in potassium phosphate buffer (0.1 M) and stored at -20°C. Cell-free extracts were prepared from the frozen

pellets. After resuspension in an appropriate buffer, the cells were disrupted by either passage through a French pressure cell at 10 lb in^2, or by ultrasonication at an amplitude of 7 microns for three periods of 30 s interspersed with 30 s rest intervals. The disrupted cells were then centrifuged using a Beckman L8-80M centrifuge in a TY65 rotor at 150 000 x g for 45 min. The protein concentration of purified or cell-free extracts was estimated using a biuret kit, Total Protein Reagent, from Sigma-Aldrich according to the manufacturer's instructions.

2.4.3 His-trap protein purification

For the over-expression of a recombinant $(His)_6$ fusion protein and subsequent affinty purification a target gene was cloned into the pET28a his-tag vector. His-Trap chelating affinity columns supplied by Amersham Pharmacia Biotech were used as the affinity matrix according to the manufacturer's instructions. Protein over-expression batch cultures (2L LB km 2.5 μg ml^{-1}) of *E. coli* Rosetta transformed with the pET28a clone were grown at 37°C until an OD_{600} of ~0.3 was reached. At this point the growth temperature was reduced to 20°C to circumvent the formation of protein inclusion bodies. When the cultures reached an OD_{600} of ~ 0.8 protein expression was induced by the addition of 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG). Incubation was continued overnight at 20°C for approx. 16 h. In preparation for column purification the cell pellets were resuspended in imidazole (10 mM) phosphate buffer pH 7.5, including 1 mM β-mercaptoethanol (βME), harvested, sonicated and centrifuged as described in section 2.4.2. The supernatant was filtered through a 0.4 μM filter to remove any residual particulate matter prior to loading the sample onto the column. The bound protein was eluted with an imidazole concentration gradient. The eluted fractions were visualized for an over-

expressed protein band corresponding to the molecular weight of the protein (48 kDa) by SDS-PAGE electrophoresis gel anaylsis.

2.4.4 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein overexpression was analysed by SDS-PAGE 12% (w/v) using a Mini-Protean® 3 Electrophoresis cell (Bio-Rad). SDS-PAGE is carried out with a discontinuous buffer system in which the buffer in the reservoirs is of a different pH and ionic strength (Tris-glycine pH 8.3) from the buffer used to cast the gel. SDS in combination with β-mercaptoethanol (βME) and heat is used to denature the proteins before loading onto the gel. The denatured polypepetides bind SDS and become negatively charged enabling the SDS-polypepetide complexes to migrate through the polyacrylamide gel in accordance to the molecular weight of the protein (Sambrook, *et al.* 1989).

2.4.5 Enzyme assays

Enzyme assays and wavelength scans to examine salicylate hydroxylase activity were performed in a Uvikon 943 double beam UV/vis spectrophotometer Bio-Tek Kontron. The decrease in absorbance of NADH in the presence of substrate was monitored at 340 nm in 1 ml acryl cuvettes using Tris-HCl buffer (25 mM) pH 7.5, or sodium phosphate buffer (0.1M) pH 7.5, NADH (100 μM) and salicylate (100 μM). To correct for background NADH oxidase activity where necessary, in cell free extracts and purified protein wavelength scans or initial velocity measurements were performed with assay constituents, buffer, enzyme and NADH in both the blank and the test cuvette. Reactions were initiated by addition of the substrate to the test cuvette only. The rates and specific activities were calculated from the initial ΔA_{340}/min

taking the molar extinction coefficient of NADH at 340 nm to be 6 220 $M^{-1}cm^{-1}$.

2.4.6 Enzyme activity - cell free extracts and purified protein

Initial velocities were measured as a decrease in the absorbance of NADH at 340 nm. The straight-line slope of the reaction was measured using the first 15 seconds (approx.).

2.4.6.1 [E] vs rate

The effects of enzyme concentration (μg ml^{-1}) on the initial reaction rates were assayed against constant amounts of the other assay constituents.

2.4.6.2 Specific activity

Specific activity was measured from initial rates and calculated using the following equation:

Volume activity = $V \times \Delta A / \varepsilon \times d \times v$

Specific activity = Volume activity / protein concentration (mg^{-1} ml^{-1})

Where V = assay volume, ΔA = change in Absorbance/min, ε = Molar extinction co-efficient, d = light path distance (1 cm), v = volume of enzyme in assay. The molar extinction co-efficient of NADH was taken to be 6220 mol^{-1}.

2.4.6.3 pH profile

The pH profile was measured similarily using a range of buffered solutions, pH 4.0-6.0 - citrate buffer (50 mM), pH 6.0 - 9.0 - sodium phosphate buffer (0.1 M), pH 6.0 - 9.0 - tris-HCl (50 mM), pH 8.0-9.0 - Bicene (50 mM) and pH 10-11 - CAPS (50 mM).

2.4.6.4 Stoichiometry

The stoichiometry between NADH and salicylate of the salicylate hydroxylase reaction was studied at saturating levels of NADH and varied concentrations of salicylate. Changes in the oxidation rates of NADH at λ340 nm were monitored over time. The stoichiometric usage of NADH (μmol) per μmol of salicylate was calculated from the point of substrate exhaustion, where the NADH oxidation traces reached a plateau, deducted from the original concentration of NADH and divided by the absorbance of 1 μM NADH (0.00622 cm^{-1}). Apparent V_{max} and K_m were calculated using Enzpack, a kinetic data-modelling programme (Biosoft).

2.4.6.5 Enzyme kinetics

The kinetic profile of the enzyme was investigated in the presence of all substrates, varying their concentration sequentially as appropriate to two substrate or multisubstrate enzyme reactions. In this case, the initial velocities were determined under conditions of varying salicylate at several fixed concentrations of NADH. The same data set was used to reverse the conditions

i.e. varied NADH at several fixed concentrations of salicylate. The data were analysed using Hanes-Woolf reciprocal primary plots.

Chapter 3

3.1 Introduction

Acinetobacter baylyi strain ADP1 is highly competent for natural transformation and can take up DNA from any source. The ease of ADP1 competence induction facilitates investigations of horizontal gene transfer events in both laboratory and microcosm studies. A study conducted by the university of Aberysthwyth inserted the ADP1 salicylate hydroxylase gene (*salA*) into tobacco plants and quantified the ability of the transgenic *salA* to transform an ADP1 *salA*$^-$ knockout strain in a worm/soil microcosm (Chapter 1 – 1.4). In the natural environment however, it is likely that bacterial homologues of the salicylate hydroxylase gene are present in soil that may also be taken up as donor DNA, repairing the recipient *salA*$^-$ mutation. Therefore, to detect and quantify transformation frequencies that could occur between ADP1 and closely related bacterial *salA* homologues, a non-antibiotic laboratory based assay was developed. This chapter describes the creation of five *A. baylyi* ADP1 Δ*salA* recipient strains (ADPW257 – ADPW261) that were impaired in their ability to utilize salicylate as a sole carbon and energy source by the incorporation of small 4 bp deletions.

Chapter 3 preface

For the ease of detection of the Δ*salA* transformants, a marker gene (*salA*::*sacB*-Km) was constructed, which after transformation into ADP1 would produce a positive selection marker system (Chapter 2 - 2.3.3). This construct placed the *sacB*-Km marker in the centre of *salA*, however in this position, the

counter selectable marker could not be successfully established within the chromosome. As a result an alternative selectable marker strain, ADPW249, was used as the recipient marker strain to create the five Δ*salA* mutants. In the case of ADPW249, the *sacB*-Km cassette was stably inserted at the 3' termini of *salA*.

Table 3.2.1 Plasmids used and constructed in the formation of ADP1 *salA* mutations

Plasmids	Description	Reference
pGem-T-Easy	Ap^R TA cloning vector	Promega
pADPW34	pUC18 with 4.1 kbp *Sac*I / *Xba*I insert, containing *salA*,-*R*,-*E* and part of *sal*D, isolated from *Acinetobacter baylyi* strain ADP1.	Jones *et al.*, 2000
pRMJ1	$Ap^R Km^R$ *sacB*-Km cassette	Jones and Williams, 2003
pADPW250	pGem-T-easy with OEPM 1.2 kb ADP1 *salA* insert with central engineered *Bam*HI site	This study
pADPW251	pADPW250 with 3.6 kbp *sacB*-Km cassette	This study
pADPW257	pGem-T-easy with 1.2 kb *salA* 4 bp deletion insert. The deletion is located 121-124 bp down-stream of the translational start site	This study
pADPW258	pGem-T-easy with 1.2 kb *salA* 4 bp deletion insert. The deletion is located 332-335 bp down-stream of the translational start site	This study
pADPW259	pGem-T-easy with 1.2 kb *salA* 4 bp deletion insert. The deletion is located 618-621 bp downstream of the translational start site	This study
pADPW260	pGem-T-easy with 1.2 kb *salA* 4 bp deletion insert. The deletion is located 874-877 bp downstream of the translational start site	This study
pADPW261	pGem-T-easy with 1.2 kb *salA* 4 bp deletion insert. The deletion is located 1132-1135 bp downstream of translational start site.	This study

3.2 Materials and methods

3.2.1 Bacterial strains, plasmids and overlap-extension PCR mutagenesis (OEPM) primers

The plasmids and bacterial strains used are described in Tables 3.2.1 and 3.2.2, respectively. Bacterial growth conditions unless stated otherwise are as described in Chapter 2 (section 1 - 2.1.2 - 2.1.3). The primers used to construct mutant *salA* genes are listed in Tables 3.2.3 and 3.2.4.

3.2.2 Construction of pADPW250

In order to create a *salA*$^-$ mutant of ADP1 containing the selective *sacB*-Km marker cassette, oligonucleotide primers were designed to introduce a unique *Bam*HI restriction site in the central region of ADP1 *salA* using OEPM. The wild-type forward primer P5F and the reverse primer carrying the *Bam*HI mutation, P3R, were paired to form one PCR fragment. The forward internal mutagenesis primer, P4F, which carried the *Bam*HI site mutation, and the reverse external wild type primer, P6R, formed the other (Fig. 3.2.1; Table 3.2.4). pADPW34 was used as the template DNA (Table 3.2.1). The PCR conditions for the generation of these separate fragments were performed as programme 1 (Table 3.2.3). The two PCR products were excised together from an agarose gel, processed as described in Chapter 2 (section 2 - 2.2.9), and used as the DNA template for the overlap stage of the PCR mutagenesis. The external primers, P5F and P6R, were used for the overlap extension reaction (Fig 3.2.1).

Table 3.2.2. Bacterial Strains

Bacterial strains	Description	Reference
Escherichia coli DH5α	[ϕ80dlacZΔM15 recA1 endA1 gyrA96 thi-1 hsdR17($r_k^- m_k^+$) supE44 relA1 deoR Δ(lacZYA-argF)U169]	Novagen
***Acinetobacter baylyi* strains**		
ADP1	(BD413)Wild type	P. A. Williams collection
ADPW251	ADP1 *salA*⁻::*sacB*-Km. *sacB*-Km cassette inserted at unique engineered *Bam*HI site within *salA*	This study
ADPW249	ADP1 *salA*⁻::*sacB*-Km. *sacB*-Km cassette inserted at the natural *Hin*dIII site within *salA*	Gift from Dr. G. Elliot, University of Wales, Aberystwyth, UK
ADPW257	ADPW249 transformed with pADPW257	This study
ADPW258	ADPW249 transformed with pADPW258	This study
ADPW259	ADPW249 transformed with pADPW259	This study
ADPW260	ADPW249 transformed with pADPW260	This study
ADPW261	ADPW249 transformed with pADPW261	This study

The OEPM PCR conditions were performed according to PCR programme 2 (Table 3.2.3). The resulting 1.2 kbp PCR product was excised from an agarose gel and ligated into Ap^R pGem®-T-easy. The resulting plasmid was transformed into *E. coli* DH5α. Transformant colonies were selected for plasmid isolation and screened using standard procedures. A plasmid containing a 1.2 kbp insert that cut once with *Bam*HI was isolated, termed pADPW250 and stored at -20°C.

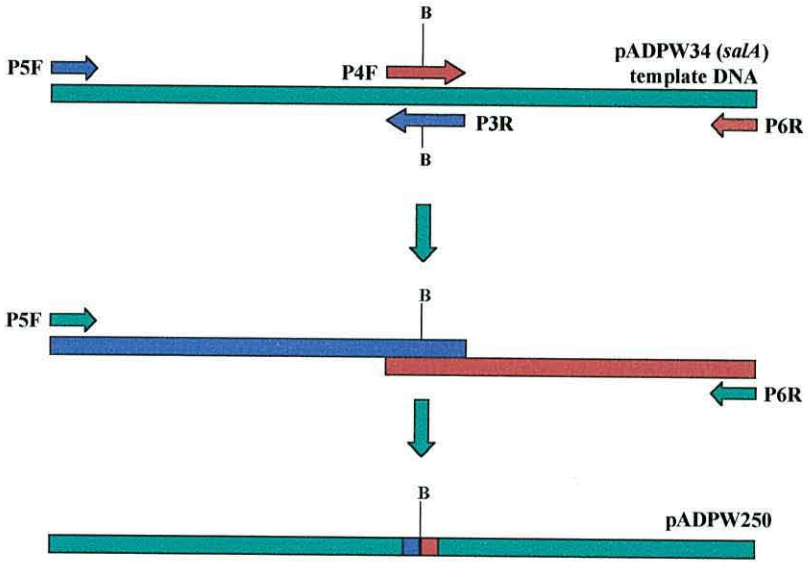


Fig. 3.2.1. OEPM procedure used to create pADPW250. pADPW34 was used as the *salA* DNA template. External primer P5F was paired with P3R, containing the reverse *Bam*HI site, to construct one PCR fragment. The external primer P6R was paired with P4F, containing the forward *Bam*HI site, constructing the other. The two fragments generated were annealed together using both external primers to form the complete mutated *salA* PCR product which was cloned into pGem-T-easy to form pADPW250.

Table 3.2.3. PCR reaction conditions to create a central *Bam*HI site within the ADP1 *salA* gene by OEPM using Hotstar Master mix.

	PCR programmes	
	1	**2**
Initial denaturation	95°C - 15 min	95°C - 15 min
Loop 1	**20 cycles**	**15 cycles**
Denaturation	94°C - 1 min	94°C - 1 min
Touchdown annealing temperature	55-47°C - 1 min	62-52°C - 1 min
Extension	72°C - 2 min	72°C - 2 min
Loop 2	**15 cycles**	**15 cycles**
Denaturation	94°C - 30 s	94°C - 1 min
Annealing temperature	48°C - 1 min	48°C - 1 min
Extension	72°C - 2 min	72°C - 2 min
Final extension	72°C - 10 min	72°C - 10 min

3.2.3 Construction of pADPW251

The 3.6 kbp promoterless *sacB*-Km cassette carried in pRMJ1 (Table 3.2.1) was excised as a *Bam*HI fragment and ligated into the engineered central *Bam*HI site of pADPW250, creating pADPW251. Transformant *E. coli*, containing putative pADPW251 were selected on LBA Ap (100 µg/ml) Km (50µg/ml) and screened via restriction digest and AGE for the presence of the sacB-Km cassette. The orientation of the cassette relative to the direction of transcription of the *salA* gene was also determined by restriction digest and AGE.

3.2.4 Construction of pADPW257-261

Five Δ*salA* plasmids were created using OEPM. Internal primers were designed, as listed in Table 3.2.4, to incorporate small 4 bp deletions within the ADP1 *salA* gene. Amplification of the separate PCR fragments was performed using the Expand High fidelity PCR system (Roche) using the relevant internal primers paired with the external primers P5F or P6R (Table 3.2.4; Fig. 3.2.2) in a 50 µl reaction volume. The reaction conditions were as PCR programme 1 (Table 3.2.5). These conditions were applied to generate all the separate PCR fragment except for P5F - PJR, and P6R - PIF which were amplified using the Expand High Fidelity PCR System in 50 µl reaction volume using PCR programme 2 (Table 3.2.5).

The overlap extension to join the relevant fragments was performed using *Pfu* DNA polymerase and the external primers P5F and P6R, (Table 3.2.4) in a 50 µl reaction volume. The reaction conditions are as described in Table 3.2.5: pADPW257 was amplified using programme 3, pADPW258 - 260 were amplified using programme 4 and pADPW261 was amplified using programme 5.

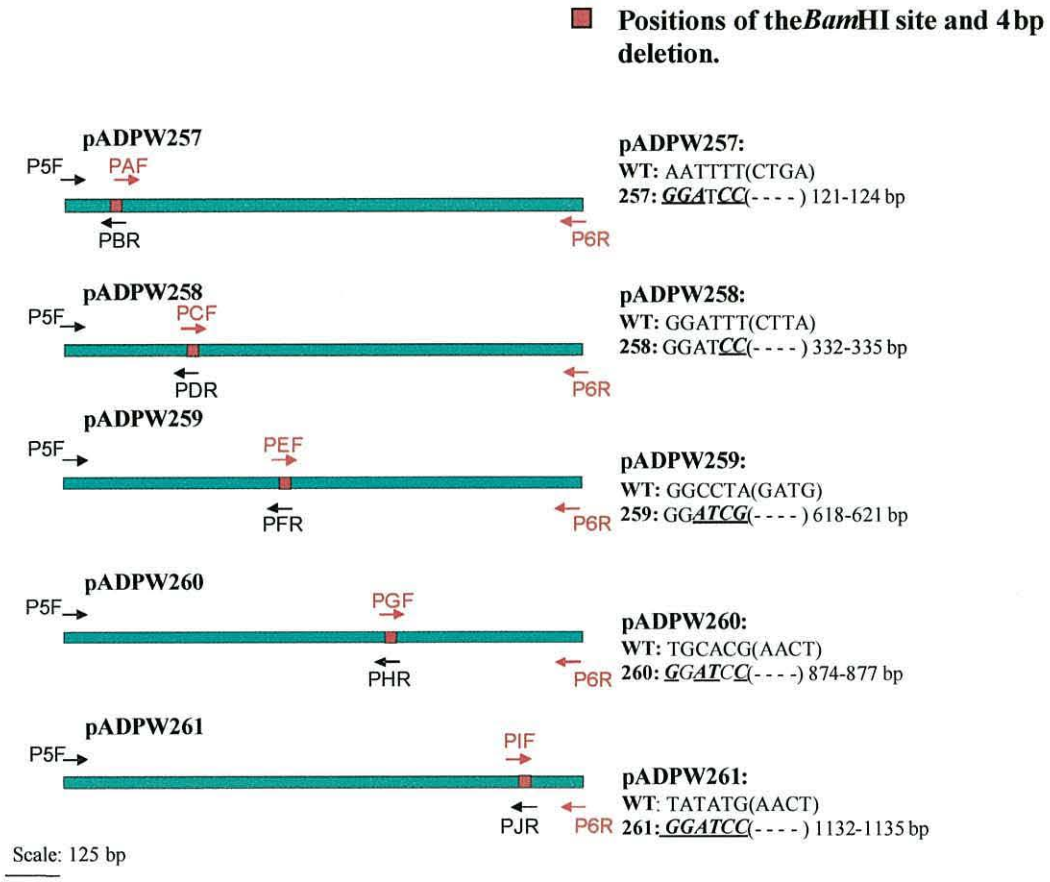


Fig.3.2.2. Positions of the engineered *Bam*HI restriction site and the 4 bp deletion incorporated within the ADP1 *salA* gene cloned into pADPW257-261. WT: ADP1. Bases in italics, bold and underlined are those altered from the wild type sequence to create the unique restriction site, bases deleted from the wild type sequence are represented by the bracketed dashed lines. Numbered base-pairs adjacent to dashed lines indicate the position of the deletion relative to the position of the GTG translational start site of *sal*A. Arrows of matching colour represent PCR primer pairs.

Table 3.2.4 *Δ salA* OEPM primer sequences

	External wild type primers
P5F	5'- AAAAATTAGTATAGCCATTATTG-GTGGTGGTAT-3'
P6R	5'- GCGCGGCTAACATGTCTTGTGCTA -3'
	Internal overlap primers designed to incorporate a unique *Bam*HI site within ADP1 *salA*
P4F	5'- AGCCGG***GGATCC***TGTAGAAATTGCCGACGT
P3R	5'- CAATTTCTACA***GGATCC***CCGGCTGCGACAA-3'
	Internal overlap primers designed to incorporate a 4 bp deletion Forward
PAF	5'- GCACCTC***GGATCC***----AATTGGTGCCGGAATTTCATTTG -3'
PCF	5'- CGTGC***GGATCC***----GATGCTATTTTAGGGAATATTCCA -3'
PEF	5'- CAGCC***GGATCC***----TAGAAATTGCCGACGTACCAC -3'
PGF	5'- GTGGGCAC***GGATCC***----GGCCGAATTGCCGACTTATC -3'
PIF	5'- GGGCAAA***GGATCC***----TAACTCAGCACTGTATC -3'
	Internal overlap primers designed to incorporate a 4 bp deletion Reverse
PBR	5'- GCACCAATT***GGATCC***----GAGGTGCTGCTTCATAC -3'
PDR	5'- TAGCATC***GGATCC***----GCACGATGCACCGAAG -3'
PFR	5'- CGGCAATTTCTA***GGATCC***----GGCTGCGACAATGGC -3'
PHR	5'- ATTCGGCC***GGATCC***----GTGCCCACAAGGTCGGAC -3'
PJR	5'- GAGTTA***GGATCC***----TTTGCCCAGACTCACGC -3'

Bases underlined and italicized have been altered to form the *Bam*HI restriction site. Dashed lines represent the position of deleted bases.

	PCR programmes					
	1	**2**	**3**	**4**	**5**	**6**
Initial denaturation	94°C - 2 min	94°C - 2 min	95°C - 2 min	95°C - 2 min	95°C - 2 min	95°C - 15 min
Loop 1	**Loop 1 25 cycles**	**Loop 1 25 cycles**	**Loop 1 20 cycles**	**Loop 1 15 cycles**	**Loop 1 15 cycles**	**Loop 1 15 cycles**
Denaturation	94°C - 15 s	94°C - 15 s	95°C - 1 min	95°C - 1 min	95°C - 1 min	94°C - 1 min
Touchdown annealing temperature	71 - 62°C -1 min	71 - 61°C - 1 min	68 - 58°C - 30 s	69 - 59°C - 30 s	69 - 59°C - 1 min	64 - 54°C - 1 min
Extension	72°C - 2 min	72°C - 2 min	73°C - 3 min	73°C - 3 min	73 °C - 3 min	72°C - 2 min
Loop 2	**Loop 2 15 cycles**	**Loop 2 15 cycles**	**Loop 2 10 cycles**	**Loop 2 15 cycles**	**Loop 2 15 cycles**	**Loop 2 20 cycles**
Denaturation	94°C - 15 s	94°C - 15 s	95°C - 1 min	95°C - 1 min	95°C - 1 min	94°C - 1 min
Annealing temperature	65°C - 30 s	61°C - 30 s	60°C - 30 s	61°C - 30 s	61°C - 1 min	54°C - 1 min
Extension	72°C -1 min 30 s	72°C - 2 min	72°C- 3 min	72°C- 3 min	72°C - 3 min	72°C - 2 min
Final extension	72°C -10 min	72°C -10 min	74°C - 5 min.	74°C - 5 min.	74°C - 5 min	72°C - 10 min

Table 3.2.5. PCR programmes used in the construction of OEPM ADP1 Δ*salA* and colony PCR (Chapter 3 Section 3.2.7). Programmes 1 and 2 were performed using Expand High Fidelity PCR System (Roche) to construct the initial single overlapping PCR fragments. Programmes 3, 4 and 5 were performed using *Pfu* DNA polymerase (Promega) to construct the OEPM PCR products cloned into pGem-T-easy to form pADPW257-261. Programme 6 was used to amplify the ADP1 *salA* gene from single colonies.

3.2.5 Creation of strain ADPW251

pADPW251 DNA was digested with *Sph*I and *Sac*I to cleave the insert from the vector and used to transform ADP1 according to Chapter 2 (section 3 - 2.3.1). The overnight transformation culture was plated onto LBA-Km (10 μg/ml) and grown at room temperature for transformant selection. Single colonies from this media were toothpicked in series onto LBA-sucrose 5% (w/v), salicylate minimal agar, (2.5 mM), LBA-Km (10 μg/ml) and succinate minimal agar (10 mM).

3.2.6 Creation of strains ADPW257-261

The plasmid constructs pADPW257-261 (Table 3.2.1) were used to transform *Acinetobacter baylyi* strain ADPW249 (Table 3.2.2). All procedures involving cultures of ADPW249 were carried out at room temperature, the permissive temperature for the expression of the *sacB* marker gene in *Acinetobacter*. Overnight cultures of ADPW249 were diluted 1/20 and incubated with *Eco*RI-digested pADPW257-261 DNA (Table 3.2.1), and incubated for approx. 18 h. The transformation culture was plated on 5% (w/v) sucrose selective medium and incubated at room temperature. Single colonies were selected from this growth medium and toothpicked in series onto LBA-Km (10 μg/ml), salicylate minimal agar (2.5 mM), LBA-sucrose (5% w/v) and succinate minimal agar (10 mM).

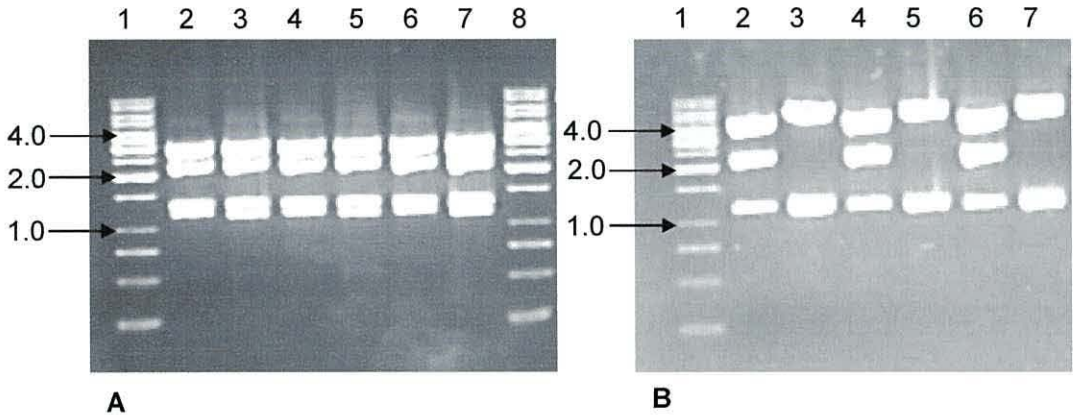


Fig.3.3.1. A. Agarose gel of the pADPW251 *Eco*RI digests; screening for the *sacB*-Km cassette insertion of pADPW250. Lanes: 1 and 8 - 1 Kb DNA marker. Lanes: 2 - 7 pADPW251 clones.

Fig.3.3.1. B. Digestion of pADPW251 clones with *Hin*dIII. Lane1: 1 Kb DNA ladder. Lanes 2-7 *Hin*dIII digested pADPW251 clones. In the *salA*::*sacB*-Km orientation plasmids (lanes 3, 5 and 7) the 1.3 and 1.1 kbp bands are seen as a large double band on the gel.

3.2.7 Colony PCR

Colony PCR to isolate the *salA* gene from the wild type strain ADP1 and the mutant strains ADPW257-261 was performed using the wild type external primers P5F and P6R (Table 3.2.5) and the Hotstar Master Mix kit. The PCR conditions are described in Table 3.2.5 programme 6.

3.3 Results

3.3.1 Creation of ADP1 *salA* mutants

3.3.1.1 Construction of pADPW251

pADPW251 was constructed by the ligation of the *sacB*-Km cassette into the unique engineered *Bam*HI site of the *salA* gene carried in pADPW250. As the cloning procedure was bi-directional, the *sacB*-Km cassette could be inserted into the gene in either orientation. For gene expression, the promoterless *sacB* gene has to be transcribed in the same direction as the *salA* gene. Restriction digest with *Eco*RI identified the presence of the cassette, giving fragment sizes of 3.0, 2.2, 1.5, 1.2 kbp, (Fig. 3.3.1A and Fig. 3.3.2), but a second digest using *Hin*dIII is required to identify the orientation of the cassette, (Fig. 3.3.1B. and Fig. 3.3.2). The *sacB*-Km orientation of the cassette can be easily distinguished from the Km-*sacB* insertion. If the cassette is inserted as *sacB*-Km then the *Hin*dIII digest produces DNA fragment sizes of 5.5, 1.3 and 1.1 kbp (Fig. 3.3.1B lanes 3, 5 and 7) and the reverse orientation produces DNA fragment sizes of 4.6, 2.4 and 1.1, (Fig. 3.3.1B lanes 2, 4 and 6).

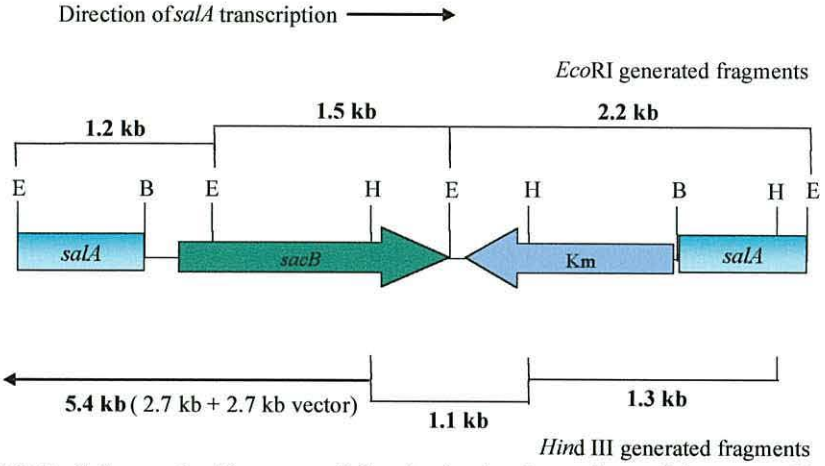


Fig.3.3.2. Schematic diagram of the desired orientation of the *sacB*-Km cassette within the ADP1 *salA* gene. Restriction of pADPW251 with *Eco*RI produces fragment sizes of approx. 3.0 kbp (vector), 2.2, 1.5 and 1.2 kbp. Digestion of the *sacB*-Km cassette in this orientation with *Hin*dIII, produces fragment sizes of 5.4, 1.3 and 1.1 kbp.

3.3.1.2 Attempted creation of ADPW251

pADPW251 DNA was digested with *Sph*I and *Sac*I to cleave the insert from the vector and used to transform ADP1 according to Chapter 2 (section 3 - 3.2.1). The overnight transformation culture was plated onto LBA-Km (10 μg/ml) and grown at room temperature for transformant selection. Single colonies from this media were toothpicked in series onto LBA-sucrose 5% (w/v), salicylate minimal agar (2.5 mM), LBA-Km (10 μg/ml) and succinate minimal agar (10 mM) to detect the correct transformant phenotype (sucrose^S, *salA*$^-$, Km^R and succinate$^+$). A colony was isolated with this phenotype, but colony PCR and Southern hybridizations failed to detect the insertion of the *sacB*-Km cassette despite resistance to kanamycin being observed in this strain, indicating that the cassette had inserted incorrectly into the chromosome. However this strain was

not healthy and did not survive well on LBA-Km (10 μg/ml). Attempts to isolate another strain with the desired phenotype were unsuccessful. As a last resort, *Acinetobacter baylyi* strain ADPW249 was acquired as a gift from Dr. G. Elliot (see reference in Table 3.2.1), in which the *sacB*-Km cassette was carried chromosomally having been inserted into the natural *Hin*dIII site at the terminal end of *salA* (Table 3.2.1). This strain was then used for all further ADP1 Δ*salA* transformations requiring the *sacB*-Km marker eviction procedure.

3.3.1.3 Construction of pADPW257-pADPW261

Five Δ*salA* plasmids were created using OEPM oligonucleotide primers designed to incorporate small 4 bp deletions within five different regions of the gene. The OEPM primer design also included the creation of an adjacent unique restriction site, *Bam*HI, for identification purposes (Table 3.2.4).

Using P5F and P6R as the external wild type primers, sets of OEPM fragments were generated by pairing with a series of internal primers (Fig.3.2.2). Primer P5F was paired with the internal reverse primers, PBR, PDR, PFR, PHR and PJR respectively; the external reverse primer, P6R was paired with the internal forward primers, PAF, PCF, PEF, PGF and PIF, respectively. For example, P5F and PBR, P6R and PAF, are paired to produce two overlapping DNA fragments which are subsequently amplified together forming the pADPW257 Δ*salA*. These two PCR product fragments are subsequently mixed and used as the template for the overlap PCR reaction using the wild type external primers, P5F and P6R.

PCR products displaying the correct size (1.2 kbp) were cloned into pGem-T-easy, transformed into DH5α, and selected for via ampicillin antibiotic resistance and X-gal chromogenic detection. The plasmids isolated were screened for the presence of a *Bam*HI restriction site that created fragments of the correct sizes relative to the position of the deletions. Each deletion construct was sequenced to verify that the engineered genes contained only the desired mutations before being used to transform the recipient strain ADPW249 (Table 3.2.2).

3.3.1.4 Creation of strains ADPW257 - ADPW261

The plasmid constructs pADPW257-261 containing the Δ*salA* genes were digested with *Eco*RI and used to transform *Acinetobacter baylyi* strain ADPW249 creating strains ADPW257-261. The *sacB*-Km marker selection eviction strategy was utilized for transformant detection. Colonies displaying the phenotype suggesting the *sacB*-km cassette had been removed, i.e. sucrose resistant, kanamycin sensitive and salicylate negative were examined by colony PCR. The resultant PCR products were digested with *Bam*HI and visualized by AGE (Fig.3.3.3). Positive colonies were stored as agar slopes and as -70°C glycerol cultures.

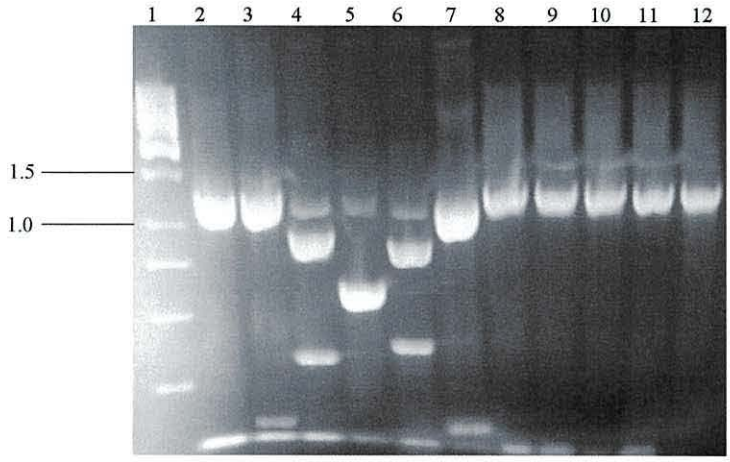


Fig.3.3.3. ADPW257-261 colony PCR. The *salA* gene was amplified using the external primers P5F and P6R. **Lane 1**. Fermentas 1 kb Generuler DNA ladder. **Lane 2**: ADP1 colony PCR digested with *Bam*HI. **Lanes 3-7** respectively show the results of the colony PCR after digestion with *Bam*HI. **Lane 3**. ADPW257. **Lane 4**. ADPW258. **Lane 5.** ADPW259. **Lane 6.** ADPW260. **Lane 7**. ADPW261. **Lane 8-12** respectively show the undigested colony PCRs of strains ADPW257-ADPW261.

As can be seen in Fig.3.3.3 the *salA* genes containing the deletions and engineered *Bam*HI sites cut the gene into different sizes depending on the location of the deletion and *Bam*HI site. The wild type ADP1 gene contains no *Bam*HI restriction site and serves as a negative control. The genes isolated from strains ADPW257 - ADPW261 clearly contained the *Bam*HI site. The phenotypic growth patterns and the colony PCR provided evidence that the mutated gene had successfully replaced the gene that contained the *sacB*-Km cassette.

3.4 Discussion

The creation of *Acinetobacter baylyi* strains ADPW257-261 facilitates the use of a simple natural transformation assay using homologous and heterologous donor DNA to examine, and quantify if possible, the transformation frequencies that may occur between the salicylate hydroxylase genes of closely related species. The insertion of specific mutations within a specific gene allows a defined approach to gene transfer/repair between bacterial species with an identified ability to utilize a certain growth substrate as a single carbon source.

The initial design to place the *sacB*-Km cassette within the central region of the *salA* gene led to problems with the expression of levansucrase. Attempts to verify the insertion of the cassette in strain ADPW251 by Southern hybridisation and PCR were unsuccessful despite an observed kanamycin resistant phenotype in the transformants. This may be due to the incorrect insertion of the cassette possibly caused by a deletion of the *sacB* counterpart prior to homologous recombination and a random insertion of the kanamycin resistance gene in the ADP1 chromosome. The position of the central *Bam*HI site may not have been conducive to homologous recombination. As will be seen in chapter five, transformations involving this area of the gene did not give expected transformant frequencies in comparison to the other deletions.

Jones and Williams (2003a) have successfully used the *sacB*-Km counterselectable marker in ADP1 to create deletions in the AreR regulated promoter region of the ADP1 *are* genes. In this case, the insertion of the *sacB*-Km cassette lies upstream of the coding region. When the cassette was inserted at the natural *Hin*dIII restriction site of *salA* at the C-terminal end of the gene, as in strain ADPW249, no problems with *sacB* expression were encountered,

(G. Elliot, personal communication) and the counterselectable marker operated efficiently. Therefore it is possible that the insertion of the cassette in the centre of the gene was unsuccessful due to topological effects of the DNA or possibly the occurrence of a recombinational 'cold spot'. As a result, strain ADPW249 was used for the insertion of the engineered *salA* deletion genes into the chromosome of ADP1.

Chapter 4

4.1 Introduction

The genus *Acinetobacter* (Chapter 1 - 1.2) is widespread in the natural environment and readily isolated from soil, water and activated sludge (Juni 1978). This genus has diverse nutritional capabilities and certain species are able to catabolize a wide range of aromatic hydrocarbon compounds to TCA cycle intermediates via the β-ketoadipate pathway. The occurrence and distribution of *A. baylyi* strain ADP1 gene homologues coding for salicylate 1-monooxygenase was examined within 34 *Acinetobacter* environmental isolates. These isolates were donated for the purpose of this study by Dr. D. Young and Prof. L. N. Ornston, Yale University, USA and Prof. J. Fry, University of Wales, Cardiff, UK. In order to provide donor DNA for the natural transformation and Δ*salA* gene repair studies of the five Δ*salA* recipient *A. baylyi* strains, ADPW257 – ADPW261, *Acinetobacter* spp. were identified that were capable of growth with salicylate as the sole carbon and energy source. This chapter describes the identification, isolation and characterization of *Acinetobacter* spp. salicylate hydroxylase genes present within the 34 environmental isolates.

Chapter 4 preface

Initially, the 34 *Acinetobacter* spp. were tested for single substrate growth with salicylate, *p*-hydroxybenzoate, benzoate, succinate, and ethyl-salicylate. Of these, 8 were selected for further analysis on the basis of showing varying degrees of growth with salicylate as the sole carbon source. Southern

hybridizations, shotgun cloning and degenerate PCR experiments were conducted to identify the presence of the putative salicylate hydroxylases. Southern hybridizations using the ADP1 *salA* as a probe identified *salA* homologues in two USA isolates – *Acinetobacter* sp. strains AD3-1 and AD321, and in two of the Cardiff *Acinetobacter* sp. isolates, strains FS30 and BS6. Shotgun cloning experiments isolated two putative *sal* operons from the genomes of *Acinetobacter* sp. strains AD3-1 and FS50. The AD3-1 *sal* operon was fully sequenced to a high quality. The FS50 shotgun clone was partially sequenced and was shown to be highly homologous to the operon isolated from AD3-1. Degenerate PCR was used to isolate the ADP1 *salA* and *salA* homologues from *Acinetobacter* sp. strains BS6, AD3-1 and FS50. Sequence alignments of these PCR products revealed an amino acid sequence identity of ~60%.

Table 4.0 Primers and PCR reaction conditions

Primers	**Name**	**Primer sequence 5' - 3'**
16S Forward	27F	AGAGTTTGATC(AC)TGGCTCAG
16S Reverse	1492R	TACGG(CT)TACC-TTGTTACGACTT
PCR reaction conditions		
Initial denaturation	95°C - 5 min	
Loop 1	**20 cycles**	
Denaturation	95°C - 30 s	
Annealing temperature	Touchdown 55-45°C - 30s	
Extension	72°C - 1 min 30s	
Loop 2	**15 cycles**	
Denaturation	95°C - 30 s	
Annealing temperature	45°C - 30 s	
Extension	72°C 1 min 30 s	
Final extension	72°C - 5 min	
***SalR* primers**	**Name**	**Primer sequence 5' - 3'**
Forward	salRF	GCGCGAGCAGCAATACAAC
Reverse	salRR	GCTTAGCCCGCGGTCTTACA
***salR* PCR reaction conditions 35 cycles**		
Initial denaturation	95°C - 15 min	
Denaturation	94°C - 30 s	
Annealing temperature	Touchdown 62-55°C - 1 min	
Extension	72°C - 2 min	
Final extension	72°C - 10 min	
***salA* degenerate primers**		
***SalR* primers**	**Name**	**Primer sequence 5' - 3'**
Forward	SalAFM1	ATTGGTGG(AGCT)GG(AGCT)ATTGCAGG
Reverse	SalARM1	ATGCTGCCA(AGCT)AACCAGTCAAAACG

4.2 Materials and methods

4.2.1 *Acinetobacter* spp.

Acinetobacter spp. used are listed in Table 4.1. The growth phenotypes of the environmental isolates were examined and compared to the wild type naturally transformable strain ADP1 using *p*-hydroxybenzoate 2.5 mM, benzoate 2.5 mM, salicylate 2.5 mM and 1.5 mM, succinate 10mM and ethyl salicylate as single carbon sources, the growth conditions were as described in Chapter 2 (section - 2.1.2).

4.2.2 Basic 16S rRNA gene isolation

A selection of *Acinetobacter* spp. including ADP1 underwent 16S rRNA amplification using *Taq* master mix, (Promega) and the universal primers 27F and 1492R (Table 4.0). The reaction conditions were as described in Table 4.0. The PCR products were digested without purification with *Msp*I, *Rsa*I, and *Cfo*I and visualized by AGE.

4.2.3 Southern hybridization analysis

Southern hybridizations were performed as described in Chapter 2 (section 2 - 2.2.13). The primers and PCR reaction conditions used to amplify the ADP1 *salR* gene using Hotstar master mix and pADPW34 as the DNA template are described in Table 4.0. Visualization of the gel after the capillary transfer revealed no residual DNA, indicating that complete transfer of the DNA from the gel to the membrane was achieved.

4.2.4 Isolation of salicylate 1-hydroxylase gene homologues

The isolation of putative *salA* gene homologues was attempted using shotgun cloning techniques, as described in Chapter 2 (section 2 - 2.2.12) and PCR amplification techniques using degenerate primers (Table 4.0). The forward and reverse degenerate primers SalAF1 and SalARM1, respectively. were designed using the DNAstar Megalign software (Martinez-NW pairwise nucleotide alignment) using *salA* sequences from ADP1 and *Acinetobacter* sp. strain AD3-1 (isolated as a shotgun clone). The PCR reaction was performed using Hotstar Master Mix, under the same conditions as described for colony PCR (Chapter 3. 3.2.7). The sequencing and analysis of isolated salicylate genes was performed as described in Chapter 2 (section 2 - 2.2.8). A putative salicylate operon isolated as a shotgun clone from the *Acinetobacter* sp. strain AD3-1 chromosome underwent high quality sequencing. Bacteria and plasmids used and constructed for this section are listed in Tables 4.1 and 4.2.

Table 4.1. ***Acinetobacter*** **spp.**

Bacteria	Growth phenotype					Reference
***Acinetobacter* spp.**	**p-HBA**	**Ben**	**Succ**	**Sal**	**E-Sal**	
ADP1 Wild type	++	++	++	++	++	P.A.Williams collection
LUH540[a] (*A. johnsonii*)	-	++	-	-	-	[a] Gifts from D. Young & L. N. Ornston Yale University, New Haven, CT, USA
81A1[a]	++	++	++	-	-	
85A1[a]	++	++	++	-	-	
AZR3577[a]	-	++	-	-	-	
59A1[a]	++	++	++	-	-	
AD3-1[a]	++	++	++	++	++	
O1BO[a]	++	++	++	-	-	
AD321[a]	++	++	++	++	++	
AC511B[a]	++	++	++	-	-	
AZR583[a]	++	++	++	+	+	
40B4F[a]	++	-	++	-	-	
(*A. haemolyticus*)						
AA1-1[a]	++	-	++	-	-	
AD513A[a]	-	++	++	-	-	
71A1[a]	++	++	++	-	-	
66A1[a]	++	++	++	+	+	
AZR3410[a]	++	++	++	-	-	
AC423D[a]	++	-	++	-	-	
AZR54[a]	++	++	++	-	-	
62A1[a]	++	++	++	-	-	

Table 4.1 continued

Acinetobacter spp.	p-HBA	Ben	Succ	Sal	E-Sal	
63A1[a]	++	++	++	+	+	
A3-6[a]	++	++	++	-	-	
P1-6[a]	++	++	++	++	++	
AΦ6[a]	++	++	++	++	++	
AC223[a]	-	++	++	-	-	
48A1[a]	++	++	++	+	+	
AZR2865[a]	-	++	++	-	-	
FS30[b]	++	++	++	++	++	[b] Gifts from J. Fry. University of Wales, Cardiff, UK
FS43[b]	++	++	++	++	++	
BS6[b]	++	++	++	++	++	
FS50[b]	++	++	++	++	++	
FS47[b]	-	-	++	-	-	
FS57[b]	-	-	++	-	-	
FS60[b]	-	-	++	-	-	
FS54[b]	++	++	++	-	-	

pHBA - *p*-hydroxybenzoate, Ben - benzoate, Succ - succinate, Sal - salicylate, E-sal - ethyl salicylate.
(+ +), (+), (-): indicates the level of growth shown by each strain. (+ +) = grew well, (+) = grew poorly, (-) = no growth

Table 4.2. Plasmids used and constructed

Plasmids	Description	Reference
pUC19	Ap^R cloning vector	Promega
pGem-T-easy	Ap^R cloning vector	Promega
pWW0-6000	contains *P. putida* C23O gene	P.A.Williams collection
pAD3-1A	5.78 kbp *Xba*I / *Sac*I AD3-1 chromosomal fragment generated by shotgun cloning ligated into the pUC19 MCS	This study
pFS50-1	approx. 7.0 kbp *Sac*I chromosomal fragment generated by shotgun cloning ligated into the pUC19 MCS	This study
pADP1A	Incomplete ADP1 *salA* gene degenerate-PCR fragment cloned into pGem-T-easy.	This study
pAD3-12A	Incomplete AD3-1 *salA* gene degenerate-PCR fragment cloned into pGem-T-easy.	This study
pFS50-pA	Incomplete FS50 *salA* gene degenerate-PCR fragment cloned into pGem-T-easy.	This study
pBS6-A	Incomplete FS50 *salA* gene degenerate-PCR fragment cloned into pGem-T-easy.	This study

4.3 Results

4.3.1 Growth phenotypes

The 34 *Acinetobacter* strains were tested for growth on salicylate, ethyl salicylate, *p*-hydroxybenzoate, benzoate and succinate as single carbon sources. The results are summarized in Table. 4.1. Strains AD3-1, AD321, P1-6, AΦ6, FS50, FS30, BS6, FS43 which grew well on salicylate 2.5 mM and strains AZR583, 66A1, AZR3410, AC423D, 63A1, 48A1, which showed moderate or poor growth on 1.5 mM salicylate were selected for further examination for the presence of salicylate genes.

The majority of strains utilized succinate except for *A. johnsonii* LUH540 and AZR3577. Growth with *p*-hydroxybenzoate was observed in 23 of the 34 strains indicating the presence of a protocatechuate branch of the β-ketoadipate pathway. Strains 48A1, AD321, O1BO, *A. haemolyticus* 40B4f (ATCC 17906) have been previously described to catabolize p-hydroxybenzoate via protocatechuate dioxygenase (*pcaHG*) and the β-ketoadipate pathway (Young and Ornston, 2001). Growth with benzoate was observed in 29 of the 34 strains indicating a broad distribution of benzoate dioxygenase genes. Growth with *p*-hydroxybenzoate, benzoate and succinate was observed in 20 out of 34 strains. Growth with salicylate was observed in 8 out of 34 strains. These strains also grew with the other substrates as single carbon sources indicating a broad distribution of catabolic genes among these *Acinetobacter* strains.

4.3.2 Basic 16S rRNA analysis

The 16S rRNA genes of 8 selected *Acinetobacter* strains were amplified from single colonies and subjected to Amplified 16S Ribosomal DNA Restriction enzyme Analysis (ARDRA). The PCR results and restriction digest patterns produced by digestion with *Msp*I, *Rsa*I and *Cfo*I are presented in Figure 4.1.

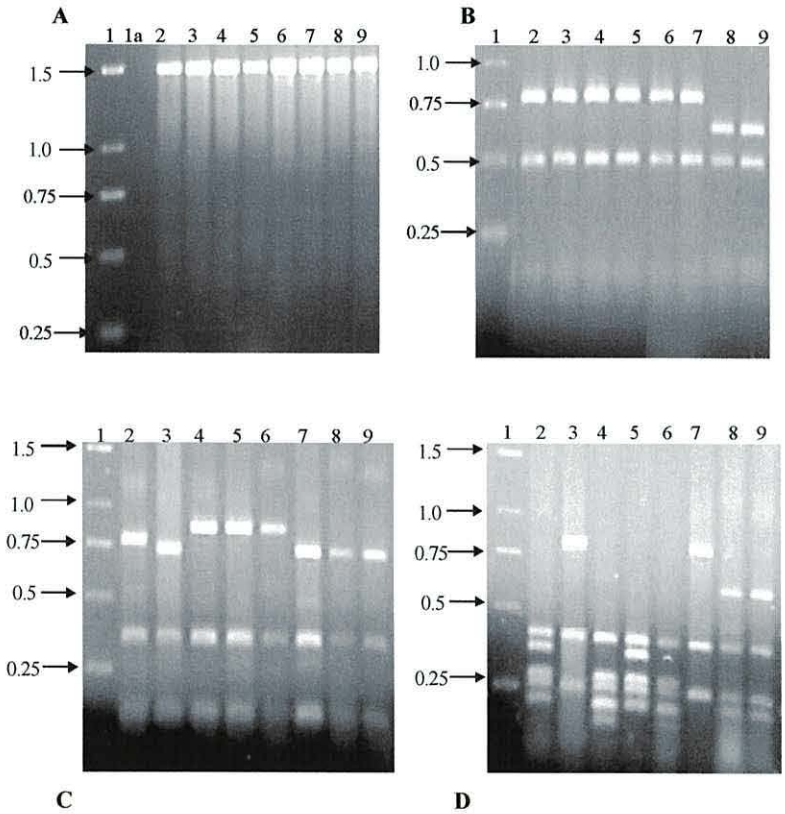


Fig. 4.1. 16S PCR and RFLP. **Lane 1.** 1 kb ladder. **Lane 1a (A).** No template control. **Lane2.** ADP1. **Lane 3.** AD3-1. **Lane 4.** BS6. **Lane 5.** P1-6. **Lane 6.** FS30. **Lane 7.** FS50, **Lane 8.** LUH540. **Lane 9.** FS57.
A. Undigested PCR products.
B. *Msp*I restriction digest.
C. *Rsa*I restriction digest
D. *Cfo*I restriction digest.

The correct ARDRA ribosomal gene typing for *Acinetobacter* involves the use of five enzymes, *Cfo*I, *Alu*I, *Mbo*I, *Rsa*I and *Msp*I. The combination of the ARDRA profiles provides a reference which is generally strain specific. Differentiation between strains requires the use of all five enzymes with further differentiation provided by ARDRA digest with *Bfa*I and *Bsm*AI (http://allserv.rug.ac.be/~mvaneech/ARDRA/Acinetobacter.html). In order to investigate the genus status of the donated strains, selected strains were digested with the single enzymes *Msp*I, *Rsa*I and *Cfo*I (Fig.4.1). These digests produced fragment sizes in agreement with profiles expected for an *Acinetobacter* species. Assigning a simple number code to the RFLPs allowed the production of a simple dendogram on the basis of the three digest RFLPs (Fig.4.1a).

Strain	*Msp*I	*Rsa*I	*Cfo*I
ADP1	101	01011	00111110
P1-6	101	10011	00111110
AD3-1	101	00111	10100100
FS50	101	00111	10100100
FS30	101	10011	00101111
BS6	101	10011	00101111
LUH540	011	00111	01100110
FS57	011	00111	01100110

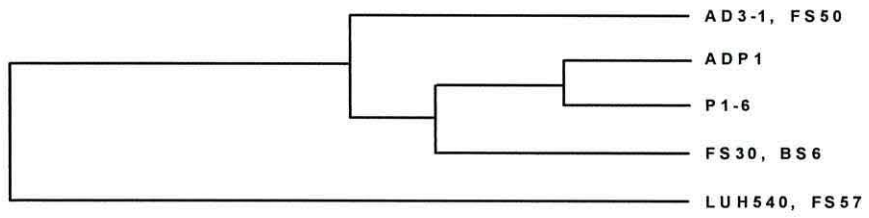


Fig.4.1a. Simple dendogram derived from the *Acinetobacter* spp. 16S rRNA gene RFLP data. 1 or 0 represents the presence or absence of a restriction fragment for each digestion and each strain.

To check that the basic tree (Fig.4.1a) was a good interpretation of the restriction data a phylogeny computer package PAUP (phylogenetic analysis using parsimony) was used. Using the tabulated data in Fig.4.1a, a heuristic branch and bound algorithm that searches for all possible trees produced two trees that contained groups of the same strains except that the placement of strain P1-6 varied. The consensus (Fig.4.1b) of these trees was calculated from 1000 replicates. The grouping pattern indicates to a basic extent the relationship between these eight *Acinetobacter* spp..

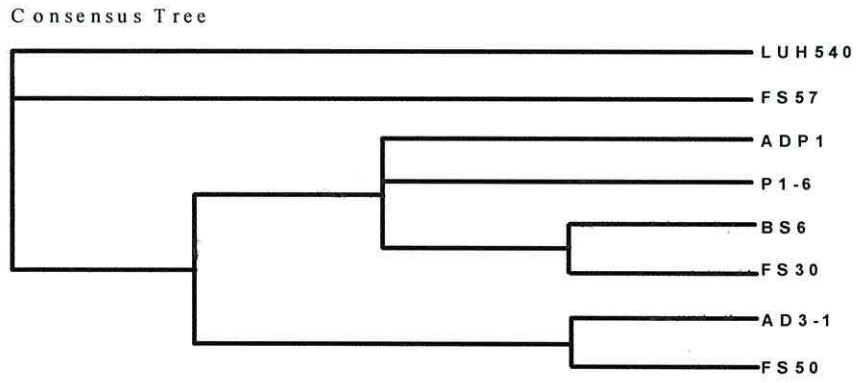


Fig.4.1b. PAUP computer package branch and bound consensus tree derived from RFLP data of Fig.4.1 and Fig.4.1a.

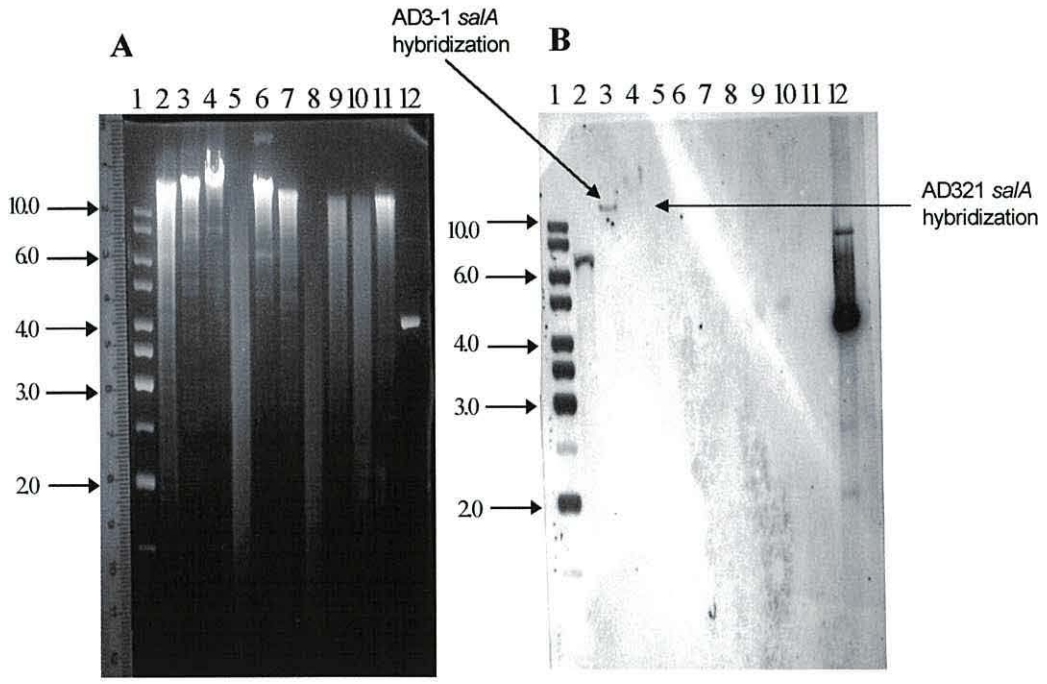


Fig.4.2 A. Southern hybridization gel of *Acinetobacter* spp. (USA strains) *SacI* digested genomic DNA. **Fig. 4.2 B**. Hybridization blot probed with the ADP1 *salA* gene. **Lane 1**. 1 kb DNA ladder. **Lane 2**. ADP1. **Lane 3**. AD3-1. **Lane 4**. AD321. **Lane 5**. 66A1 (15). **Lane 6**. AZR3410 (16). **Lane 7** AC423D (17). **Lane 8**. 63A1 (20). **Lane 9**. P1-6 (22). **Lane 10**. 48A1 (25). **Lane 12**. pADW34 - digested with *Xba*I / *Sac*I. Hybridized DNA bands are observed in lanes 2, 3, 4 and 12.

4.3.3 Genomic DNA Southern hybridization

Southern hybridizations were performed using the *salA* DNA fragment purified from pADP1A as the probe (Fig.4.2). Fig.4.2 B shows the hybridization results for eight of the USA *Acinetobacter* sp. strains. DNA bands illuminated by the probe are visible in lanes 2, 3 and 12. Lane 2 is the positive control wild type ADP1 genomic DNA digested with *Sac*I which gives a clear signal band of ~7.0 kbp. Lane 12 contains a second positive control, pADPW34 DNA digested with *Xba*I / *Sac*I, a digest which cleaves the insert from the vector. A large chemiluminescent signal band of 4.1 kbp in this lane corresponds to the size of the insert DNA. In lane 3 a definite signal band for strain AD3-1 is observed above the molecular weight standards. The large size of this fragment indicates that a DNA *Sac*I fragment larger than 10 kbp carries a hybridized *salA* homologue.

Lane 4 (strain AD321) contains a very faint band the same size as that observed for AD3-1. This indicates the presence of a *salA* gene homologue in this strain that shows very weak hybridization to the probe. Above this faint band there is another area lit more strongly by the probe suggesting the presence of undigested genomic DNA that has also hybridized to the probe. This may be non-specific hybridization or actual *salA* detection in an undigested or partially digested fraction. The weakness of the signal for strain AD321 compared to that detected for AD3-1 may be due to less homology between the *salA* genes. No hybridization signals were detected for any of the other USA strains examined. This may be due to the lack of salicylate hydroxylase genes in these strains or to low gene homology in the chromosomal DNA.

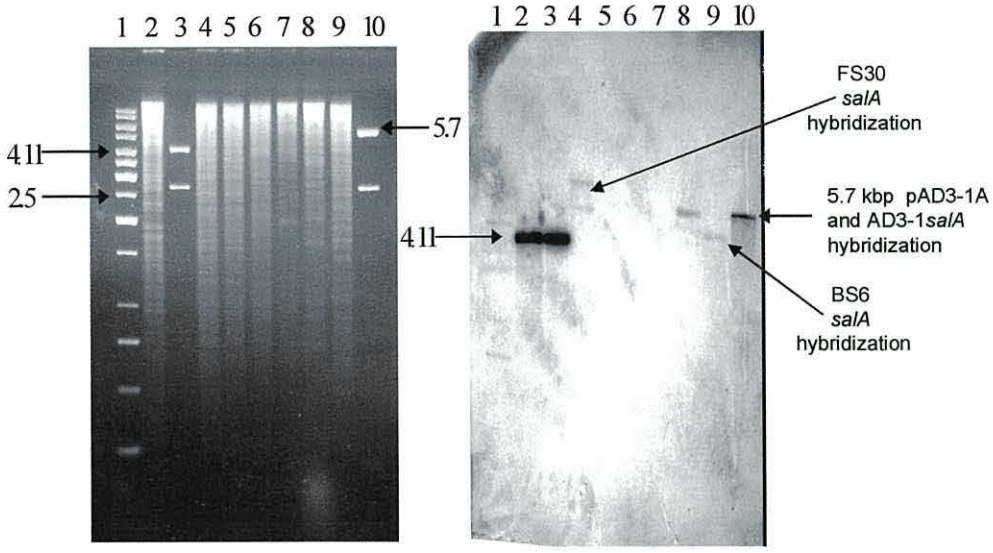


Fig.4.3. **A.** *Xba*I / *Sac*I genomic digest of *Acinetobacter* spp. (UK strains) ADP1, FS30, FS43, FS50, FS57, AD3-1 and BS6. **B.** Hybridization blot of A. **Lanes 1**. 1 kbp Molecular weight marker. **Lanes 2.** ADP1. **Lanes 3**. pADPW34. **Lanes 4**. FS30. **Lanes 5**. FS43. **Lanes 6**. FS50. **Lanes 7**. FS57. **Lanes 8-**. AD3-1. **Lanes 9**. BS6. **Lanes 10**. pAD3-1A.

Strains P1-6 and Aϕ6, which grew well with salicylate as the single carbon source did not produce any hybridization signals under any variations of the reaction conditions, e.g. variations of the stringency conditions and DNA concentration. This result suggests that either the DNA homology of these genes with the ADP1 *salA* was not sufficient to produce a signal or possibly that salicylate was catabolized via an alternative route such as gentisate in these strains. All genomic DNA digests were visualized by AGE to assess the extent of the restriction digest.

The genomic DNA of the Cardiff *Acinetobacter* spp. strains FS30, FS43, FS50, BS6 and AD3-1 (a USA strain), which grew well with salicylate were digested with *Xba*I / *Sac*I, blotted and probed with the ADP1 *salA* gene (Fig.4.3). A negative control was included in the form of strain FS57, which produced no growth with salicylate as a single carbon source.

Chemiluminescent signals were detected for the ADP1 and AD3-1 positive controls (Fig.3B lanes 2,3,8 and 10). No signal was produced for the FS57 negative control. Two other bands in lanes 4 and 9 were detected in this hybridization (strains FS30 and BS6 respectively). Strain FS30 contains two signal bands suggesting that the *Xba*I / S*ac*I restriction digest cuts within the salicylate gene sequence. Attempts to shotgun clone from this strain using a double *Xba*I / *Sac*I digest and single digests of XbaI and *Sac*I failed to isolate a clone with salicylate hydroxylase activity. The positive result for strain BS6 indicates the presence of a salicylate hydroxylase homologue. However, shotgun cloning with this strain also failed to yield a clone displaying salicylate hydroxylase activity. The single copy number of the salicylate genes per chromosome coupled with the inefficiencies of shotgun cloning increases the difficulty in the isolation of these specific genes by this method.

4.3.4 Shotgun clones

4.3.4.1 Analysis of catechol production

Shotgun cloning into the pUC19 vector isolated two clones from the genomes of strains AD3-1 and FS50 which were selected for by the production of a brown discolouration, indicating catechol production, during growth on the indicator media, LBA-salicylate (1mM) Ap (100 μg ml^{-1}). The *E.coli* DH5α host cell cannot metabolize catechol leading to an accumulation of this compound in the medium as a result of the degradation of salicylate by the cloned monooxygenase enzyme. To determine if the brown discolouration was the result of catechol formation and accumulation, a simple assay was conducted using the catechol 2, 3-dioxygenase enzyme (C23O). C23O was isolated from *P. putida* mt-2 and is the first enzyme of the *meta*-cleavage degradation pathway, converting catechol to the distinctive bright yellow compound, 2-hydroxymuconic semialdehyde (2HMS).

E. coli DH5α strains carrying the C23O gene on pWW0-6000 and the two shotgun clones were grown up overnight in LB broth Ap (100 μg ml^{-1}) in the presence and absence of salicylate (1 mM). The shotgun clones incubated in the presence of salicylate produced a dark brown discoloration; those incubated without salicylate did not. Aliquots, (1 ml), of pWW0-6000 and shotgun clone cell cultures were lysed with 20 μl of toluene and microcentrifuged to pellet the cells. The pWW0-6000 cell pellets were resuspended in the supernatants of the shotgun clones.

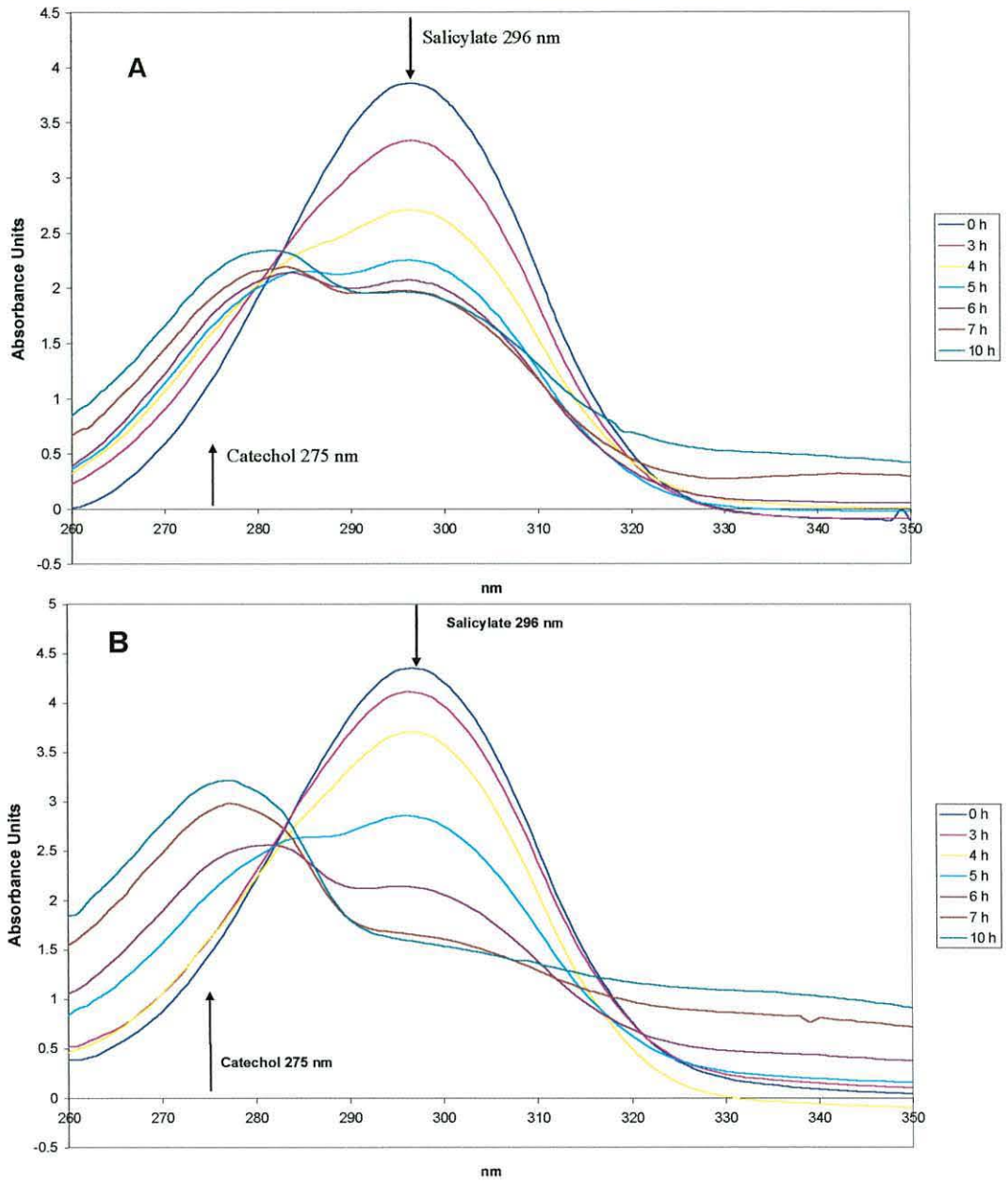


Fig.4.4 A. and B. Wavelength scan time series of metabolite transformation of salicylate to catechol by A) pAD3-1A and B) pFS50-1 respectively. The salicylate absorbance peak visible at 296 nm can be seen to decrease as a catechol peak at 275 nm increases.

Catechol production was confirmed by the immediate production of a bright yellow product assumed to be 2HMS. Negative controls, i.e. pWW0-6000 cell pellets resuspended in the culture supernatant of cells incubated in the absence of salicylate or resuspended in sterile phosphate-buffered saline, produced no yellow product. This initial simple assay provided strong evidence that these clones carried an *Acinetobacter* sp. salicylate hydroxylase gene.

The shotgun cloned plasmids were termed pAD3-1A and pFS50-1 for their source *Acinetobacter* sp. strains AD3-1 and FS50, respectively. Further characterizations of the clones were performed using restriction digestion mapping and Southern hybridization techniques. pAD3-1A, which was isolated first, was sequenced to publication quality by MWG Biotech (Chapter 2. 2.2.8). pFS50-1 was partially sequenced using MWG value reads. Salicylate hydroxylase enzyme activity of strain AD3-1 was examined using cell free extracts of pAD3-1A, wild type AD3-1 and the purified, overexpressed AD3-1 SalA (See Chapter 7).

To further characterize the dark brown compound produced by growth of the shotgun clones in the presence of salicylate, cultures of pAD3-1A and pFS50-1 were grown in 1/10 of the normal concentration of LB media and in the presence of salicylate. Samples were taken at regular intervals throughout growth to monitor the metabolite transformation. Wavelength scans were performed on a Uvikon UV/vis spectrophotometer with 1 ml samples taken from each culture and diluted 1/20. The blank cuvette contained a 1/10 dilution of culture medium removed from a control flask containing no salicylate for each time point. The results of the wavelength scans may be seen in Fig. 4.4 A and B.

4.3.4.2 AD3-1 salicylate genes sequence

The pAD3-1A sequencing results (Fig.4.6) revealed a putative salicylate operon comprised of three consecutive open reading frames (ORFs 2-4). Upstream of these ORFs, transcribed in the same orientation but separated by 0.9 kbp, lies another ORF (ORF 1) of 0.5 kbp. Blastx (NCIB database) identity / similarity searches were performed and the closest gene homologues identified are summarized in Table 4.3. (For a more detailed analysis see Chapter 6).

4.3.4.3 Isolation of *salA* homologues by degenerate PCR

Degenerate PCR amplifications performed on *Acinetobacter* spp. genomic and shotgun clone plasmid DNA isolated three putative salicylate hydroxylase genes from strains ADP1 (wild type), AD3-1, FS50, and BS6. Fragments of these genes were successfully cloned into pGem-T-easy (Table 4.2) and the inserts were sent for sequencing. NCIB Blastx analysis identified the closest gene homologue to be the ADP1 salicylate hydroxylase. These genes however, were incomplete as the degenerate primers were designed from DNA regions with close identity within the coding region in a Martinez-NW pairwise alignment of the ADP1 and the AD3-1 *salA* genes and could not be analysed for salicylate hydroxylase activity.

Table 4.3. pAD3-1A gene homologues

ORF	Gene	Closest homologue in the database	% amino acid identity / similarity	Reference (See Chapter 6 for more details)
1	Putative ferric siderophore	*Acinetobacter baylyi* strain ADP1	67 / 82	Barbe *et al.*, 2004
2	Salicylate 1-hydroxylase (*salA*)	*Acinetobacter baylyi* strain ADP1 *salA* salicylate hydroxylase	60 / 72	Jones *et al.*, 1999; 2000. Barbe *et al.*, 2004
3	Putative salicylate transporter. (MFS-superfamily) (*salK*)	*Acinetobacter baylyi* strain ADP1 *benK* benzoate transporter	58 / 75	Collier *et al.*, 1997; Barbe *et al.*, 2004
4	LysR transcriptional regulator (*salR*)	*Acinetobacter baylyi* strain ADP1 *salR* salicylate hydroxylase regulator.	54 / 72	Jones *et al.*, 1999; Barbe *et al.*, 2004

The fragment sizes amplified using the degenerate primers were 1185 bp for both ADP1 and AD3-1, 1182 bp for BS6 and 1174 bp for FS50. Clustalx alignments were analysed statistically using the Genedoc multiple alignment software, using the amplified sequence (Fig.4.5). The level of homology shared with the wild type ADP1 sequence at the DNA and the amino acid level (% identity / % similarity) was reported as 61% for AD3-1, 63% for BS6 and 60% for strain FS50. AD3-1 and FS50 *salA* sequences shared 90% identity at the DNA level, whereas the strain BS6 homologue shared 64% DNA identity with the AD3-1 sequence and 67% DNA identity with strain FS50.

```
                    *        20         *        40         *        60         *        80
AD3-1A   : ATTGGTGGCGGGATTGCAGGATTAGCACTCACCACTCAATTGGTCAAAAATAAACATTTAGATGTACAGCTTTTGAGTCTGC :   83
pFS50-A  : ATTGGTGGGGGGATTGCAGGCTTAGCACTCACGACTCAATTGGTCAAAAATAAACATTTAGATGTACATCTCTTTGAGTCTGC :   83
pBS6-A   : ATTGGTGGTGGTATTGCAGGTCTGGCATTGGCAAGTAATTTAAGTAAACATGCTCATTTAGATGTCCAAATGTTTGAATCTGC :   83
ADP1     : ATTGGTGGTGGTATTGCAGGGTCGCGCTTGCAGCCAACTTATTTAAGCAACCACATTTAGAGGTTTGCTTGTATGAAGCAGC :   83
           ATTGGTGG GG ATTGCAGG  T GCacT  C act A TT  t AAa At  aCATTTAGAtGT ca  T TtTGA tCtGC

                 *       100         *       120         *       140         *       160
AD3-1A   : ATCACAGTTTTCTGAAATTGGTGCAGGGATTTCGTTTGGTGCCAATGCAGTTAAAGCCATTCAATTATTAGGTCTAAGCCAAG :  166
pFS50-A  : ATCACAGTTTTCTGAAATTGGTGCAGGGATTTCGTTTGGTGCGAATGCAGTTAAAGCCATTCAATTATTAGGTCTAAGCCAAG :  166
pBS6-A   : ACCTCAGTTTTCTGAAATTGGGGCAGGAATTTCTTTTGGAGCAAATGCAGTCAAAGCGATTGAACTGTTAGGTTTAGCCAACG :  166
ADP1     : ACCTCAATTTTCTGAAATTGGTGCCGGAATTTCATTTGGCGCGAATGCGGTTCGTGCCATTGAGCTACTTGGTTTGGCTTCGC :  166
           A C CAgTTTTCTGAAATTGGtGCaGG ATTTC TTTGG GC AATGCaGTtaaaGCcATT Aa TatTaGGT Ta  c a g

                *       180         *       200         *       220         *       240
AD3-1A   : AGTATGAATCTATTGCAGATCAGGTAAAGACACCATACACTGATATTTGGTTTCAATGGCGTAATGGTTATACCGATGAGTAT :  249
pFS50-A  : AGTATGAATCTATTGCAGATCAGGTAAAGGCACCATACACCGATATTTGGTTTCAATGGCGTAATGGTTATACCGATGAGTAT :  249
pBS6-A   : AATATCACGCAATTGCAGATAAGGTATCTGCACCATTTCAAGATGTTTGGTTTCAATGGCGCAATGGTTATACCGATGAATAT :  249
ADP1     : AATATACCGCGATTGCAGATCAAGTATCTGCGCCATTTCAGGATGTGTGGTTTCAATGGCGAAATGGTTATAACGATGATTAT :  249
           A TAT a  C ATTGCAGATcAgGTA   gCaCCAT     GAT TtTGGTTTCAATGGCG AATGGTTATAcCGATGA TAT

           *       260         *       280         *       300         *       320         *
AD3-1A   : CTATCTGCATCGATTGCGCCACAAGTTGGACAATCTTCGGTACATCGTGCAGATTTTCTCGATCGCTTAATTCCGTTAGTGCC :  332
pFS50-A  : CTATCTGCATCGATTGCACCACAAGTCGGACAGTCTTCGGTACATCGTGCAGATTTTCTCGATCGTTTAATTCCGTTGGTGCC :  332
pBS6-A   : TTATCTGCATCGGTAGCGGCAGGTGTAGGACAGTCTTCTGTTCATCGTGCTGATTTTCTAGATGCCATTATTCCTCACATGCC :  332
ADP1     : TTGTCGAGTTCAATTTCTCCTCAGGTTGGTCAGTCTTCGGTGCATCGTGCCGATTTCTTAGATGCTATTTTAGGGAATATTCC :  332
            TaTCtgcaTCgaTtgC cCaca GT GGaCAgTCTTCgGT CATCGTGC GATTTtcT GAT    T aTtccg    TgCC

                 340         *       360         *       380         *       400         *
AD3-1A   : TTTA-ACAAATGTACATTTTAATAAACGTGTACAAGAAATTGAGGCAGATGAAGAACAAGCGACAATTCATTTTATCGATGGT :  414
pFS50-A  : TTTA-ACAAATGTACATTTTAATAAACGTGTACAAGAAATTGAGGCAGATGAAGAACAAGCGACAATTCACTTTATCGATGGT :  414
pBS6-A   : TACG-CAAAATGTCCACTTTTCTAAACGACTTGAGTCTATTGAAGAACAAGAAGATCAAGTCATTTTACATTTCAATGATGGC :  414
ADP1     : ACAGCACCAGTGTA-AGTTTAATAAAAAACTCAAATCGATTCAGGAGTACGACACACATATTGAATTGAGTTTTGAAGATGGT :  414
           t    acaAaTGTacA TTTaaTAAAcg  T  Aa   ATTgAgG a A GAagaaCAag  a a T catTTta  GATGGt

            420         *       440         *       460         *       480         *       5
AD3-1A   : CAACATATTACATTTGATTATGTGATTGGTTGTGACCGGATCCGTTCAGTCGTCCGAAACCATGTGCTTGACTCTAATCAATT :  497
pFS50-A  : CAACATATTACATTTGATTATGTGATTGGTTGTGATGGAATTCGGTCAGTTGTTCGAAATCATGTGCTTGACTCTAATCAATT :  497
pBS6-A   : AGTCGCCATGAGTGCGATTATTTGATTGGGGCTGATGGTATTCGCTCTGTTGCCCGCCAATATGTTTTAGCTACTCATCATTT :  497
ADP1     : ACATGCGCCGAAGCCGATTATGTGATTGGTGCAGATGGCATACACTCAGCCACGCGTGATTATGTGCTGCAAACCCATCAGTT :  497
             ac    t  at  GATTATgTGATTGGt  tGAtGG AT Cg TCaGt g  CG  A  ATGTgcT ga  Ct ATCA TT
```

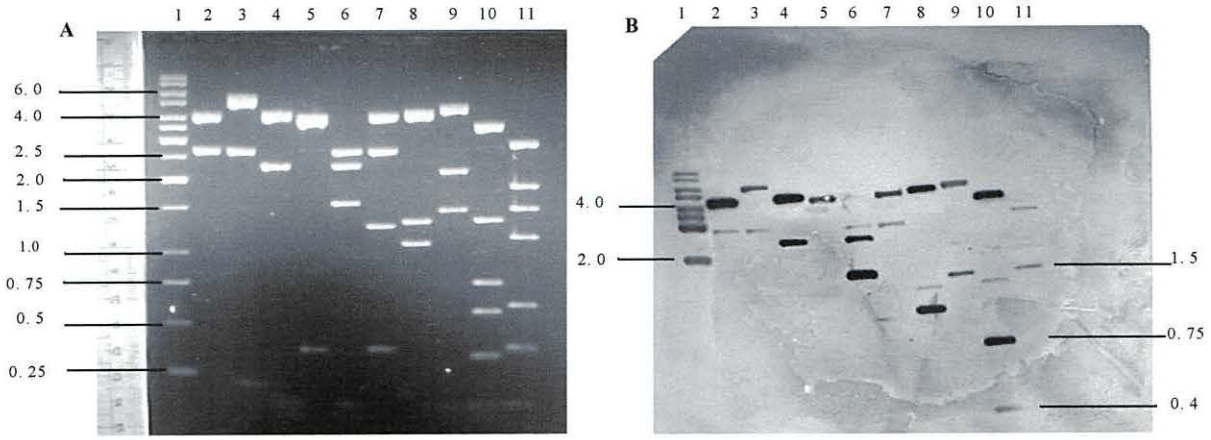
```
           00         *        520         *        540         *        560         *        580
AD3-1A   : AGCGCGTGTTGAACCTCAGTTTTCAGGGACGTGGGCATATCGTGGAATTATTAAATTTCAAGATTTTAAGCAAGCTATTG-CA :  579
pFS50-A  : AGCGCGTGTTGAACCTCAGTTTTCAGGGACTTGGGCATATCGTGGAATTATTAAATTTCAAGATTTTAAGCAAGCCATTG-CA :  579
pBS6-A   : ACTACCTGTTCAACCACGCTTTTCAGGAACTTGGGCATATCGTGGCATCATTTCACATGCTTCATTTAAACAAGCCATTGTCG :  580
ADP1     : TGCTCCAGTGCGTCCTAATTTTACTGGAACATGGGCTTACCGAGGCATTATTAAAGCAGCAGAATTTAGGCAAGCCATTGTCG :  580
           agc C tGTt aaCCtca TTTtCaGG AC TGGGCaTAtCGtGG ATtATTaaA  t  aga TTTAagCAAGCcATTG C

                    *        600         *        620         *        640         *        660
AD3-1A   : AAACTGGGAAATGATGCCGAAATTGCCGATGTACCGCAAATGCTTTTAGGGAAAGATAAACATATCTTAACTTTCCCAATCCG :  662
pFS50-A  : AAACTGGGAAATGATGCCGAAATTGCCGATGTACCGCAAATGCTTTTAGGGAAAGATAAACATATCTTAACTTTCCCAATCCG :  662
pBS6-A   : CAAATGG----CGATACAGATTTAGCAGATATTCCCCAAATGTTATTGGGTAAAGATAAGCATATTTTGACGTTCCCAATTCG :  659
ADP1     : CAGCCGGCCT-AGATGTAGAAATTGCCGACGTACCACAAATGTTTTTAGGCCAAAACAAGCACATTTTAACCTTTCCGATTCG :  662
            AactGG     GATgc GAaaTtGCcGAtgTaCC CAAATG TtTTaGG aAAgAtAA CAtAT TTaAC TTcCCaAT CG

                *        680         *        700         *        720         *        740
AD3 1A   : TCAGGGTGAAGAAATTAACATTGTGGCATTTAAGTCAGATCGCACGCAAACCGTATTAGCAGAAAATACACCATGGACATCGC :  745
pFS50-A  : TCAGGGTGAAGAAATTAACATTGTCGCATTTAAGTCGGATCGCACTCAAACCGTATTAGCAGAGAATACACCATGGACATCGC :  745
pBS6-A   : TAAGGGTGAACAAATCAATATTGTTGCCTTTTGCTCAAATCGTGAAGATACGGTACTACCAGCCGATACACCATGGACACGCC :  742
ADP1     : TCAAGGTGAAGACATTAACATTGTGGCGTTTAAGACAAACCCTGAGCAGCGTACGCTTCCAGAACATACCCCATGGACACGTG :  745
           TcAgGGTGAAgAaATtAAcATTGT GC TTTaagtCa AtCg    cA ac gta Ta CAGa  ATACaCCATGGACA   c

                *        760         *        780         *        800         *        820         *
AD3-1A   : CTGTTTCAAAACAAACGATGTTGGCTGATTTTGCAGACTGGAGCGAGAGTTGTAAGGCATTATTGGATCTGATCGACTCACCA :  828
pFS50-A  : CTGTTTCAAAACAAACGATGCTGGCTGATTTTGCAGACTGGAGCGAGAGTTGTAAGGCATTATTGGATCTGATCGACTCACCA :  828
pBS6-A   : CTGTCGATAAAGCTCAAATGCTTTCTGATTTTTCAGACTGGAGTGAGAGCTGTCAAAAGTTACTGGGTTTAATTGAGCAACCT :  825
ADP1     : CGGTAGACAAACAGGAAATGTTGGACGATTTTCAGGACTGGAGCGAAAGCTGCCGAATTTTACTCGGTTTGATTGAGCGTCCG :  828
           CtGT    AAAca    ATG TggctGATTTT caGACTGGAGcGAgAG TGt a    TTA TgG T TgAT GA   aCC

                840         *        860         *        880         *        900         *
AD3-1A   : ACAATTTGGGCATTACATGAAATTCAACCTTTAGAAAGCTATAAAACAAAAGCAATAATGTGATTTTGATTGGCGATGCTGC :  911
pFS50-A  : ACAATTTGGGCATTACATGAAATTCAACCTTTAGAAACATATAAAAACAAAAGCAATAATGTGATTTTAATTGGCGATGCTGC :  911
pBS6-A   : ACCATTTGGGCTTTACATGAAATTGAGGAGCTTTCCACCTACCAAAGTGCCTCGGGTCACGTGATACTTATGGGTGATGCCGC :  908
ADP1     : ACCTTGTGGGCACTGCACGAACTGGCCGAATTGCCGACTTATCAAAGCCACTCTGGCCGCGTCATATTAATGGGAGATGCTGC :  911
           AC aTtTGGGCatTaCAtGAAaTt a    tT    Ac TAt AAA c a      t a GTgAT tT AT GG GATGCtGC

               920         *        940         *        960         *        980         *
AD3-1A   : TCATGCCATGTTACCGCATCAAGGTGCTGGTGCCGGACAAGGTTTAGAGGATGCTTTGATTTTAGCTAAATTATTAGAAAACC :  994
pFS50-A  : TCATGCCATGTTACCGCATCAAGGTGCTGGTGCCGGACAAGGTCTAGAGGATGCTTTGATTTTAGCTAAATTATTAGAAAACC :  994
pBS6-A   : ACATGCCATGTTACCTCACCAAGGGGCTGGAGCAGGGCAAGGTCTTGAAGATGCTTTGATCCTTGCAGCCTTACTAGCGA-CA :  990
ADP1     : GCACGCCATGCTTCCACATCAGGGTGCCGGAGCAGGACAAGGGCTTGAAGATGCACTCACGCTCAAAGTATTGTTTGAGCACA :  994
            CAtGCCATGtTaCC CAtCAaGGtGCtGG GC GGaCAAGGtcT GA GATGCttTgAt  T gc   aTTatTaGa aaC

           1000        *       1020        *       1040        *       1060        *       108
AD3-1A   : CTGAGCTGACATCTGAATCTA-TCAAAGCTGTATCAGAAATCTATGAGCAAGTTCGATTAGCGCGTGCCGTAAAAGTTCAAAA : 1076
pFS50-A  : CTGAGCTGACATCTGAATCTA-TCAAAGCTGTATCAGAAATCTATGAGCAAGTTCGATTAGCGCGCGCCGTAAAAGTTCAAAA : 1076
pBS6-A   : AAAAACTTACACTCGGATCAACTCACCGATGTGTCTTCAATTTATGAAAAGCTGCGTGTAAAACGAGCTTGCCGTGTTCAACA : 1073
ADP1     : CTGAGCTGACTGTTGAAGATTTACCGCGA-GTTTCTGCAATTTATGAACAGATCCGAAAAGAACGCGCCTGCAAGTTCAGCG : 1076
           ctgAgCTgACa  tGaAtcta tCa  G tGT TC g AAT TATGA cA  T CGa tAg  CG GCc   aa GTTCAa a

           0         *       1100        *       1120        *       1140        *       1160
AD3-1A   : TACCTCGCGTGAATCAGGTGAAATCTATGAGCTATATTCATCGAAATATCCTGATTTTGAATCAATAGGCCAACATTTGGCAC : 1159
pFS50-A  : TACCTCGCGTGAATCAGGTGAAATCTATGAGCTATATTCATCGAAATATCCTGATTTTGAATCAATAGGCCAACATTTGGCAC : 1159
pBS6-A   : AACCTCTCGTGAGTCTGGTGAAATTTATGAGTGCTATTCAACTCAGTACTCAACGTTTGCAGAAATTGGCGAGCACCTTGAAC : 1156
ADP1     : TACTTCGCGTGAGTCTGGGCAAATATATGAACTTAACTCAGCACTGTATCCAAGCTTTGAAGCAGTGGGTGCACATTTGCAAA : 1159
           tACcTCgCGTGA TC GGtgAAAT TATGAgct tAtTCA C  a TAtcC    TTTGaA cAaT GGc aaCAttTgg Ac

                *       1180        *       1200        *       1220        *       1240
AD3-1A   : AGCGTTTTGACTGGTTGTGGCAGCATTCGCTTGAAGATGATATTCAGACAGCTAAAGCAGCATTAGCAAAACAACTACAGCAC : 1242
pFS50-A  : AGCGTTTTGACTGGT------------------------------------------------------------------- : 1174
pBS6-A   : ATCGTTTTGACTGGTTATGGCAGCAT-------------------------------------------------------- : 1182
ADP1     : ACCGTTTTGACTGGTTATGGCAGCATGATTTAGCACAAGACATGTTAGCCGCGCGAGCAGCAATACAACCTGTTGCAACGATT : 1242
           A CGTTTTGACTGGTt tggcagcat

                *       1260
AD3-1A   : ACTGAAATGACGACTTAA : 1260
pFS50-A  : ------------------ :    -
pBS6-A   : ------------------ :    -
ADP1     : TAA--------------- : 1245
```

Fig.4.5. Clustalx software generated *salA* alignment of homologues isolated by degenerate PCR. The known sequence, extending beyond the degenerate primer sequences, for the ADP1 and AD3-1 *salA* have been included in this figure to illustrate the position of the degenerate primers relative to that of the complete gene sequence.

4.3.4.4 pAD3-1A Southern hybridization

Identical restriction digests of pAD3-1A and pADPW34, were probed alternately with the ADP1 *salA* gene (Fig.4.6A and B) and the ADP1 *salR* gene (Fig.4.8 A and B) and the DNA-DNA hybridization patterns were compared to investigate *sal* gene homologies. Bands of hybridized DNA were observed using the ADP1 *salA* as the probe (Fig.4.6B). The hybridized bands were sized and fitted to the restriction map and the known sequences of the two salicylate operons (Fig.4.7) and corresponded to the position of *salA* in both plasmids. These results correlate with the location of the pAD3-1A *salA* homologue identified from the sequencing results and show that the ADP1 *salA* gene possesses a degree of identity with that of AD3-1.

The ADP1 *salR* gene probe showed only very weak interaction with the pAD3-1A DNA whilst hybridising strongly with pADPW34 DNA (Fig. 4.8, Fig. 4.9). These very faint signal bands indicate the presence of a *salR* gene homologue with a low level of DNA identity with the ADP1 *salR* sequence. The pAD3-1A sequence data show a DNA identity of 50% between the ADP1 and AD3-1 *salR*s which is quite borderline for hybridization to occur. The strongly hybridized bands visible in the digests of pADPW34 (Fig.4.8 and Fig.4.9) correlate with the position of *salR* within the known ADP1 sequence (Jones *et al.*, 2000).



Restriction digest fragments (kbp)				
Restriction enzymes	**Lane**	**pADPW34**	**Lane**	**pAD3-1A**
*Xba*I - *Sac*I	2	4.1, 2.7	3	5.7, 2.7
*Hin*dIII	4	4.3, 2.5	5	4.2, 4.11, 0.38
*Hin*dIII - *Eco*RI	6	2.7, 2.3, 1.6, 0.23, 0.1	7	4.11, 2.7, 1.3, 0.38
*Cla*I	8	4.1, 1.3, 1.1	9	4.6, 2.2, 1.5, 0.15
*Cla*I - *Hin*dIII	10	3.5, 1.4, 0.7, 0.6, 0.4, 0.1	11	2.8, 1.8, 1.5, 1.1, 0.6, 0.38, 0.15

Fig.4.6 A and B. Agarose Southern gel (A) and blot (B) of pADPW34 and pAD3-1A digests probed with ADP1 *salA*. The adjoining table describes the restriction fragments (kbp) generated for each digest. Coloured numbers indicate hybridized bands: red - strong hybridization, blue - weak or non-specific hybridization, un-highlighted - no hybridization.

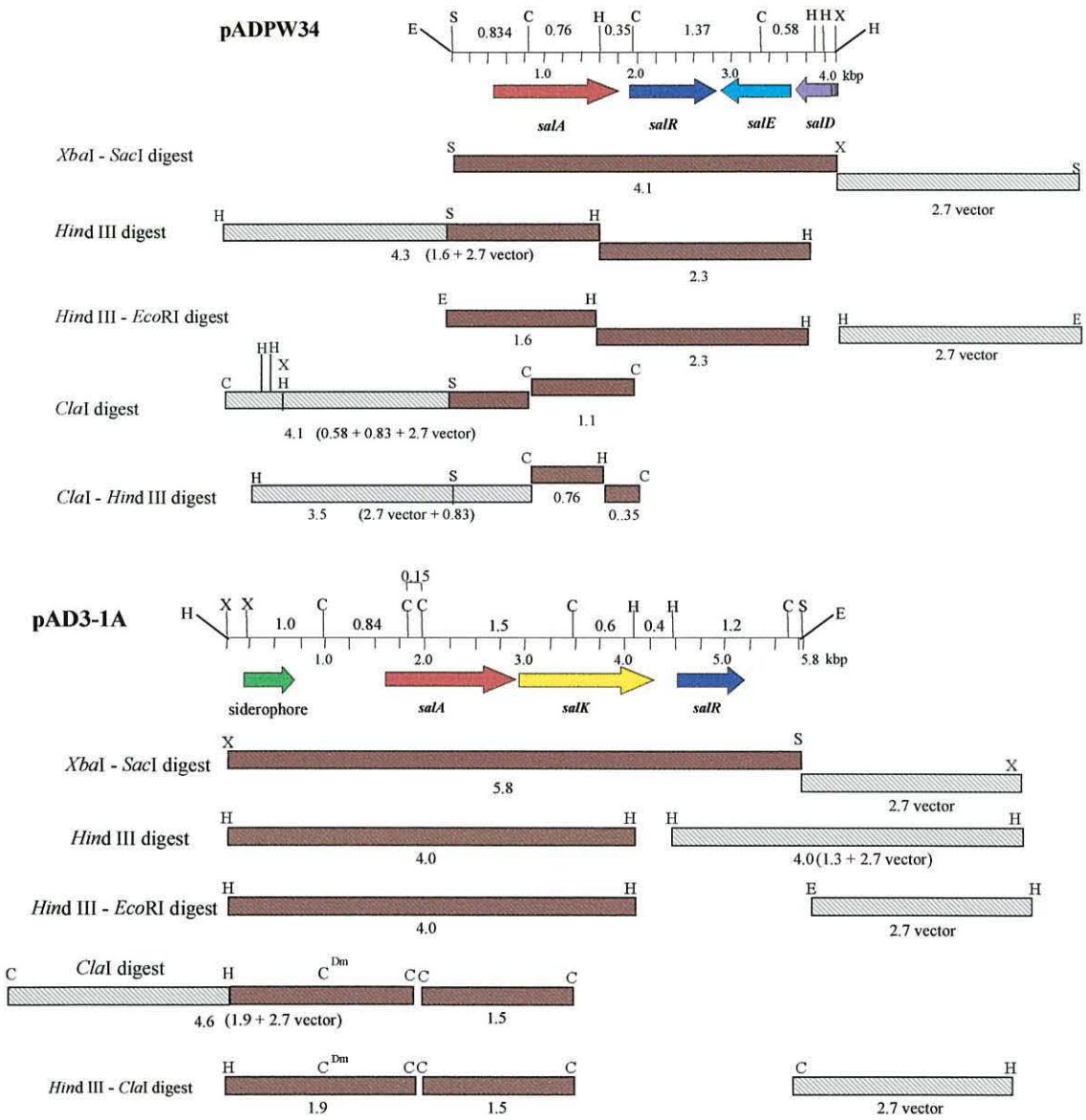
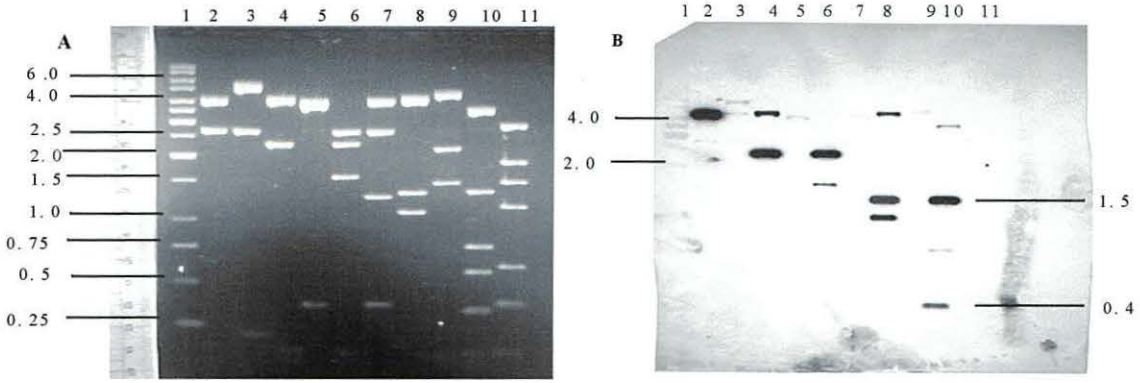


Fig.4.7. Hydbridization patterns of pADPW34 and pAD3-1A probed with ADP1 *salA*. Green-red boxes: strong hybridization. Grey-white boxes: weak/ non-specific hybridization. Letters: restriction sites, X - *Xba*I, S - *Sac*I, H - *Hin*dIII, C - *Cla*I, E - *Eco*RI, C^{Dm} - *Cla*I site protected by Dam methylation.



Restriction digest fragments (kbp)				
Restriction enzymes	**Lane**	**pADPW34**	**Lane**	**pAD3-1A**
*Xba*I - *Sac*I	**2**	4.1, 2.7	**3**	5.7, 2.7
*Hin*dIII	**4**	4.3, 2.5	**5**	4.2, 4.11, 0.38
*Hin*dIII - *Eco*RI	**6**	2.7, 2.3, 1.6, 0.23, 0.1	**7**	4.11, 2.7, 1.3, 0.38
*Cla*I	**8**	4.1, 1.3, 1.1	**9**	4.6, 2.2, 1.5, 0.15
*Cla*I - *Hin*dIII	**10**	3.5, 1.3, 0.7, 0.6, 0.35, 0.1	**11**	2.8, 1.8, 1.5, 1.1, 0.6, 0.38, 0.15

Fig. 4.8 A and B. Agarose Southern gel (A) and blot (B) of pADPW34 and pAD3-1A digests probed with ADP1 *salR*. The adjoining table describes the restriction fragments (kbp) generated for each digest. Highlighted numbers indicate strength of hybridization: red - strong hybridization, blue - weak or non-specific hybridization, black - no hybridization.

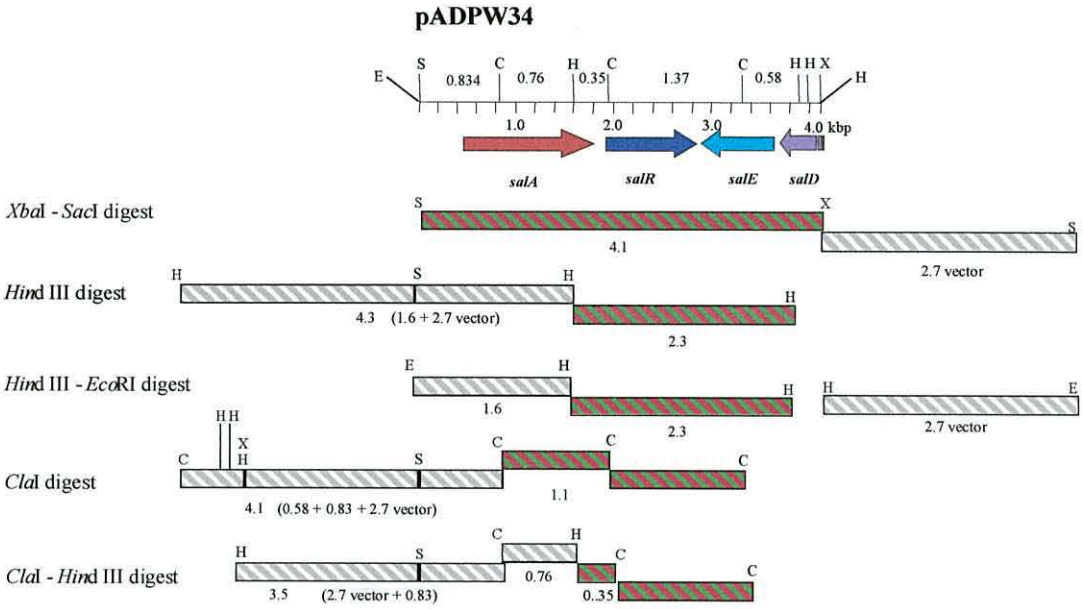


Fig. 4.9. Map and hybridization patterns of pADPW34 probed with the ADP1 *salR*. Green and red striped boxes - strong hybridization. Grey and white striped boxes weak or non-specific hybridization. Letters represent restriction sites: X - *Xba*I, S - *Sac*I, H - *Hin*dIII, C - *Cla*I, E - *Eco*RI.

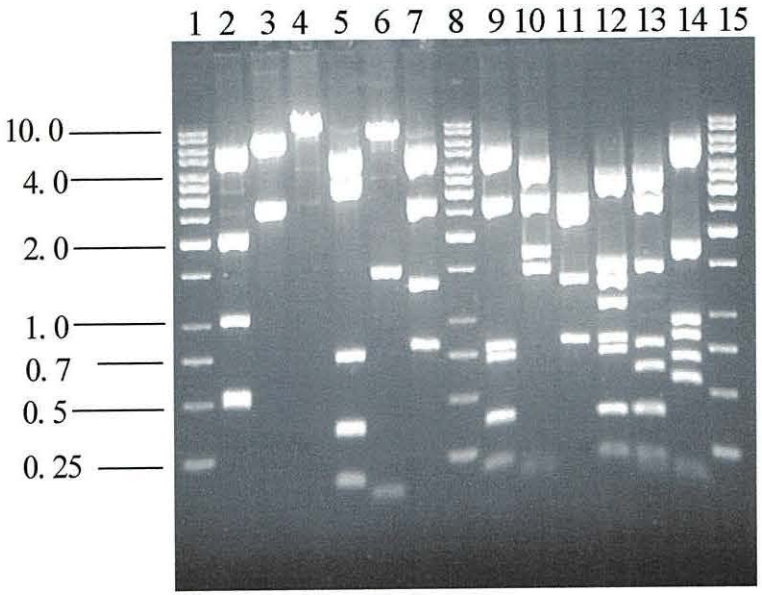
4.3.4.5 pFS50-1 analysis

The shotgun clone, pFS50-1 isolated from the genome of *Acinetobacter* sp. strain FS50 was digested with a variety of restriction enzymes (Fig.4.10) in order to create the plasmid map depicted in Fig.4.10.2. To confirm the location of the putative salicylate hydroxylase gene within the sequence Southern hybridization analysis was performed on selected digests (Fig.4.10.1). The

putative *salA* homologue clone pFS50-pA and the wild type *salA* gene contained in pADP1A were used as probes (Table 4.2). When the pFS50-1 blot was probed with pFS50-pA positive control strong hybridization signals were observed (Fig.4.10.1B). For the blot probed with pADP1A weaker hybridization signals of the same size were observed. A signal corresponding to the length of the pFS50-1 insert (~7.0 kbp) was detected in the genomic DNA when probed with pFS50-pA (Fig.4.10.1B lane 2). A band of this size for pFS50-1 was detected with both probes. The genomic DNA signal was not detected with the wild type *salA* probe. This indicates that the gene copy number present in the genomic DNA was not sufficient for hybridization to occur with the pADP1A probe.

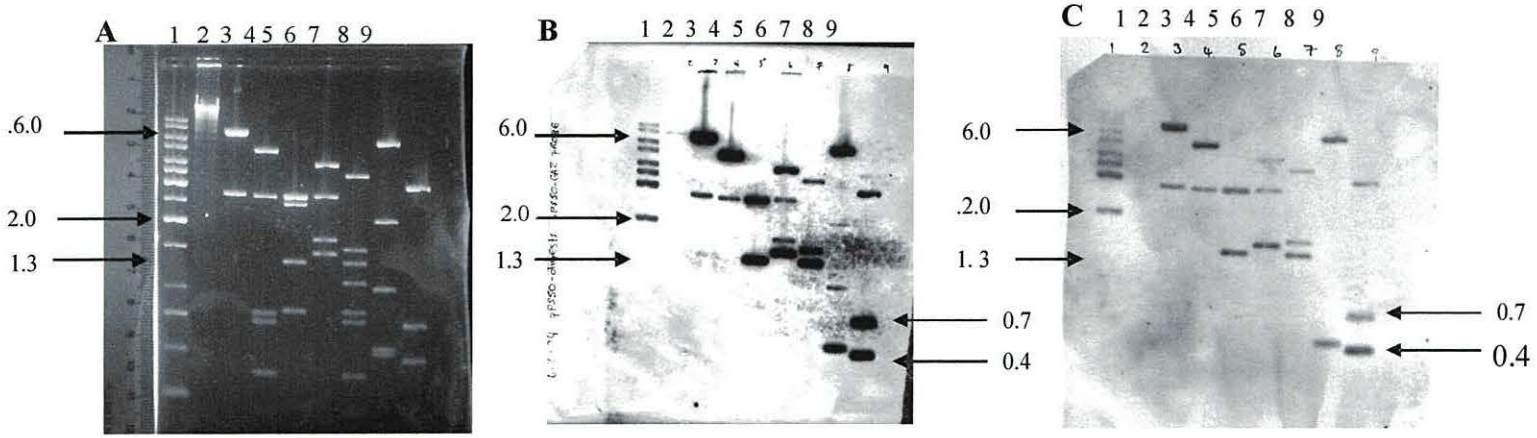
Both probes hybridized to a large 4.5 kbp fragment generated by the *Sac*I / *Hin*dIII digest (lane 4 Fig. 4.10.1A and Fig. 4.10.1B and C). It was known from the partial pFS50-1 sequencing results that the last 1.8 kbp of the sequence contained 0.8 and 0.4 kbp *Hin*dIII fragments. Therefore the large 4.5 kbp fragment can be deduced to encompass the central region of the insert.

In the *Sac*I / *Xba*I digest a 2.7 kbp fragment and a 1.3 kbp *Xba*I generated fragment hybridized to both probes. From the sequencing results of the *salA* of pFS50-pA it is known that a *Xba*I site is present within the gene. Therefore, the1.3 kbp *XbaI* fragment contains the majority of the *salA* gene and the 2.7-2.6 kbp fragment contains the last ~300 bp of the gene.



Lane	Restriction digest	Approximate fragment size (kbp)
1, 8, 15	1 kb DNA ladder	10.0, 8.0, 6.0, 5.0, 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.0, 0.7, 0.5, 0.25
2	*Eco*RI	4.5, 2.0, 1.0, 0.6,0.5
3	*Sac*I	7.0, 2.7
4	*Bam*HI	9.0
5	*Hin*dIII	4.5, 3.0, 0.8, 0.4, 0.2
6	*Cla*I	8.0, 1.5, 0.1
7	*Xba*I	5.0, 2.7, 1.3, 0.8
9	*Sac*I / *Hin*dIII	4.5, 2.7, 0.8, 0.7, 0.4, 0.2
10	*Sac*I / *Cla*I	4.0, 2.7, 1.8, 1.5, 0.2
11	*Sac*I / *Xba*I	2.7, 2.6, 1.3, 0.8
12	*Hin*dIII / *Xba*I	3.5, 1.5, 1.3, 1.1, 0.8, 0.7, 0.4, 0.2
13	*Hin*dIII / *Cla*I	3.5, 2.5, 0.8, 0.6, 0.4, 0.2
14	*Xba*I / *Cla*I	4.5, 1.8, 1.0, 0.8, 0.65, 0.6, 0.1

Fig. 4.10. Agarose gel of pFS50-1 multiple restriction digests. Adjoining table: Restriction enzymes used and the approximate fragment sizes generated (kbp).



Lane	DNA and restriction enzymes	Restriction fragments (approx. kbp) hybridized to the pFS50-pA probe	Restriction fragments (approx. kbp) hybridized to the pADP1A probe
1	1 kbp DNA ladder		
2	FS50 genomic DNA - *Sac*I	7.0	-
3	pFS50-1 - *Sac*I	7.0, 2.7	7.0, 2.7
4	pFS50-1 - *Sac*I / *Hin*dIII	4.5, 2.7, 0.8, 0.4, 0.2	4.5, 2.7, 0.8, 0.4, 0.2
5	pFS50-1 - SacI / XbaI	2.7, 2.6, 1.3, 0.8	2.7, 2.6, 1.3, 0.8
6	pFS50-1 - SacI / ClaI	4.0, 2.7, 1.8, 1.5, 0.2	4.0, 2.7, 1.8, 1.5, 0.2
7	pFS50-1 - XbaI / HindIII	3.5, 1.5, 1.3, 1.1, 0.8, 0.7, 0.4, 0.2	3.5, 1.5, 1.3, 1.1, 0.8, 0.7, 0.4, 0.2
8	pFS50-1 - EcoRI	5.0, 2.0, 1.0, 0.5	5.0, 2.0, 1.0, 0.5
9	pFS50-pA - EcoRI	3.0, 0.7, 0.4	3.0, 0.7, 0.4

Fig. 4.10.1 A.pFS50-1 restriction digest agarose gel. **B**. Southern hybridization blot of A using the pFS50-pA probe. **C**. Southern hybridization blot of **A** using the pADP1A probe. Highlighted in red - strong hybridization. Highlighted in blue - non-specific hybridization.

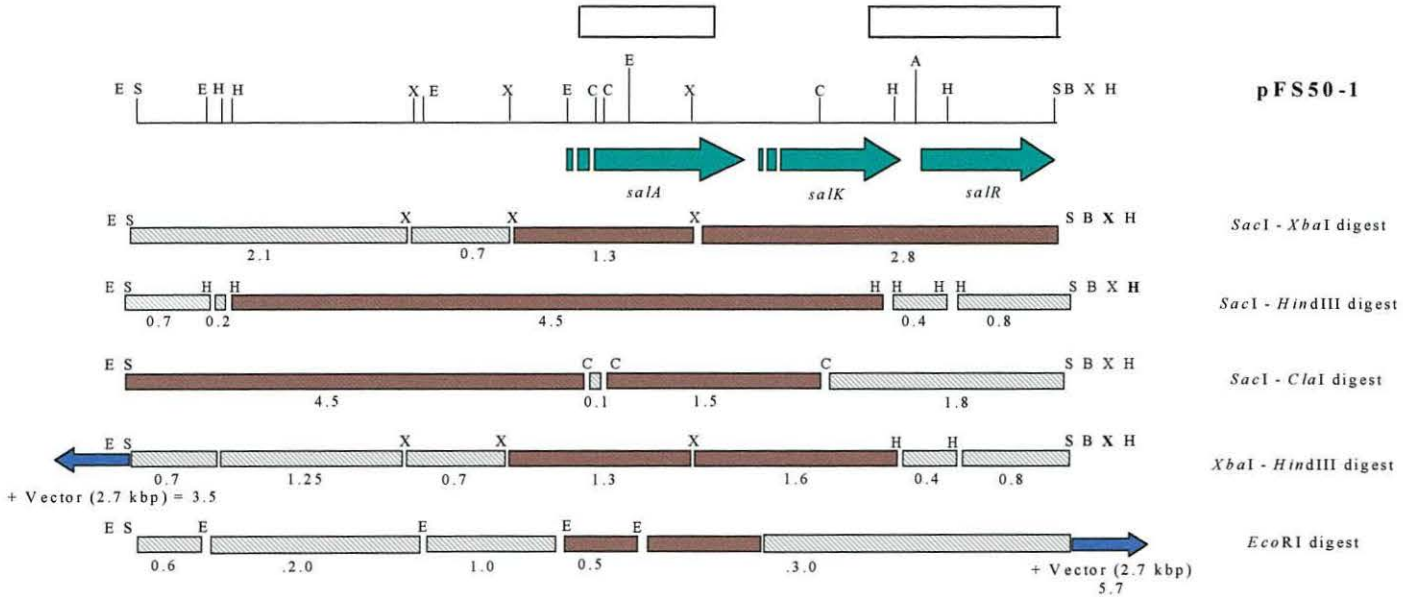


Fig.4.10.2 Approximate plasmid map of pFS50-1 as deduced using data from Fig.4.10 and Fig.4.10.2. Green and red striped boxes: areas that hybridized to pFS50-pA and pADP1A probes in Southern hybridization studies (Fig.4.10.2).. Grey and white striped boxes - no or non-specific hybridization. Letters represent restriction sites: S - *Sac*I, B - *Bam*HI, X - *Xba*I, H - *Hin*dIII, E - *Eco*RI, C - *Cla*I. Open boxes above the map indicate the regions of the plasmid insert that have been sequenced.

The hybridization signals of the *Xba*I / *Hin*dIII digest localize the *salA* gene to 1.6 and 1.3 kbp fragments. The 1.3 kbp fragment is *Xba*I generated as described above and the 1.6 kbp fragment is cleaved as a *Xba*I / *Hin*dIII fragment containing the terminal end of the *salA* gene (Fig.4.10.2). The *Eco*RI cleavage of the pFS50-pA insert produced two fragments sized 0.7 and 0.4 kbp that hybridized to each probe (Fig.4.10.1 lanes 9). *Eco*RI cleavage of pFS50-1 produced two hybridization signals of 4.5 and 0.5 kbp. From the partial sequencing results it is known that the last 1.8 kbp and the first 0.6 kbp of the sequence contains no *Eco*RI site and that an *Eco*RI site is present ~0.5 kbp from at the 5' end of the degenerate *salA*. Therefore, the 4.5 kbp band contains a 0.7 kbp fragment of *salA* and downstream sequence DNA sequence including the vector (Fig.4.10.2), an *Eco*RI site must lie immediately upstream of 5' *salA* to produce the 0.5 kbp signal.

This analysis identifies the presence of an ADP1 *salA* gene homologue in *Acinetobacter* sp. strain FS50. The restriction digestion and Southern hybridization results have provided an approximate map of the sequence cloned into pUC19 (Fig.4.10.2). The partial sequencing results reveal a high level of homology between the *sal* genes of strains FS50 and AD3-1, both of which share approx. 60 % DNA identity with the ADP1 *salA* gene.

4.4 Discussion

Acinetobacter strains can degrade a wide and diverse range of organic compounds such as aliphatic alcohols, dicarboxylic and fatty acids, alkanes and many aromatic hydrocarbons (Juni, 1978; Parke *et al.*, 2001). The majority of the *Acinetobacter* strains used in this study grew with *p*-hydroxybenzoate and or benzoate as sole carbon sources. However, only 8 of the environmental isolates displayed growth with 2.5 mM salicylate. Four of these strains, AD3-1, AD321, FS30 and BS6 were shown to hybridize to the ADP1 *salA* gene probe in Southern hybridization analysis. Those strains which showed poor to moderate growth on 1.5 mM salicylate did not produce any hybridization signals in Southern blots which may be due to a lack of homology between these salicylate genes and the ADP1 *salA* gene used as the probe.

Shotgun cloning isolated two putative salicylate operons from the genomes of strains AD3-1 and FS50. Analysis of the pFS50-1 clone by partial sequencing revealed a high level of homology with the insert of pAD3-1A. In genomic DNA Southern blot analysis, hybridizaion did not occur with strain FS50 when probed with the ADP1 *salA* gene. However, hybridization was observed to occur when the pFS50-pA *salA* homologue, isolated by degenerate PCR from this strain, was used as the probe. PCR using degenerate primers also isolated a putative salicylate hydroxylase gene from the genome of strain BS6. Homologies between the three putative *salA* genes and the ADP1 wild type were approx. 60% at both the DNA and amino acid level. The putative salicylate operons of strains AD3-1, an USA isolate and FS50, an isolate from Cardiff, Wales UK, are very closely related, despite the geographical difference in their isolation. The basic ARDRA profiles performed on these strains show identical restriction patterns when digested with *Cfo*I, *Rsa*I and *Msp*I. In

addition, DNA and amino acid sequence alignments revealed 90 % identity between the *salA* genes. Alignments between the putative *salR* genes and approx. 150 bp of the region upstream of the putative translational start codon of the shotgun cloned inserts gave a DNA identity of 96%. An alignment of the AD3-1 and FS50 *salR* genes with the ADP1 *salR* gene revealed 60% DNA identity. The identity between the putative AD3-1 and FS50 salicylate *salA* and *salR* genes including the upstream region suggest that the putative salicylate transporter gene, termed *salK* is probably also present within this sequence. In the ADP1 salicylate operon the *salA* and *salR* genes are co-transcribed convergent to the genes *salE* and *salD*, for salicylate ester degradation and the putative salicylate ester porin gene, respectively which are also co-transcribed (Jones *et al.*, 2000). No specific salicylate permease gene, such as the putative *salK* gene, has as been described in ADP1. The organisation and regulation of the AD3-1 salicylate genes will be described further in Chapter 6.

The AD3-1 and FS50 *sal* genes isolated as putative operons and single *salA* genes isolated by degenerate PCR were used as the heterologous donor DNA to transform the ADP1 deletion strains ADPW257-261 to prototrophy. The wild type ADP1 salicylate genes were also exploited as homologous donor DNA to set a baseline for the frequency of transformation.

Chapter 5

5.1 Introduction

A. baylyi strain ADP1 becomes competent for natural transformation during the exponential growth phase. The ability of ADP1 to take up DNA from any source has facilitated research into the mechanisms of horizontal gene transfer. The incoming DNA may be incorporated into the recipient genome in a RecA dependent manner via homologous recombination, providing a sufficient region of flanking homology exists or is provided within the cell (Chapter 1 – 1.3.1.3). This chapter describes the natural transformation of strains ADPW257 – ADPW261 using wild type ADP1 DNA as a standard for gene repair of each mutant strain, against which the transformation frequencies obtained using heterologous DNA as the donor could be compared.

Chapter 5 preface

To quantify the frequency of transformation and horizontal gene transfer events between salicylate hydroxylase gene homologues isolated from *Acinetobacter* spp. a gene rescue assay was employed. Transfer events were detected by restoration of the $salA^+$ phenotype demanded by growth on salicylate as a sole carbon and energy source. The *salA* gene (encoding salicylate 1-hydroxylase) of *Acinetobacter baylyi* strain ADP1 was mutated to incorporate small deletions (4 bp) within five different gene regions, creating *A. baylyi* strains ADPW251 – ADPW261 (Chapter 3). These five strains were used as the DNA recipients in natural transformation assays to quantify the transformation frequency between

the *A. baylyi* strain ADP1 *salA* and homologous genes originating from *Acinetobacter* sp. strains AD3-1 and FS50. These *Acinetobacter* spp. were identified as containing a functional salicylate hydroxylase with 60% identity to that of *A. baylyi* strain ADP1 (Chapter 4). Three different types of both the homologous and the heterologous donor DNA were used in the natural transformation assays – total purified genomic DNA, plasmid DNA containing a salicylate operon, and plasmid DNA containing just *salA*. The DNA sources used are listed below.

- **Homologous *A. baylyi* strain ADP1 DNA**

1). The wild type *A. baylyi* strain ADP1 salicylate operon cloned into the high copy number vector, pUC18, as a 4.1 kbp *Xba*I – *Sac*I fragment (pADPW34) containing the complete coding sequences, *salA*, *salR*, *salE* and a truncated copy of *salD*.

2). A truncated copy (~1.1 kbp) of the *A. baylyi* strain ADP1 *salA* isolated by degenerate PCR and cloned into pGem-T-easy (pADP1A).

3). Total purified genomic DNA isolated from *A. baylyi* strain ADP1.

- **Heterologous *Acinetobacter* sp. strain AD3-1 DNA**

1). The putative salicylate operon isolated from *Acinetobacter* sp. strain AD3-1 cloned into the high copy number vector, pUC19, as a 5.78 kbp *Xba*I – *Sac*I fragment (pAD3-1A) containing four open reading frames designated *orf1* (encoding a putative siderophore protein), *salA*, *salK* and *salR*.

2). A truncated (~1.1 kbp) *salA* gene (pAD3-12A), isolated by degenerate PCR from the genome of *Acinetobacter* sp. strain AD3-1 and cloned into pGem-T-easy.

3). Total purified genomic DNA isolated from *Acinetobacter* sp. strain AD3-1.

- **Heterologous *Acinetobacter* sp. strain FS50 DNA**

1). The putative salicylate operon isolated from *Acinetobacter* sp. strain FS50 cloned into the high copy number vector, pUC19, as ~7.0 kbp SacI fragment (pFS50-1) putatively containing *salA*, *salK* and *salR*.

2). A truncated copy (~1.1 kbp) of the *salA* gene, (pFS50-pA), isolated by degenerate PCR from the genome of *Acinetobacter* sp. strain FS50, cloned into pGem-T-easy.

3). Total purified genomic DNA isolated from *Acinetobacter* sp. strain FS50.

5.2 Materials and methods

5.2.1 Donor DNA

The homologous and heterologous donor DNA is described in Fig.5.2.1 and Table 5.2.1. Homologous donor DNA derived from the wild type strain ADP1 was provided in the form of plasmid and genomic DNA. Heterologous donor DNA was provided by plasmid and genomic DNA isolated from strains AD3-1 and FS50 (Fig.5.2.1. See also Chapter 4 - 4.1. Tables 4.2 and 4.3). The donor DNA was provided as homologous or putative salicylate operons, a single truncated *salA* gene or homologue (approx 1180 bp in length) and purified chromosomal DNA from each strain.

5.2.2 Transformation frequency assays

The natural transformation assay method is described in Chapter 2 - (section 3 - 2.3.3). Following a 1h incubation period with the donor DNA, the transformation cultures were placed directly on ice and when necessary, serially diluted in phosphate-buffered saline, then plated onto selective minimal medium supplemented with 2.5 mM salicylate. Viable cell counts were determined after plated onto 10 mM succinate minimal medium. As a control for spontaneous reversion of the created mutants, all transformations included replicates performed without the addition of donor DNA. Three replicate transformations were each performed in duplicate. The transformation frequency was calculated as the number of transformants / number of viable cells.

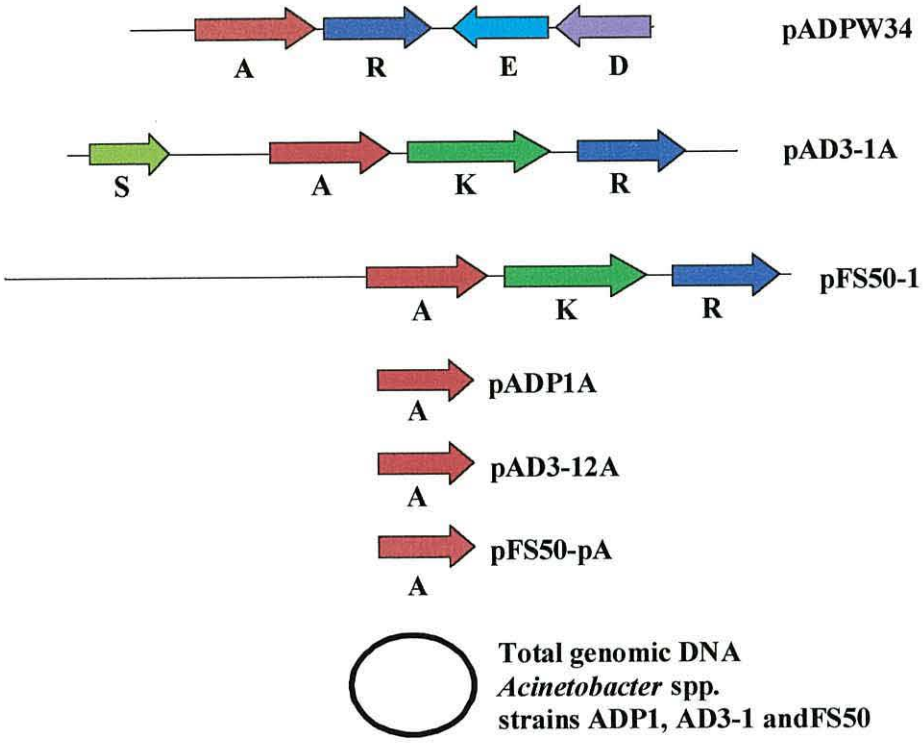


Fig. 5.2.1. Sources of homologous and heterologous donor DNA.

Homologous DNA - pADPW34, pADP1A, ADP1 genomic DNA.

Heterologous DNA - pAD3-1A, pAD3-12A, AD3-1 genomic DNA; pFS50-1, pFS50-pA and FS50 genomic DNA.

Abbreviations: S – siderophore sequence, A - *salA* wild type or homologue salicylate hydroxylase, R- *salR*, wild type or putative heterologous LysR type transcriptional regulator gene, E - *salE*, salicylate esterase gene, D - *salD*, salicylate porin gene, K - *salK*, putative salicylate transporter gene.

Donor DNA			
Salicylate operons	**Genes**	**Source and size**	**Restriction enzyme**
pADPW34	*salA*, *salR*, *salE*, *salD*	ADP1 (WT) 4.1 kbp	*Sal*I
pAD3-1A*	*salA*, *salK*, *salR*	AD3-1 (HT) 5.7 kbp	*Sal*I
pFS50-1*	*salA*, *salK*, *salR*	FS50 (HT) ~ 7.0 kbp	*Bam*HI
Salicylate hydroxylase genes			
pADP1A	*salA***	ADP1 (WT) ~1.1 kbp	*Sac*I
pAD3-12A	*salA***	AD3-1 (HT) ~1.1 kbp	*Sac*I
pFS50-pA	*salA***	FS50 (HT) ~1.1 kbp	*Sac*I
***Acinetobacter* spp. Genomic DNA**			
ADP1	total chromosomal DNA	ADP1 (WT) 4.6 Mbp	*Sal*I
AD3-1	total chromosomal DNA	AD3-1 (HT) NK	*Sal*I
FS50	total chromosomal DNA	FS50 (HT) NK	*Sal*I

Table 5.2.1. Type and source of the donor DNA used in natural transformation assays of the five *Acinetobacter baylyi* strain ADP1 Δ*salA* mutants, ADPW257-ADPW261. * putative salicylate operons. ** truncated *salA* genes amplified from pADPW34, pAD3-1A and pFS50-1 by degenerate PCR (Chapter 4 - 2.4.2.4). Restriction enzymes listed were used to linearize the plasmid donor DNA or completely digest the genomic donor DNA. Abbreviations: WT - wild type DNA, HT - heterologous DNA, NK - not known.

5.3 Results

5.3.1 Effects of increasing DNA concentration

In order to determine an optimum DNA concentration for use in subsequent transformation assays, strain ADPW258 was incubated with increasing concentrations of pADPW34 donor DNA (Fig.5.3.1). For a better representation of the data, the transformation frequency is plotted on a semi-logarithmic scale. The graph shows that the transformation frequency with increasing amounts of donor DNA approaches saturation kinetics. These data are in agreement with the results of the effects of plasmid concentration on the transformation frequency observed by Palmen *et al.* (1993). As a result of these data, 2.5 μg of DNA per approx 10^8 cells was chosen as an optimal concentration of saturating donor DNA.

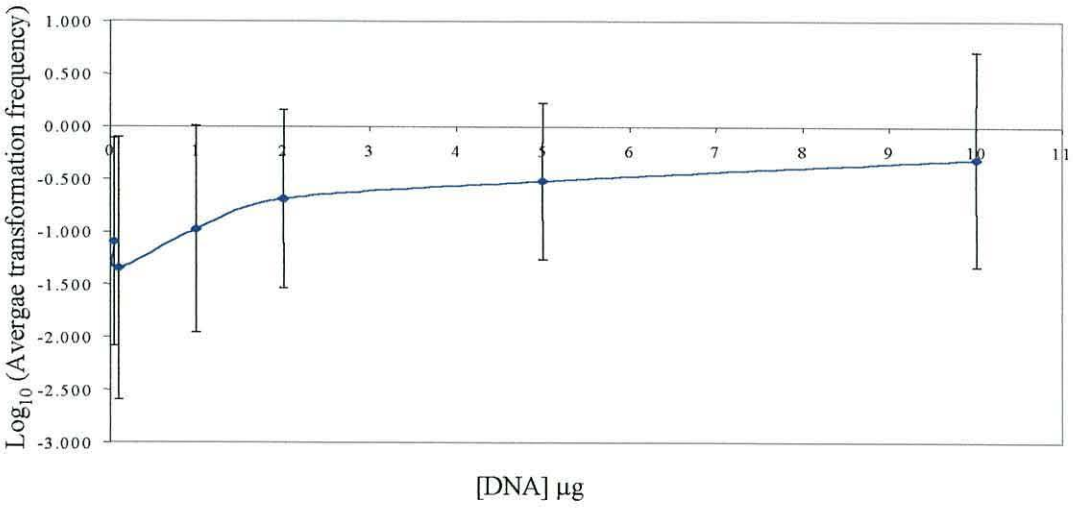


Fig.5.3.1. Effects of increasing concentration of pADPW34 (4.11 kbp) DNA on the transformation efficiency of *Acinetobacter baylyi* strain ADPW258. Log_{10} ± standard deviation, n=9 (average transformation frequency).

5.3.2 Standardization of donor DNA

Aliquots of donor DNA were quantified and combined to give a sufficient amount of DNA for each assay being performed. The combined DNA was then digested and used as donor DNA. This step was performed for all transformation assays in order to standardize the DNA concentration between replicates. The typical donor DNA concentrations used are shown in Fig.5.3.2. The genomic DNA was digested with *Sal*I, as this restriction enzyme produced quite large fragment sizes and was not present within the salicylate operon sequences of the donor DNA. Wild type transformations using genomic DNA digested with *Hin*dIII, performed in parallel with the *Sal*I digested genomic donor DNA, produced no transformants whereas detectable levels of transformants were produced using DNA prepared with the latter enzyme.

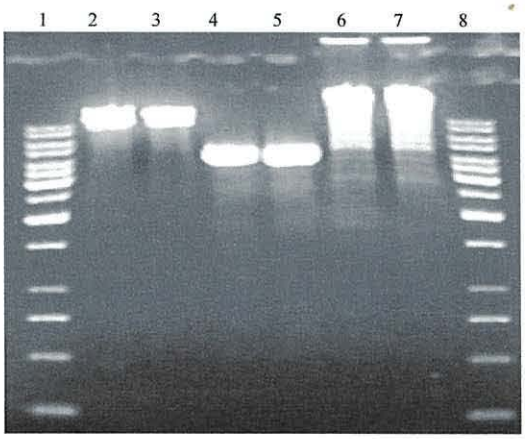


Fig.5.3.2. Standardized 2.5 μg donor DNA used in an average transformation assay. **Lanes 1** and **8**. Fermentas 1 kb Molecular weight marker. **Lanes 2** and **3**. pAD3-1A digested with *Sac*I. **Lanes 4** and **5**. pAD3-12A digested with *Sal*I. **Lanes 6** and **7**. AD3-1 genomic DNA digested with *Sal*I.

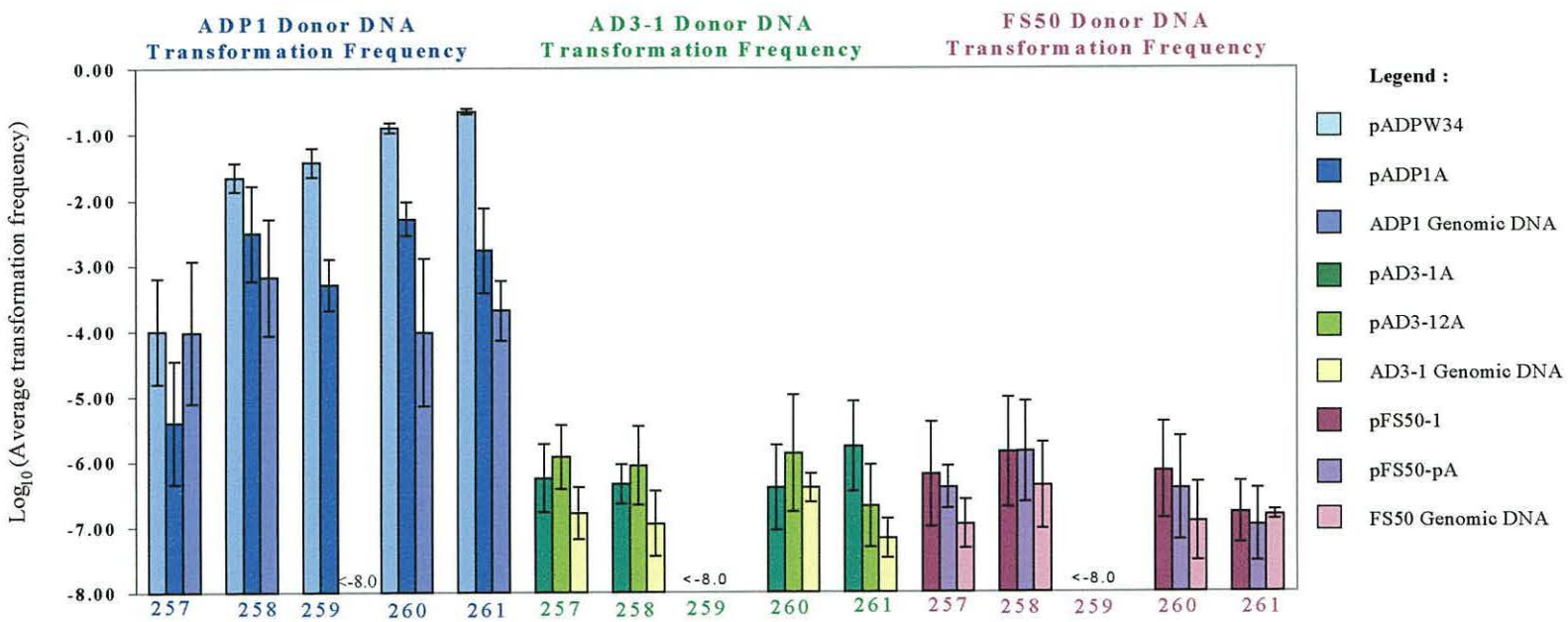


Fig.5.3.3. Log_{10} (average transformation frequencies) of Δ*salA* ADP1 mutants ADPW257 - ADPW261 using wild type ADP1 DNA and heterologous DNA from *Acinetobacter* sp. strains AD3-1 and FS50 respectively. Mean of three replicates. The data are depicted logarithmically to enable results from three separate transformation experiments to be visualized simultaneously. Standard deviations are the mean of three replicate experiments represented as the percentage deviation of the log_{10} mean values.

5.3.3 Homologous donor DNA transformation frequency

The results of mean transformation frequencies using wild type (ADP1) homologous donor DNA and the heterologous donor DNA from *Acinetobacter* sp. strains AD3-1 and FS50 are depicted in Fig.5.3.3 (-A, -B, -C) and listed in Tables 5.3.1 - 5.3.4. The frequency of transformation differed considerably in profile between homologous and heterologous donor DNA. The transformation frequency using heterologous donor DNA (approx. 10^{-7}) decreased by several magnitudes compared with that achieved using homologous DNA.

5.3.3.1 pADPW34 donor DNA transformations

When pADPW34 was used as the donor DNA an increasing gradient of transformation frequency was observed in strains ADPW257-261 (Fig.5.3.3A). This gradient correlates with increased two-sided homology present between donor DNA and the positions of the chromosomal *salA* deletions (Table 5.3.1). ADPW261 shows the greatest transformation frequency with the pADPW34 donor [2.24×10^{-1} ($\pm$ 0.16)]. The mutation in this strain has the greatest extent of two-sided flanking homology with pADPW34. For strains ADPW260, -259, and -258 the average transformation frequency decreases by approx. 2-fold, 6-fold and 10-fold respectively, corresponding with a decrease in flanking homology. Strain ADPW257 has 561 bp upstream and 3547 bp downstream homology with the pADPW34 donor and has the lowest transformation frequency. This is a 2000-fold decrease compared with the transformation frequency of strain ADPW261. These results indicate that with the pADPW34 donor the extent of homology bordering the mutation site between the recipient and transforming DNA has a direct effect on the frequency of transformation.

pADPW34	5' position of the 4bp deletion [a]	Upstream homology (bp)	Downstream homology (bp)	Transformation frequency [b]	Log_{10} [c]	Fold reduction transformation frequency [d]
ADPW257	121-124	561	3547	$1.0 (\pm 0.2) \times 10^{-4}$	-4.00	4.5×10^{-4}
ADPW258	332-335	772	3336	$2.2 (\pm 0.3) \times 10^{-2}$	-1.65	0.1
ADPW259	618-621	1058	3050	$3.7 (\pm 0.05) \times 10^{-2}$	-1.43	0.17
ADPW260	874-877	1314	2794	$1.25 (\pm 0.11) \times 10^{-1}$	-0.90	0.6
ADPW261	1132-1135	1566	2536	$2.24 (\pm 0.16) \times 10^{-1}$	-0.65	1

Table 5.3.1. Extent of the two-sided homology flanking the position of the mutation in strains ADPW257-261 and the relevant transformation frequency detected using pADPW34 (2.5 μg) as the homologous donor DNA.

[a] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[b] Mean transformation frequency (± Standard deviation) (n=3).

[c] Log_{10} mean transformation frequency.

[d] Fold reduction in pADPW34 transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW261.

5.3.3.2 pADP1A donor DNA transformations

When pADP1A was used as the homologous donor DNA, a pattern of transformation frequency was observed (Fig.5.3.3A) which, with the exception of strain ADPW259, again appeared to correlate with the smaller extent of the two-sided homology of the sequences flanking the mutation (Table 5.3.2). Strains ADPW258 and ADPW260 (Fig.5.3.4) have almost identical sizes of flanking DNA sequence bordering the mutations in a symmetrical orientation, and show similar transformation levels within 2-fold of each other (Table 5.3.2), 7- and 24-fold lower, respectively than those observed with the pADPW34 donor DNA (Table 5.3.4).

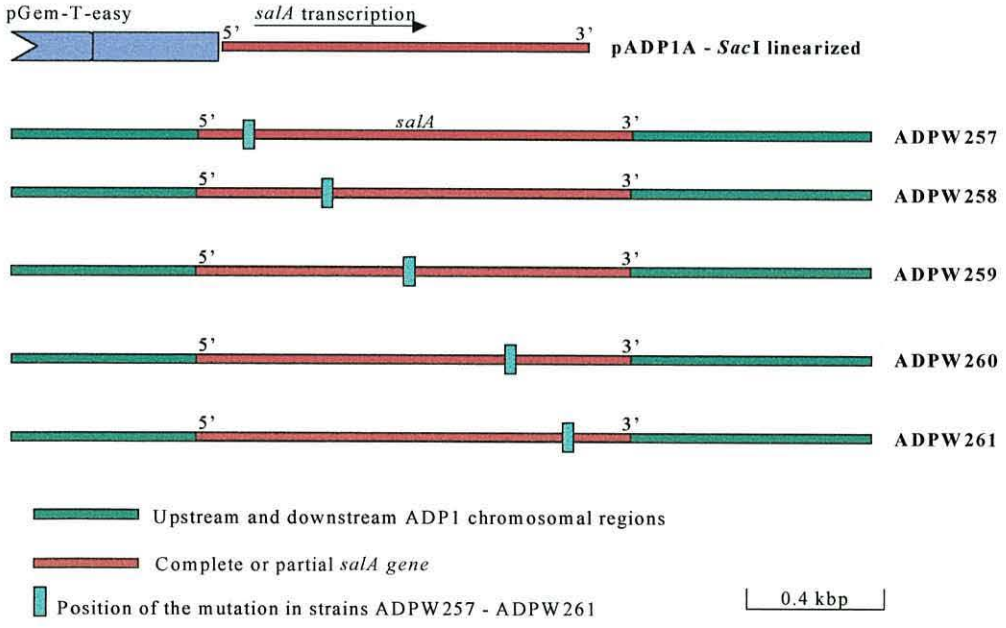


Fig.5.3.4. The extent of the homologous *salA* region contained in the pADP1A donor DNA aligned with the chromosomal Δ*salA* of the mutated strains ADPW257-ADPW261.

pADP1A	5' position of the 4bp deletion [(a)]	Upstream homology (bp)	Downstream homology (bp)	Transformation frequency [(b)]	Log_{10} [(c)]	Fold reduction transformation frequency [(d)]
ADPW257	121-124	95	1087	$4.0 (\pm 0.7) \times 10^{-6}$	-5.40	7.6×10^{-4}
ADPW258	332-335	305	876	$3.1 (\pm 1.4) \times 10^{-3}$	-2.51	0.6
ADPW259	618-621	591	590	$5.1 (\pm 0.6) \times 10^{-4}$	-3.28	0.1
ADPW260	874-877	847	334	$5.25 (\pm 0.6) \times 10^{-3}$	-2.28	1
ADPW261	1132-1135	1105	76	$1.7 (\pm 0.4) \times 10^{-3}$	-2.77	0.33

Table 5.3.2. Extent of the two-sided homology flanking the position of the mutation in strains ADPW257-261 and the relevant transformation frequency detected using pADP1A (2.5 μg) as the homologous donor DNA.

[(a)] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

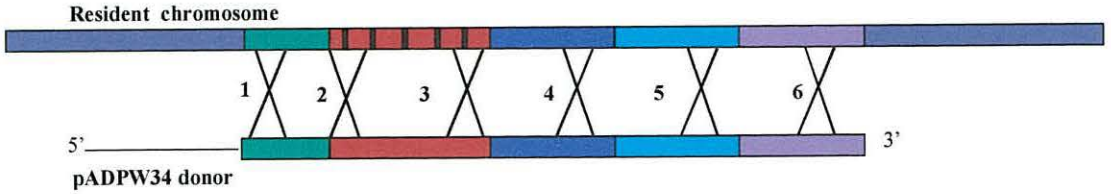
[(b)] Mean transformation frequency (± Standard deviation) (n=3).

[(c)] Log_{10} mean transformation frequency.

[(d)] Fold reduction in pADP1A transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW260.

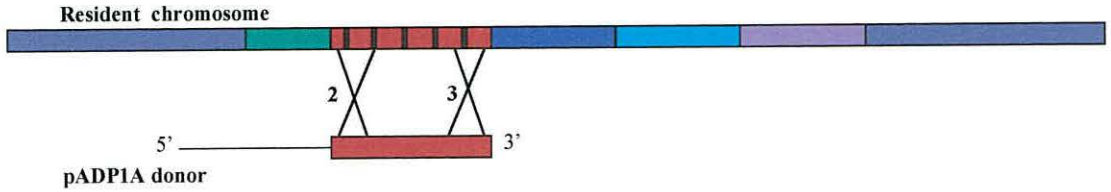
Given the central position of the Δ*salA* mutation within strain ADPW259 (Table 5.3.2 and Fig.5.3.4) and the equal length of the flanking sequences provided by the pADP1A donor DNA surrounding the mutation, the frequency of transformation was expected to be higher. The transformation frequency for strain ADPW259 with pADP1A donor DNA was 10-fold lower than that of strain ADPW260 and 6-fold lower than strain ADPW258. When strain ADPW259 was transformed using homologous genomic donor DNA the transformation frequency fell below the level of detection of this assay. This and the lower than expected results obtained for transformations with the pADP1A donor DNA indicated that some other factor was affecting transformation frequencies of the ADPW259 mutation. This factor has not been determined but could possibly be connected to DNA topology or the presence of a coldspot for recombination.

In ADPW257 transformations using pADP1A as the donor DNA, a flanking homology borders the mutation of 95 bp upstream and 1087 bp downstream (Table 5.3.2 and Fig.5.3.4). This situation is reversed with ADPW261 where the mutation has bordering homologies of 1105 bp upstream and 76 bp downstream with the pADP1A donor. Despite this similarity in the surrounding sequence, the transformation frequency of ADPW257 is approx. 400-fold lower than that of ADPW261. Cleavage of pADP1A, which contains the truncated wild type *salA* gene, linearizes the ds donor DNA so that the plasmid fraction lies upstream of the donor DNA region that repairs the ADPW257 deletion. Therefore, there is 88 bp of upstream homology adjacent to 3.0 kbp of the heterologous plasmid DNA. The large region of heterologous plasmid DNA immediately upstream of the 5' end of the *salA* insert in pADP1A may negatively affect the ability of the DNA to undergo homologous recombination at the ADPW257 deletion site. Integration of the vector sequence by homology-



Possible homologous recombination events between the homologous pADPW34 donor and recipient DNA.

E.g. cross over can occur at any point: 1 can combine with 2, 3, 4, 5 or 6 etc replacing the resident DNA with the integrating single stranded DNA.



Possible homologous recombination events between the homologous pADP1A donor and recipient DNA.

E.g. cross-over can occur only within the target gene region.

Fig.5.3.5. Possible homologous recombination and reciprocal cross over events between pADPW34 and pADP1A plasmid donors and the recipient chromosome. Homologous recombination cross over events for pADP1A DNA are limited to within the *salA* allele resulting in lower transformation frequencies. HFIR events may occur integrating the pADP1A vector region upstream of *salA*.

facilitated illegitimate recombination (HFIR) with the pADP1A *salA* allele acting as an anchor region could occur, disrupting the resident upstream *salA* region and producing a 5' truncated *salA* gene that could not be selected for under the conditions of this assay (Fig.5.3.5).

ADPW261 transformations are not affected by the presence of the plasmid DNA as it is located at the opposite end of the donor DNA (Fig.5.3.4). The low transformation frequencies of ADPW261 with pADP1A donor DNA (Table 5.3.2) may be accounted for by the small downstream region of homology (88 bp) present between the incoming pADP1A donor DNA and the resident chromosome. Repair of the ADPW261 mutation with the pADP1A donor occurs with comparatively one-sided homology resulting in lower levels of transformation (Fig.5.3.4).

When strain ADPW257 was transformed with pADPW34 561 bp of homology exists upstream of the mutation and a transformation frequency of 1.0×10^{-4} was observed (Table 5.3.1). Approximately the same amount of upstream homology is present when strain ADPW259 is transformed with the pADP1A donor DNA. The transformation frequency of this strain with pADP1A donor DNA is 5 times higher than strain ADPW257 but within the same order of magnitude. The possibility that the transformation frequencies of these strains are affected by other factors such as recombinational coldspots or the effects of repetitive sequences that could cause DNA slippage or misalignment during homologous recombination cannot be ruled out. The DNA sequence TTGCAG occurs four times within 165, 60, 75 and 84 bp of the position of the ADPW257 deletion.

ADP1 genomic DNA	5' position of the 4bp deletion [a]	Transformation frequency [b]	Log_{10} [c]	Fold reduction transformation frequency [d]
ADPW257	121-124	9.5 (± 2.6) x 10^{-5}	-4.02	0.14
ADPW258	332-335	6.7 (± 1.9) x 10^{-4}	-3.17	1
ADPW259	618-621	<1.0 x 10^{-8}	-8.0	>1.5 x 10^{-5}
ADPW260	874-877	9.5 (± 2.7) x 10^{-5}	-4.02	0.14
ADPW261	1132-1135	2.1 (± 0.26) x 10^{-4}	-3.68	0.3

Table 5.3.3. Transformation frequencies observed for strains ADPW257 - ADPW261 using ADP1 genomic DNA (2.5 μg) as the donor DNA.

[a] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[b] Mean transformation frequency (± Standard deviation) (n=3).

[c] Log_{10} mean transformation frequency.

[d] Fold reduction in ADP1 genomic DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW258.

5.3.3.3 Wild type genomic donor DNA transformations

ADP1 wild type transformations of ADPW257-261 with the exception of ADPW259, gave similar levels of transformation frequencies within 3- to 7-fold of each other when ADP1 genomic DNA was provided as the donor DNA (Fig.5.3.3A; Table 5.3.3). The observed values fall within 10-fold of chromosomal donor DNA transformation frequencies observed by Palmen *et al.* (1993). For ADPW259 the transformation frequency using wild type genomic DNA fell below the level of detection ($< 1.0 \times 10^{-8}$). The highest transformation frequencies observed for ADPW260 and -261 with pADPW34 as donor DNA (Table 5.3.1) were reduced a 1000-fold when genomic DNA was used as the DNA source (Fig.5.3.3A and Table 5.3.4). The differences between the transformation frequencies for the three different DNA sources are described in Table 5.3.4. The dramatic decrease in transformation frequency noted for strain ADPW257 with the two plasmid DNA sources as donors was not apparent when genomic DNA was used as the transforming DNA.

The ADPW259 transformation assays that produced very few transformants, were repeated with the same culture and also with a fresh strain plated from the original -70°C stock culture using both homologous and heterologous donor DNA. Colony PCR was performed on this strain to isolate the *salA* mutated gene and to check for the presence of the introduced *Bam*HI site indicating the strain carried the Δ*salA* gene. These steps were performed to ensure that the deletion mutation was present and that competency for natural transformation remained viable in this strain and that no manual errors were incorporated during the transformation assay procedure. These assay repetitions produced similar transformation frequency results.

Fold reduction in the transformation frequency of the Δ*salA* ADP1 mutants between the different sources of wild type donor DNA

	pADPW34 / pADP1A	pADPW34 / ADP1 genomic DNA	pADP1A / ADP1 genomic DNA
ADPW257	0.04	1	24
ADPW258	0.14	0.03	0.22
ADPW259	0.014	$>2.0 \times 10^{-7}$	$>2.0 \times 10^{-5}$
ADPW260	0.04	7.6×10^{-4}	0.02
ADPW261	7.6×10^{-3}	1.0×10^{-3}	0.12

Table 5.3.4. Fold reduction in transformation frequencies between the different sources of homologous ADP1 donor DNA for strains ADPW257 - ADPW261.

Comparison of the transformation frequency of the ΔsalA ADP1 mutants using homologous *A. baylyi* strain ADP1 and heterologous *Acinetobacter* sp. strains AD3-1 and FS50 derived donor DNA.

	pADPW34 / pAD3-1A	pADPW34 / pFS50-1	pADP1A / pAD3-12A	pADP1A / pFS50-pA	ADP1 genomic DNA / AD3-1 genomic DNA	ADP1 genomic DNA / FS50 genomic DNA
ADPW257	180	160	3	10	560	860
ADPW258	5.0×10^4	1.6×10^4	3.5×10^3	2.0×10^3	6.0×10^3	1.6×10^3
ADPW259	$>3.7 \times 10^6$	$> 3.7 \times 10^6$	$>5.1 \times 10^4$	$> 5.1 \times 10^4$	-	-
ADPW260	3.1×10^5	1.7×10^5	4.0×10^3	1.3×10^4	240	430
ADPW261	1.3×10^5	1.3×10^7	8.0×10^3	1.7×10^4	3.1×10^3	2.0×10^3

Table 5.3.5. Comparisons of the transformation frequencies of the Δ*salA* mutants using homologous *A. baylyi* strain ADP1 donor DNA and *Acinetobacter* sp. strains AD3-1 and FS50 derived donor DNA.

5.3.4 Heterologous donor DNA transformation frequency

The transformation frequencies of the five ADPW257-261 mutants with AD3-1 plasmid or genomic donor DNA are depicted in Fig.5.3.3 B and Tables 5.3.6 - 5.3.8. The FS50 donor DNA results (Tables 5.3.9 - 5.3.11) are included in Appendix A. The transformations using AD3-1 and FS50 heterologous donor DNA gave overall very similar levels of mean transformation frequencies in comparison with the wild type donor (Table 5.3.5) and the standard deviation showed a greater variance than that of the transformations by wild type DNA. From the sequencing and alignment results the two heterologous DNA sources were shown to be very closely related (Chapter 4 - 4.3.4.2 and Table 4.2). As a result of the DNA identity and the overall similarity in the transformation frequencies of the Δ*salA* strains using the heterologous *Acinetobacter* spp. strains AD3-1 and FS50 (Table 5.3.5), only the transformation frequency results with strain AD3-1 donor DNA will be discussed.

5.3.4.1 pAD3-1A donor DNA transformations

Transformations using the putative salicylate operon donor pAD3-1A displayed no gradient of transformation frequency relative to an increase in the two-sided homology, as observed for the wild type transformations with the homologous salicylate operon carried in pADPW34. Transformation frequencies with pAD3-1A (approx. 10^{-7}) were reduced by 180-fold to 10^{5}-fold compared to the corresponding wild type plasmid DNA. In strain ADPW259 transformants were not detected and the value given is that classed as being below the level of detection (Table 5.3.5).

pAD3-1A	5' position of the 4bp deletion [(a)]	Transformation frequency [(b)]	Log_{10} [(c)]	Fold reduction transformation frequency[(d)]
ADPW257	121-124	5.6 (± 1.4) x 10^{-7}	-6.25	0.3
ADPW258	332-335	4.6 (± 0.2) x 10^{-7}	-6.33	0.3
ADPW259	618-621	< 1.0 x 10^{-8}	< -8.0	>0.006
ADPW260	874-877	4.0 (± 0.4) x 10^{-7}	-6.39	0.23
ADPW261	1132-1135	1.7 (± 0.2) x 10^{-6}	-5.77	1

Table 5.3.6. Transformation frequencies of strains ADPW257-261, using plasmid donor DNA pAD3-1A isolated from *Acinetobacter* sp. strain AD3-1.

[(a)] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[(b)] Mean transformation frequency (± Standard deviation) (n=3).

[(c)] Log_{10} mean transformation frequency.

[(d)] Fold reduction in pAD3-1A donor DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW261.

pAD3-12A	5' position of the 4bp deletion [(a)]	Transformation frequency [(b)]	Log_{10} [(c)]	Fold reduction transformation frequency[(d)]
ADPW257	121-124	1.2 (± 0.1) x 10^{-6}	-5.92	0.9
ADPW258	332-335	9.0 (± 0.9) x 10^{-7}	-6.06	0.65
ADPW259	618-621	< 1.0 x 10^{-8}	<-8.0	> 7.5 x 10^{-3}
ADPW260	874-877	1.3 (± 0.2) x 10^{-6}	-5.87	1
ADPW261	1132-1135	2.1 (± 0.2) x 10^{-7}	-6.67	0.16

Table 5.3.7. Transformation frequencies of strains ADPW257-261 using donor DNA pAD3-12A isolated from *Acinetobacter* sp. strain AD3-1.

[(a)] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[(b)] Mean transformation frequency (± Standard deviation) (n=3).

[(c)] Log_{10} mean transformation frequency.

[(d)] Fold reduction in pAD3-12A donor DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW260.

AD3-1 genomic DNA	5' position of the 4bp deletion [(a)]	Transformation frequency [(b)]	Log_{10} [(c)]	Fold reduction transformation frequency [(d)]
ADPW257	121-124	1.7 (± 0.1) x 10^{-7}	-6.78	0.4
ADPW258	332-335	1.1 (± 0.08) x 10^{-7}	-6.95	0.3
ADPW259	618-621	< 1.0 x 10^{-8}	<-8.0	> 0.025
ADPW260	874-877	4.0 (± 0.14) x 10^{-7}	-6.39	1
ADPW261	1132-1135	7.0 (± 0.3) x 10^{-8}	-7.17	0.17

Table 5.3.8. Transformation frequencies of strains ADPW257-261 using genomic donor DNA isolated from *Acinetobacter* sp. strain AD3-1.

[(a)] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[(b)] Mean transformation frequency (± Standard deviation) (n=3).

[(c)] Log_{10} mean transformation frequency.

[(d)] Fold reduction in AD3-1 genomic donor DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW260.

5.3.4.2 pAD3-12A and AD3-1 genomic donor DNA transformations

When pAD3-12A or AD3-1 genomic DNA was used as the donor transformation frequencies similar to those of pAD3-1A were observed, between 10^{-6} and 10^{-8} (Fig.5.3.3B and Tables 5.3.6 - 5.3.8). ADPW259 produced transformants at a frequency below 10^{-8} for all types of the heterologous donor DNA used. Between strains ADPW257, -258, -260 and -261, the transformation frequencies varied between 3-4-fold for pAD3-1A, 1-6 -fold for pAD3-12A and 2-6-fold using AD3-1 genomic DNA.

This general frequency of transformation indicates that in the case of heterologous donor DNA, the replacement of the mutation was not affected by its position within the *salA* gene or the length of transforming DNA. When digested with *Sac*I the plasmid fraction in pAD3-12A lies in a 5' orientation relative to the direction of transcription of the cloned *salA*. This fact, however, appears to have no effect on the transformation frequency of ADPW257. The similarity of the transformation frequencies with the three types of heterologous donor DNA suggests homologous recombination occurred solely within the *salA* gene region, possibly through the formation of micro-homologies within small regions of the gene possessing identical sequences. The AD3-1 genomic DNA transformation frequency was slightly lower than the high copy number plasmid DNA. The possibility of larger stretches of this donor DNA being incorporated outside of the *salA* coding sequence also through the formation of microhomologies cannot be ruled out, but were not identified in transformant colony PCR studies.

5.3.4.3 Transformant colony PCR

For each Δ*salA* mutant strain transformed with the pAD3-1A donor DNA, five transformant CFUs were selected at random for colony PCR. The *salA* gene was isolated using primers P5F and P6R (Chapter 3 - Table 3.2.4) and digested with *Bam*HI and *Cla*I to investigate the extent of donor DNA integration and recombination. Colony PCR performed with strain AD3-1 using these primers with either plasmid or genomic DNA template did not produce a *salA* PCR fragment, in contrast with the wild type ADP1 strain, in which the *salA* PCR fragment was reliably reproduced. All of the transformant colonies selected (Fig.5.3.6) produced a PCR fragment of approx. 1.2 kbp. Proof of the integrity of the PCR reaction was provided via the use of a no template negative control (Fig.5.3.6 Lane 2).

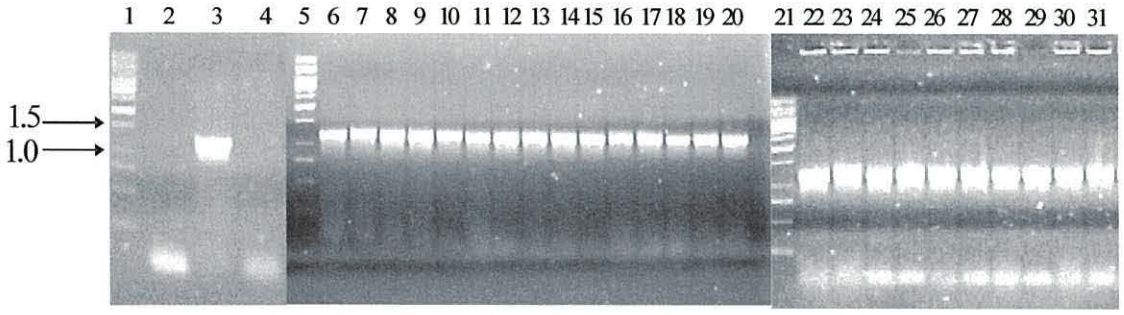


Fig.5.3.6. Colony PCR. **Lane 1, 5, 21**. 1 kb DNA marker. **Lane 2**. No template. **Lane 3**. ADP1. **Lane 4**. AD3-1. **Lanes 6-10**. ADPW257 pAD3-1A transformants. **Lanes 11-15**. ADPW258 pAD3-1A transformants. **Lanes 16-20**. ADPW259 pAD3-1A transformants. **Lanes 22-26**. ADPW260 pAD3-1A transformants. **Lanes 27-31**. ADPW261 pAD3-1A transformants.

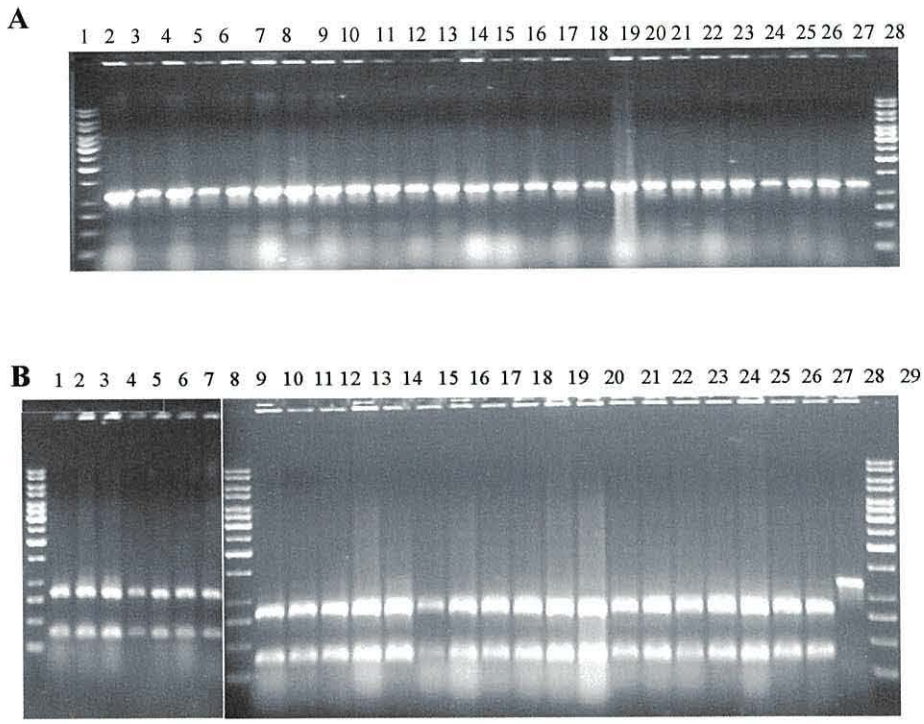


Fig.5.3.6.1 A. pAD3-1A and ADP1 colony PCR *Bam*HI digests. **Lanes 1, 28**. 1 kb DNA ladder. **Lane 2-6**. ADPW257 colonies, 1-5. **Lanes 7-11**. ADPW258 colonies, 1-5. **Lanes 12-16**. ADPW259 colonies, 1-5. **Lanes 17-21**. ADPW260 colonies, 1-5. **Lanes 22-26**. ADPW261 colonies, 1-5. **Lane 27**. ADP1 colony PCR.

Fig.5.3.6.1 B. pAD3-1A colony PCR *Cla*I digests. **Lanes 1, 9 and 29**. 1 kb DNA ladder. **Lanes 2-6**. ADPW257 colonies, 1-5. **Lanes 7-8 and 10-12**. ADPW258 colonies, 1-5. **Lanes 13-17**. ADPW259 colonies, 1-5. **Lanes 18-22**. ADPW260 colonies, 1-5. **Lanes 23-27**. ADPW261 colonies, 1-5. **Lane 28**. Undigested control.

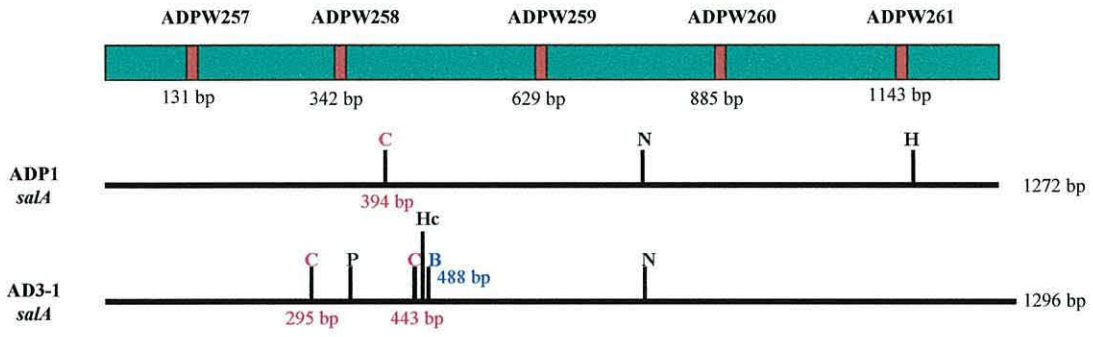


Fig.5.3.7. Positions of common restriction sites within the ADP1 and AD3-1 *salA* alleles. Above the *salA* restriction maps is a representation of the locations of each 4 bp deletions relative to the restriction sites. Base-pair numbers indicate the exact location of the deletions, restriction sites and length of the two *salA*s. C – *Cla*I, B – *Bam*HI, P – *Pvu*I, Hc – *Hin*cII, N – *Nco*I, H – *Hin*dIII.

Digestion of the transformant colony PCR products with *Bam*HI and *Cla*I enzymes yielded no restriction pattern that indicated the acquisition of restriction enzyme sites from the pAD3-1A donor DNA within the *salA* sequence (Fig.5.3.6.1A and B). Digestion with *Bam*HI provided evidence that the introduced site within the Δ*salA* genes had been replaced and that the unique *Bam*HI site present within the AD3-1 *salA* had not become incorporated within the ADP1 *salA* allele during recombination. The *Cla*I digests produced a restriction pattern identical to the wild type ADP1 *salA*, however if only the *Cla*I site at 443 bp was acquired the restriction fragment length generated would differ by only 50 bp and may have gone undetected (Fig.5.3.7). The incorporation of both *Cla*I sites would have produced fragment lengths of ~800

bp, ~150 bp and ~290 bp that would have been distinguishable from the wild type pattern, unless recombination produced *Cla*I site(s) protected by Dam methylation. Incorporation of the *Bam*HI site present in the donor DNA can be more clearly stated as to have not occurred. As the restriction sites unique to the donor DNA occur mainly within a ~200 bp region (Fig.5.3.7) and none of these sites appear to have been incorporated, the indication is that recombination occurred only between small regions of the donor and recipient DNA

5.4. Discussion

5.4.1 Effects of flanking homology on the transformation frequency

The transformation frequencies observed with homologous donor DNA indicate that the transformation frequency is directly related to the extent of two-sided homology between the donor DNA and the resident chromosome surrounding the mutation.

The maximum transformation frequency using pADPW34 donor DNA was observed for ADPW261 (Fig.5.3.3A). This strain had the greatest amount of two-sided flanking homology with the donor DNA of 1566 bp upstream of the deletion and 2536 bp downstream (Table 5.3.1). As the two-sided homology between the recipient strains and the donor DNA decreased so did the frequency of transformation as evidenced by a decrease in of 2-, 6-, 10- and 2000-fold for strains ADPW260 - 257 respectively (Table 5.3.1). This phenomenon has been observed previously for natural transformations of *Acinetobacter* sp. strain BD413 (Palmen and Hellingwerf, 1997). In this study, the effect of the length of the transforming DNA on the transformation frequency was investigated. In transformations repairing a frameshift mutation within a chromosomally inserted kanamycin resistance gene (*aphA-3*), the highest transformation frequencies observed were with DNA possessing flanking regions of 1750 bp upstream and 2071 bp downstream from the mutation. These workers demonstrated that the transformation frequency decreases upon a decrease in the fragment length of the transforming DNA. Comparisons of the transformation frequencies of the five Δ*salA* mutations

using homologous donor DNA (Fig.5.3.3A, Tables 5.3.1 and 5.3.2) have shown that the position of the mutation within the gene and the extent of flanking homology with the donor DNA has a direct affect on the transformation process.

5.4.2 Effects of gene copy number per μg of DNA

The differences in the wild type transformation frequencies observed with multicopy plasmid DNA and genomic DNA (Fig.5.3.3.A, Tables 5.3.1 - 5.3.4) may be explained by the effects of gene copy number. The target gene within the genomic DNA represents ~0.04% of the total chromosome. Within pADPW34 and pADP1A the target gene represents approx. 17% and 30% respectively of the total DNA available for repair of the mutation. Transformation with genomic DNA is enhanced by the larger extent of sequence homology that exists with the resident chromosome surrounding the mutation. Multicopy plasmid DNA has a greater transformation frequency due to the higher level of DNA that may be imported into the cell that is specific for the target gene. Chromosomal fragments that do not contain the target gene may saturate the DNA uptake system reducing the number of competent cells in the culture that take up the target allele. The fact that the transformation frequencies are so similar between the mutated strains when genomic DNA is used as the donor (Table 5.3.3) indicates a similar sized fragment is generated by the *Sal*I digestion that contains the region of the chromosome which repairs the mutation. Variation in the transformation frequencies between strains may be due to differences in the total number of cells within competent cultures. Alternatively, variations in frequency may arise due to the position of the random double-stranded break introduced in the donor DNA during DNA

binding at the cell surface prior to translocation into the cytoplasm. This also may affect transformation in terms of the length of the transforming DNA, due to nuclease degradation of the incoming DNA upon entering the cytoplasm prior to processing of the DNA for homology pairing and integration by the RecA protein and/or the recombination machinery.

5.4.3 Effects of pADP1A donor DNA on the transformation frequency of ADPW257 and ADPW261

When pADP1A was used as the donor DNA the transformation frequency of strains ADPW261, -260 and -258 differed by only a -2 to -3 -fold decrease in frequency (Table 5.3.2). ADPW259 showed a 10-fold decrease in transformation frequency. Transformations of strain ADPW257 however showed a dramatic decrease of approx. 1000-fold. When ADP1 genomic DNA was used as the donor such a dramatic decrease in frequency was not noted for the ADPW257 mutation (Fig.5.3.3A and Table 5.3.3). This indicates that transformations using the plasmid DNA donors in this strain were negatively affected by the presence of heterologous vector sequence upstream of the *salA* region. Homology-facilitated illegitimate recombination (HFIR) has been demonstrated in *Acinetobacter* (de Vries and Wackernagel, 2002) and in *Pseudomonas stutzeri* (Meier and Wackernagel, 2003). These workers used a system that demonstrated the integration of foreign DNA facilitated by one-sided homologous regions which served as anchor sites, allowing the formation of microhomologies (junction sites) between the heterologous DNA and the resident chromosome. Homologous regions down to 183 bp and 311 bp in size were demonstrated to serve as anchors for foreign DNA integration in *Acinetobacter* and *P. stutzeri* respectively. The microhomologies formed were

between 3-8 bp in length and facilitated the integration of large stretches of completely heterologous DNA replacing the native genome. The *salA* insert in pADP1A could possibly have served as a homologous anchor region facilitating the integration of the 5' oriented vector DNA. As pADP1A contains a truncated *salA* gene incorporation of the 5' *salA* region into the resident chromosome would create a non-viable salicylate hydroxylase gene, possibly replacing the natural promoter region required for LysR regulated *salA* expression. Under the conditions of this assay such a transformation event could not be identified, resulting in the decreased transformation frequencies observed for this strain with the plasmid donors.

With ADPW261 a similar amount of homology exists with the pADP1A donor to that of ADPW257 but in the opposite orientation (Table 5.3.2 and Fig.5.3.4). The transformation frequency of this strain was 425-fold greater than that of ADPW257. Relatively small fragments with one-sided homology have been observed to be capable of repairing small mutations (Palmen and Hellingwerf, 1997). Using the repair of an inactivated kanamycin gene as a marker they observed significant transformation frequencies (approx 10^{-6}) when using a donor DNA fragment of 616 bp that contained only a significant bordering fragment on one side of the mutation. Transformations with this fragment indicated that after uptake and partial degradation, the 616 bp donor DNA was completely integrated into the chromosome replacing the mutation. pADP1A which has approx. 1 kb of one-sided homology bordering the ADPW257 mutation displayed a transformation frequency similar to that observed by Palmen and Hellingwerf (1997) using donor DNA of 616 bp in length. A possible explanation of the discrepancy in frequencies between ADPW261 and ADPW257 strains is that in strain ADPW257 the frequency of mutational repair with one-sided homology of 1 kb is decreased due to the presence of the

heterologous vector sequence and possible HFIR events in the 5' *salA* upstream region. In a study performed by de Vries and Wackernagel (2002) the integration of donor DNA was shown to occur in either a 5' or a 3' orientation but with an overall preference for 5' orientation (ratio 80:20). A preference for 5' integration would also result in a larger number of transformants at the 3' end of the gene.

5.4.4 Strain ADPW259

Transformation experiments with strain ADPW259 showed significantly lower frequencies of transformation with homologous and heterologous donor DNA types than the four other strains. Considering the central position of this mutation (Fig.5.3.4) and the even amount of two-sided homology with the homologous DNA the transformation frequencies of this strain were expected to be higher. In addition, somewhat unexpectedly, heterologous donor DNA transformation frequencies fell below the level of detection. As the neighbouring mutations produced quantitative transformant numbers when transformed with heterologous donor DNA the data for strain ADPW259 suggests the presence of a recombinational cold spot.

Recombinational hot and cold spots are widely accepted as occurring within both prokaryotic and eukaryotic genomes. Over-represented GC rich nucleotide Chi sequences between 5 and 8 nucleotides in length have been identified as hotspots for recombination within bacterial genomes such as *E. coli* (Bianco and Kowalczykowski, 1997), *Lactococcis lactis* (Biswas *et al.*, 1995), *H. influenzae* (Sourice *et al.*, 1998), and *B. subtilis* (Chedin *et al.*, 1998). These repetitive sequences interact with the endonucleolytic-helicase protein complex RecBCD or analogue and cause a conformational change within the RecBCD

protein, "switching off" the nucleolytic activity whilst retaining helicase function. Therefore the Chi sequence serves to protect DNA during the process of DNA repair and replication from endonucleolytic digestion and promotes homologous recombination from the single-stranded end generated by continued helicase activity. Perhaps coldspots arise at positions distal to the location of a Chi sequence. Although no evidence has been presented to date it is interesting to speculate on the presence of a Chi-like sequence within the ADP1 genome.

5.4.5 Recombinational coldspots in yeast

In yeast, gene disruption cassettes have been effectively used in gene targeting studies. However, observations have been made that certain loci are very difficult to disrupt due to the low efficiency of integration of the disruption cassette. A recent report by Gjuracic *et al.* (2004) used the haploid yeast strain S288C to study the efficiency of gene targeting within different regions of genes, i.e. promoter, ORF and terminator regions. Transformation experiments demonstrated that the 5' and 3' flanking regions were preferred as sites for plasmid integration over the corresponding ORF regions. When plasmids bearing combinations of two or three regions were linearized to target them to a specific site of integration three out of four ORFs were found to be less preferred as sites for plasmid integration. Their results suggest that coding sequences may be considered as "cold spots" for plasmid integration in yeast. The presence of such cold-spots in bacterial species has not been especially detailed but may possibly mirror the situation observed above.

5.4.6 Heterologous donor DNA transformation frequencies

The alignment of the AD3-1 derived heterologous *salA* donor with the wild type DNA sequence (Chapter 4 - 4.3.4.3) revealed that regions of homology were distributed throughout the whole gene sequence. The largest single block of homology was 17 bp in length but the most frequently occurring regions of homology were much shorter. PCR digests of the randomly selected colonies did not detect the acquisition of restriction sites present in the donor DNA within the transformant sequence. Primers used to amplify the *salA* sequence from the transformant colonies produced no PCR products when performed with the AD3-1 DNA sources. The overall similarity in the transformation frequencies observed between the three heterologous DNA and the PCR results suggest that recombination occurred solely within the *salA* region. Downstream of the pAD3-1A *salA* homologue between *salA* and *salR* lies an ORF encoding a permease gene (*salK*) that is absent from the wild type DNA. This creates decreased downstream homology between the two sequences that may negatively affect the transformation frequency. In addition little conservation with ADP1 is observed for the region upstream of *salA* (Fig.6.3.2.6).

Heterologous donor DNA of 60 % identity at the nucleotide level gave transformation frequencies ranging from 180 - 10^5 -fold lower than wild type DNA. Under laboratory conditions wild type transformation frequencies were affected by the length of transforming DNA, the extent of two-sided homology and possibly regions of the chromosome permissive for transformation.

Chapter 6

6.1 Introduction

Salicylate is a common compound in the environment and many bacteria are capable of degrading it (Chapter 1 – 1.1.8.1 – 1.1.8.2). In *Pseudomonas* spp. salicylate hydroxylase is an integral enzyme required in the pathway of naphthalene degradation and the gene encoding this function (*nahG*) is clustered with the other genes required for naphthalene degradation (Chapter 1 - 1.1.7). In *A. baylyi* strain ADP1 the salicylate genes are clustered together as two transcriptional units comprised of *salA* and *salR* (encoding salicylate hydroxylase and the LysR-type transcriptional regulator, respectively) and *salE* and *salD* (encoding salicylate esterase and a putative porin protein, respectively). These two salicylate transcriptional units are clustered with the *ben-cat* supraoperonic cluster of catabolic genes (Chapter 1 - 1.2.2.2 – 1.2.3.1). In this study a putative salicylate operon was isolated as a 5.78 kbp DNA fragment from the genome of *Acinetobacter* sp. strain AD3-1 which has the gene order *salAKR*. Blastx protein similarity searches indicated *salA* encodes a salicylate hydroxylase, *salK* encodes a transport protein of the Major Super Facilitator family (MFS) and *salR* encodes a LysR type transcriptional regulator.

This chapter describes investigations into the *salA*, -*K* and –*R* gene functions by gene knockout analysis and considers the mechanism of regulation of the putative *sal* operon. The *Acinetobacter* sp. strain AD3-1 *salA* gene was cloned, over-expressed under optimized conditions and purified. A single protein band of 48 kDa was observed in SDS-PAGE analysis. Investigations into the kinetics

and the substrate specificity of the purified enzyme and cell extracts were carried out.

Chapter 6 preface

E. coli strains carrying pAD3-1A gene deletion constructs of *salA, -K,* and *-R* were created and analysed by measuring salicylate 1-monooxygenase catalysed NADH + H^+ oxidation rates as compared with wild type *E. coli* (pAD3-1A) oxidation rates. A pAD3-1A *salA-Bam*HI frame-shift deletion destroyed the relative specific activity of SalA, as did a *salR-Hin*dIII deletion. A 27 bp deletion introduced at the *Hin*cII site within *salR* reduced the relative specific activity of SalA by ~50%. A *salK* frame-shift mutation resulted in 60% loss of SalA relative specific activity.

Kinetic studies of the purified enzyme revealed a narrow pH optimum (pH 7-8), 1:1 stoichiometry between salicylate and NADH + H^+, and salicylate K_m and V_{max} values of 5.5μM and 30 nmol min^{-1} respectively. Hanes-Woolf analytical plots suggest an ordered compulsory-ternary complex enzyme mechanism for this enzyme. Substrate specificity studies using a range of substituted salicylate compounds revealed relative specific activities of ~80 % with 4- and 5-methylsalicylates, and ~50% with both 4-chlorosalicylate and 5-aminosalicylate when compared with the % relative specific activity of the enzyme with salicylate.

6.2 Materials and methods

Table 6.2.1 Plasmids used and constructed

Plasmid	Description	Source
pUC19	Ap^R vector used in cloning experiments	Promega
pUC19H$^-$	pUC19 with frameshift mutation deleting the *Hin*dIII restriction site from the MCS	This study
pUC19HS$^-$	pUC19H$^-$ with a frameshift mutation deleting the *Sal*I restriction site from the MCS	This study
pAD3-1A	5.7 kbp shotgun clone of chromosomal DNA isolated from *Acinetobacter* sp. Strain AD3-1 containing the putative AD3-1 salicylate operon	This study
pAD3-1A1	pAD3-1A with a *Sph*I deletion of 0.381 kbp creating a 5.4 kbp *Sph*I / *Sac*I insert in pUC19 removing 0.255 kbp of the pAD3-1A siderophore gene	This study
pAD3-1A2	*salA* knockout of pAD3-1A1 with a frameshift mutation introduced into the unique *Bam*HI site of *salA*	This study
pAD3-1A3	*salK* knockout of pAD3-1A1 with a frameshift mutation introduced into the unique *Sgf*I site of *salK*	This study **Continued overpage**

pAD3-1A4	3.6 kbp *Bam*HI - *Sac*I fragment of pAD3-1A containing part of *salA* and the entire *salK* and *salR* genes ligated into pUC19HS$^-$	This study
pAD3-1A5	pAD3-1A4 with a frameshift deletion introduced at the unique *Acc*I site in the centre of the putative differential expression loop	This study
pAD3-1A6	pAD3-1A5 with the 1.6 kb *Sph*I - *Bam*HI fragment of pAD3-1A1 reforming the complete pAD3-1A1 sequence plus the *Acc*I deletion	This study
pAD3-1A7	pAD3-1A4 with a 0.381 kbp *Hin*dIII deletion	This study
pAD3-1A8	pAD3-1A7 with the pAD3-1A *Xba*I - *Bam*HI fragment reforming the complete pAD3-1A sequence including a 0.381 kbp *Hin*dIII deletion	This study
pAD3-1A9	pAD3-1A4 with a blunt ended "chew back" *Hin*cII mutation removing the restriction site and creating a 27 bp deletion	This study
pAD3-1A10	pAD3-1A9 with the pAD3-1A *Xba*I - *Bam*HI fragment reforming the complete sequence including a 27 bp deletion in *salR*	This study
pET28a	His-tag expression vector Km^R	Novagen
p196EXA	pET28a with pAD3-1A *salA* PCR product inserted into the MCS as an *Eco*RI - *Hin*dIII fragment	This study

Abbreviations: MCS - multi-cloning site. Plasmids pAD3-1A - pAD3-1A10 were transformed into *E. coli* strain DH5α.

Table 6.2.2. Bacterial strains

Strain	Description	Source
Acinetobacter sp. strain AD3-1	Wild type isolate	Young and Ornston
E. coli DH5α	[ϕ80dlacZΔM15 recA1 endA1 gyrA96 thi-1 hsdR17($r_k^-m_k^+$) supE44 relA1 deoR Δ(lacZYA-argF)U169]	Novagen
E. coli Rosetta	λDE3 F- ompT hsdSB (rb- mB-) gal dcm lacY1 (DE3) pRARE6 (CmR)	Novagen

6.2.1 Creation of knockout mutations within pAD3-1A

In order to determine gene functions of *salA*, *salK* and *salR* within the putative operon knockout mutations were created in the relevant ORFs of the pAD3-1A shotgun clone isolated from *Acinetobacter* sp. strain AD3-1.

6.2.1.2 Disruption of the pAD3-1A siderophore gene

pAD3-1A was digested with *Sph*I and *Sac*I deleting a 381 bp *Sph*I fragment from the 5' end of the pAD3-1A insert. This deletion removed 255 bp of the 498 bp putative siderophore receptor gene. The resulting 5.4 kbp *Sph*I and *Sac*I fragment was religated into pUC19 creating pAD3-1A1 (Fig.6.1.1). The construction of pAD3-1A1 also deleted six restriction sites between *Sph*I and *SacI*, including *Bam*HI, from the MCS of pUC19. pAD3-1A1 was transformed into DH5α, grown up in LB medium as a 500 ml shake-flask liquid culture, at

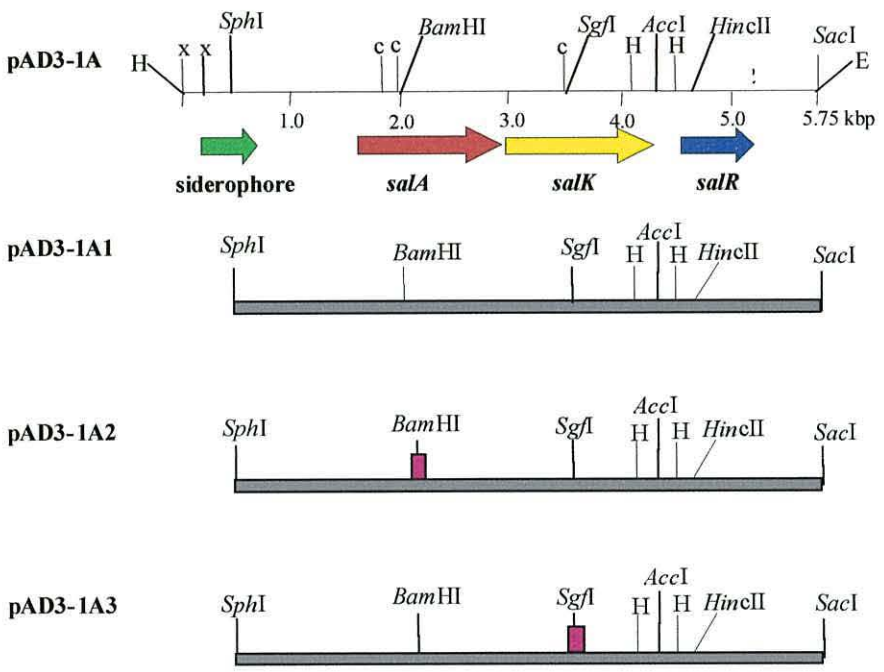


Fig. 6.1.1. Construction of knockout deletions within the siderophore, *salA* and *salK* of the pAD3-1A shotgun clone. The creation of pAD3-1A1 is described in section 6.2.2. Using pAD3-1A1 as the starter sequence a 'fill-in' frameshift deletion was introduced at the unique *Bam*HI site within *salA* creating pAD3-1A2. The unique *Sgf*I site of pAD3-1A1 was used for a 'fill-in' frameshift deletion within *salK* creating pAD3-1A3. Abbreviations: H - *Hin*dIII, x - *Xba*I, c - *Cla*I, E - *Eco*RI.

37°C and 160 rpm under inducing and non-inducing conditions (Chapter two 2.1.2). Cell free extracts were prepared (Chapter 2 section 4 - 2.4.2) and enzyme activity was monitored as wavelength scans of NADH oxidation (Chapter 2 section 4 - 2.4.5).

6.2.1.3 Disruption of the pAD3-1A *salA*

The salicylate hydroxylase gene was disrupted at the now unique *Bam*HI site within *salA* of pAD3-1A1 by performing a frameshift 'fill-in' reaction (Chapter 2 section 2 - 2.2.11) (Fig.6.1.1) creating pAD3-1A2. pAD3-1A2 was transformed into DH5α and treated as described for the pAD3-1A1 construct for growth, induction and analysis of the created mutation (6.2.1.2).

6.2.1.4 Disruption of the pAD3-1A *salK*

Based on sequence homology searches this 1.3 kbp ORF was identified as a transporter gene with closest homology at the amino acid level to the *Acinetobacter baylyi* strain ADP1 benzoate transporter BenK. Due to its location between the *salA* and *salR* genes this gene was putatively termed *salK* for salicylate transport into the cell. Disruption of pAD3-1A1 *salK* coding region was achieved by a frameshift 'fill-in' reaction at the unique *Sgf*I site of *salK* creating pAD3-1A3. This construct was transformed into DH5α and activity was analysed by monitoring the specific activity as described in Chapter 2 (section 4 - 2.4.6.2).

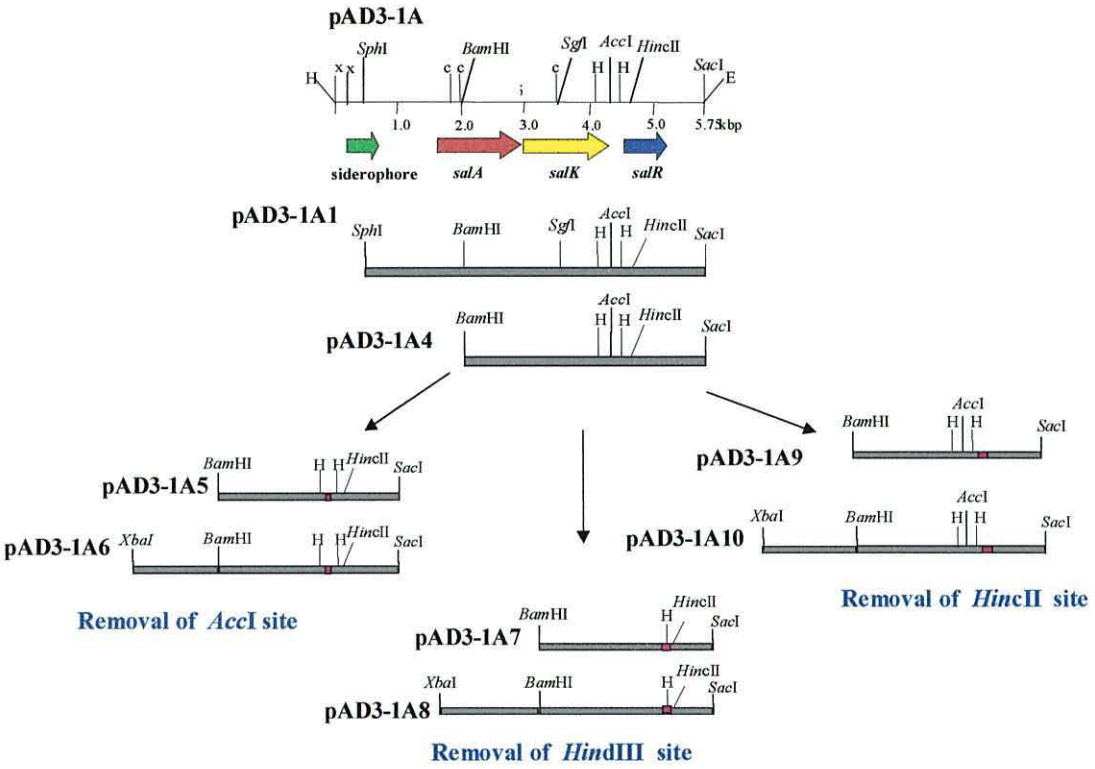


Fig.6.1.2. Construction of *salR* knockout mutations. DNA manipulation steps to create 1) an *Acc*I frameshift deletion within the stem-loop structure immediately downstream of *sal*K, pAD3-1A6. 2) a deletion of 381 bp between the two *Hin*dIII sites of pAD3-1A2 encompassing the terminal region of *salK* and upstream region of *salR*, pAD3-1A8. 3) a *Hin*cII exonuclease deletion within the *salR* gene, pAD3-1A10.

6.2.1.5 Disruption of the pAD3-1A *salR*

Knockout of the salicylate regulatory gene was somewhat more complicated than for the *salA* and *-K* genes. Regulatory gene products are generally not required at the same intracellular levels as for the structural genes. It is common for such regulators to control their own transcription. Nine bps downstream of the *salK* gene lies a sequence of DNA, AGGGAGGTCTACCTCCCT, which appears capable of forming a stem-loop structure that could possibly attenuate transcription of the downstream regulator gene (Fig.6.1.2.3). Upstream of the *salK* sequence is a direct repeat of 11 of the 18 loop base-pairs.

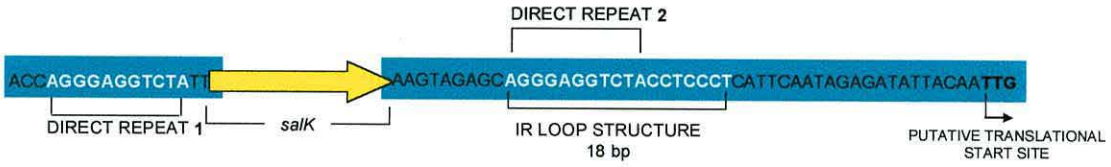


Fig.6.1.3. The putative termination loop of the AD3-1 *salK* encoding a putative salicylate transport protein. 18 bp sequence highlighted in white is the inverted repeat hairpin forming structure. Direct repeats are labeled and also highlighted in white. The alleged TTG translational start is marked with an arrow.

The translational start site of *salR* was also a little ambiguous, possibly beginning from a rarely used TTG start codon. An ATG start site located 231 bp further downstream from the putative TTG start site produced a shorter SalR product in amino acid alignments (Fig.6.3.2.4). Therefore in an attempt to shed light on the regulator gene structure and function, three knockout mutations were created. 1) an *Acc*I frameshift deletion within the stem-loop structure

downstream of *salK* and upstream of *salR* creating pAD3-1A6 (Fig.6.1.2). 2) a 381 bp *Hin*dIII deletion encompassing 132 bp of the 3' terminal end of *salK* and 249 bp of possible intergenic or *salR* DNA creating pAD3-1A8 and 3) a *Hin*cII framshift base-removal deletion (Chapter 2 Section 2.2.11) within the *salR* ORF creating pAD3-1A10 (Fig.6.1.2). Sequencing of pAD3-1A10 confirmed the introduction of a 27 bp deletion encompassing the *Hin*cII site.

To obtain these *salR* mutations the *Hin*dIII and *Sal*I sites from pUC19 were sequentially deleted by 'fill-in' frameshift deletions creating pUC19H$^-$ and pUC19HS$^-$. Destruction of the *Sal*I site also removed the isoschizimer sites for *Acc*I and *Hin*cII. Due to the lack of unique restriction sites within the pAD3-1A1 it was necessary to create some. This was achieved by sub-cloning the 3.6 kbp *Bam*HI-*Sac*I fragment from pAD3-1A1 into pUC19HS$^-$ creating pAD3-1A4 (Fig.6.1.2) which was used as the starting material for all further constructs. These were as follows: the unique *Acc*I site of pAD3-1A4 was filled-in by T4 DNA ligase creating pAD3-1A5. A *Hin*dIII deletion of 381 bp was introduced into pAD3-1A4 by restriction digest followed by re-ligation of the plasmid creating pAD3-1A7. pAD3-1A9 was constructed by the introduction of a chew-back deletion at the *Hin*cII site of *salR*. Once the desired deletions had been introduced into each construct the upstream region of the sequence was sub-cloned in as a *Xba*I-*Bam*HI fragment recreating the complete sequence plus specific deletion creating pAD3-1A6, pAD3-1A8 and pAD3-1A10 (Fig.6.1.2). These constructs were then transformed into DH5α and grown up as induced or uninduced cultures. The mutations were analysed by measuring salicylate-dependent NADH oxidation.

6.2.2 Cloning and over-expression of AD3-1 *salA*

The AD3-1 *salA* was amplified with *Pfu* DNA polymerase (Promega) using pAD3-1A1 as the template DNA. Unique *Eco*RI and *Hin*dIII restriction sites were introduced at the 5' and 3' termini of the gene respectively by PCR mutagenesis to facilitate directional cloning of the gene into pET28a in frame with and downstream of the poly His-tag. The PCR reaction conditions and primers are described in Table 6.2.3.

Purified PCR product was digested with *Eco*RI and *Hin*dIII and ligated into similarly digested His-tagged over-expression vector pET28a. The resulting ligation mixture was digested with *Sac*I to remove any self-ligated plasmids without insert and transformed into DH5a. Plasmid DNA from resulting colonies were screened by restriction digest with *Eco*RI and *Hin*dIII. Clones with an insert of the correct size were termed p196EXA and sequenced to verify the integrity of the insert (Chapter 2 section 2 - 2.2.8). This plasmid was then transformed into the overexpression strain *E. coli* Rosetta™. Over-expression, purification and SDS-PAGE analysis of AD3-1 SalA was performed as described previously (Chapter 2 section 4 - 2.4.4). Characterization of the properties of the purified SalA enzyme and kinetic studies were performed as described in Chapter 2 (section 4 - 2.4.5 to 2.4.6).

Primers			**Restriction site formed**	**Bases changed*** **(underlined)**
Forward primer p196EXAF	5'- ATTTAGTAA***GAATTC***CTAAAAATGGAC -3'		*Eco*RI	**GGAATG**
Reverse primer p196EXAR	5'- GGTCATATTGTCCTG***AAGCTT***ATCTTTGTTG -3'		*Hin*dIII	**TTGGTA**
PCR reaction conditions	***Pfu* DNA polymerase**			
Initial denaturation	95°C - 2 min	1 cycle		
Denaturation	95°C 1- min	15 cycles		
Touchdown annealing	56 - 63°C - 1 min			
Extension	72°C - 3 min			
Denaturation	95°C - 1 min	15 cycles		
Annealing	60°C - 1 min			
Extension	72°C - 3 min			
Final extension	72°C - 10 min	1 cycle		

Table 6.2.3. PCR primers and reaction conditions for the amplification of the AD3-1 *salA*. The PCR product was cloned into the protein expression vector pET28a. *Bases changed from the wild type sequence to create the unique restriction sites are in black and underlined. The restriction site formed is highlighted in blue, italicised and underlined.

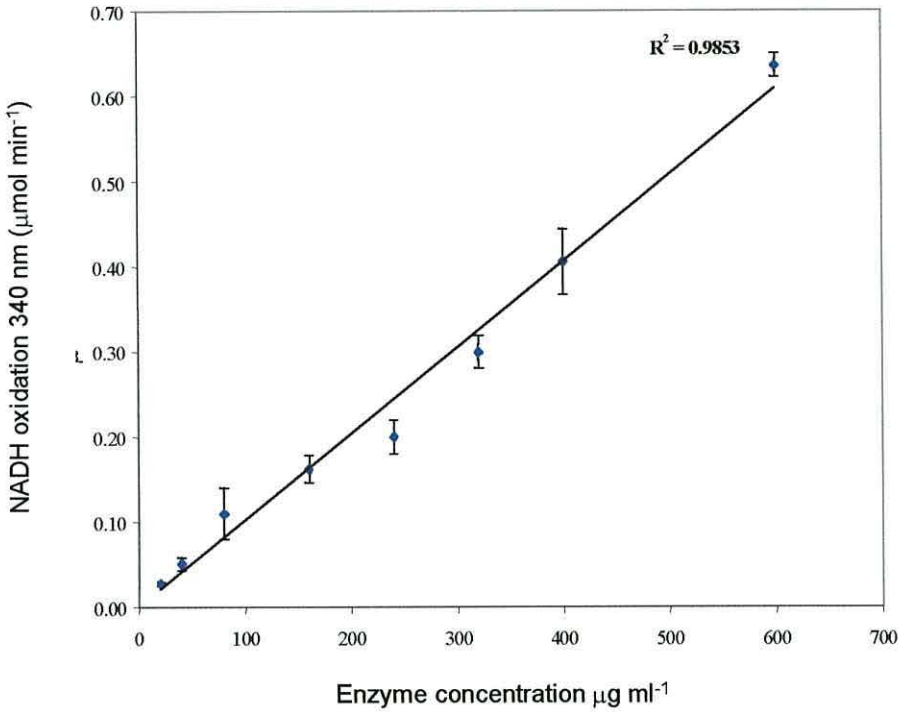


Fig.6.2. NADH oxidation (μmol min^{-1}) vs enzyme concentration (μg ml^{-1}), for example NADH oxidation resulting from increasing amounts of pAD3-1A1 CFE. Average of three replicates ± standard deviations (n=3).

6.2.3 Initial rates

Enzyme activities of both purified enzyme preparations and CFEs were determined by measuring initial rates (the first 10 - 15 s of the reaction). To determine the optimum amount of enzyme to use in the enzyme assays rates of NADH oxidation (μmol min^{-1}) were measured vs enzyme concentration (μg ml^{-1}). An enzyme concentration that lay within the linear region was used in all replicates (Fig.6.2). Further increases in enzyme concentration resulted in a plateau effect as the reaction proceeded too rapidly to allow accurate rate measurements.

6.2.4 AD3-1 salicylate hydroxylase activity against substituted salicylates

The activity profiles of the *Acinetobacter* sp. strain AD3-1 SalA against a variety of substituted salicylates were measured using three sources of the enzyme. Wild type AD3-1 grown in minimal media with salicylate as the sole carbon source, *E. coli* pAD3-1A1, grown in the presence of 1 mM salicylate and purified IPTG induced AD3-1 SalA. Specific activities (Chapter 2 section 4 - 2.4.6.2) were calculated and expressed as activity relative to that against salicylate. Each assay was performed in triplicate.

6.2.5 DNA and protein sequence analyses of pAD3-1A

Analysis of the pAD3-1A gene organization and structure was performed using the DNAstar programme software. NCBI Blastx searches were performed for each gene to identify the closest homologues in the database. Protein alignments and phylogenetic trees were constructed using clustalx, ClustalW, Genedoc and Megalign (DNAstar) software packages.

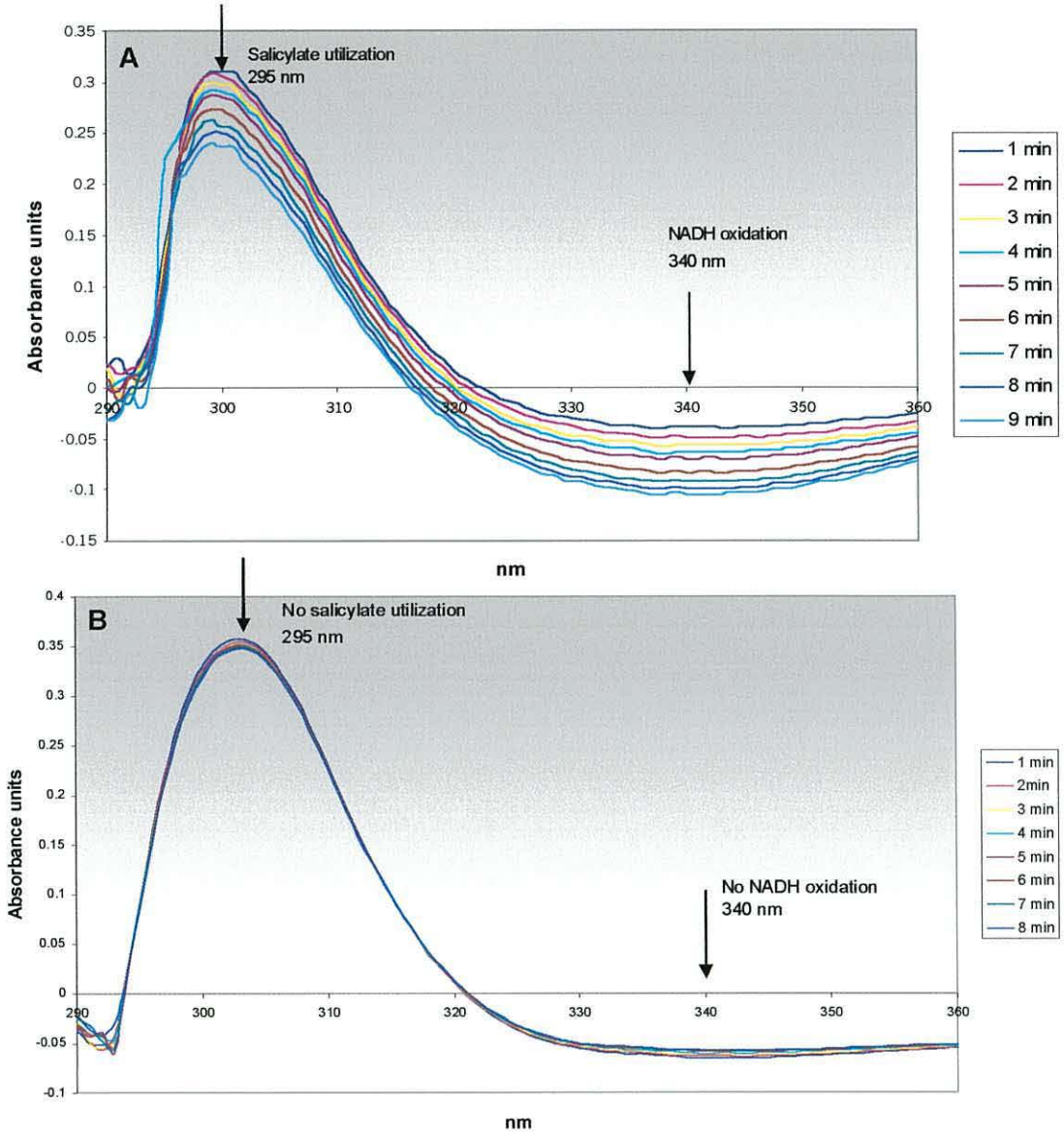


Fig.6.3.1.1 Wavelength scans of NADH oxidation of wild type AD3-1 CFE. **(A)** Induced CFE (salicylate grown cells 2.5 mM). **(B)** Uninduced CFE (succinate grown cells 10 mM).

6.3 Results

6.3.1.1 Wavelength scans of enzyme assays

Salicylate hydroxylase catalyses the oxidation of salicylate in the presence of NADH + H^+ via the introduction of one atom of dioxygen into the aromatic substrate to form catechol with the liberation of NAD^+, CO_2 and H_2O. The bound FAD of the hydroxylase is reduced via the oxidation of NADH + H^+ to NAD^+ a reaction that can be monitored as a decrease in the absorbance spectra of NADH at 340 nm. Wavelength scans of salicylate hydroxylase activity in cell-free extract of salicylate-grown wild type AD3-1 cells demonstrated enzyme activity consistent with the oxidation of NADH to NAD^+ during the reaction (Fig.6.3.1.1.A). The salicylate absorbance peak (295 nm) was observed to decrease during the time-course of the wavelength scan. The catechol product has an absorbance peak at 275 nm but the formation of such a peak was not observed in these wild type scans. A decrease in the absorbance spectra at 340 nm and 295 nm was observed only in the presence of induced enzyme. Cell-free extracts of succinate-grown cells displayed no salicylate hydroxylase activity when monitored as described above (Fig.6.3.1.1B). Reactions performed using a) NADPH as the electron donor, b) without added NADH or c) with no enzyme addition exhibited no change in the absorbance spectra. This indicates that enzyme activity is dependent on the presence of induced enzyme and NADH as the electron donor.

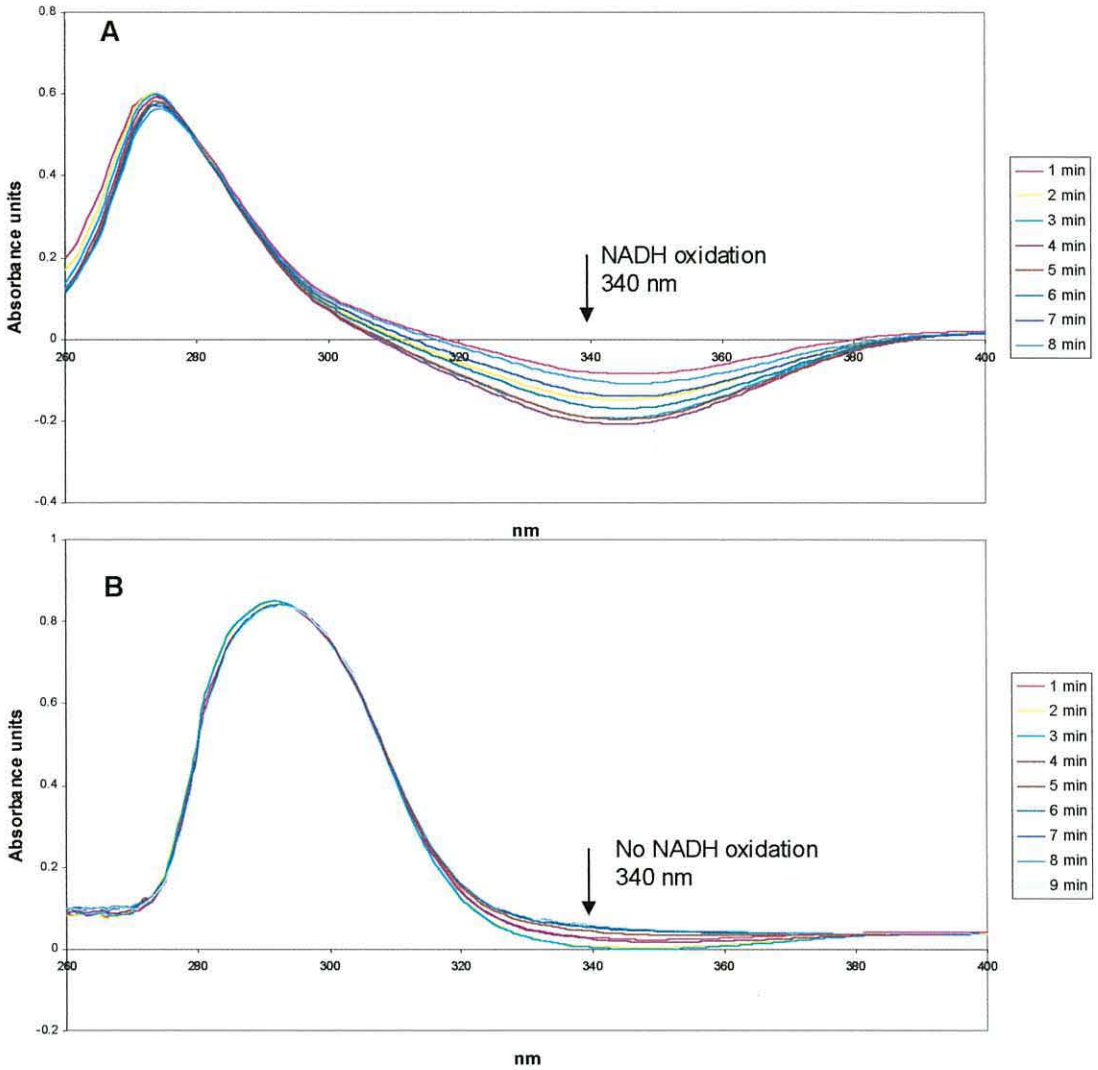


Fig.6.3.1.2 (A) Wavelength scan of salicylate hydroxylase activity in salicylate induced *E. coli* pAD3-1A1. Cuvettes contained: sodium phosphate buffer (50 mM) pH 7.0, 100μM NADH, 60 μg/ml CFE. Reaction started by addition of 100μM salicylate. NADH oxidation observed at 340 nm. **(B)** Wavelength scan of salicylate hydroxylase activity in salicylate induced *E. coli* pAD3-1A2.

6.3.1.2 Wavelength scans of pAD3-1A1

Wavelength scans of *E. coli* pAD3-1A1 cell free extracts induced by growth in the presence of salicylate (Fig.6.3.1.2A) showed NADH oxidation, salicylate utilisation and catechol accumulation. Uninduced pAD3-1A1 cell free extracts (data not shown) displayed no discernable difference to the wild type succinate grown cells depicted in Fig.6.3.1.1.B.

6.3.1.3 Wavelength scans of pAD3-1A2 (*salA*⁻)

Assays performed for both induced and uninduced cell free extracts of *E. coli* cells containing pAD3-1A2, the *salA* knockout mutation, exhibited no NADH oxidation or salicylate utilization (Fig.6.3.1.2B). This indicates that the introduced frameshift mutation had destroyed SalA function providing corroborative evidence that the putative *salA* gene encodes a functional salicylate hydroxylase protein.

6.3.1.4 Effects of the *salK* knockout deletion

The effects of the pAD3-1A3 deletion were measured by comparing the specific activity with that of pAD3-1A1 (Fig.6.3.1.3). The fill-in mutation of this gene created a frame-shift mutation altering the amino acid profile by one base forming a different C-terminal amino acid sequence. A decrease of 60% of the % relative specific activity was observed compared with that of pAD3-1A1 indicating that SalK was active in the transport of salicylate into the cells. The 40% residual activity is possibly the result of passive diffusion of salicylate into

the host cells. Aromatic acid substrates such as benzoate and 4-hydroxybenzoate are known to traverse bacterial cell membranes by diffusion at neutral pH (Collier *et al.*, 1997).

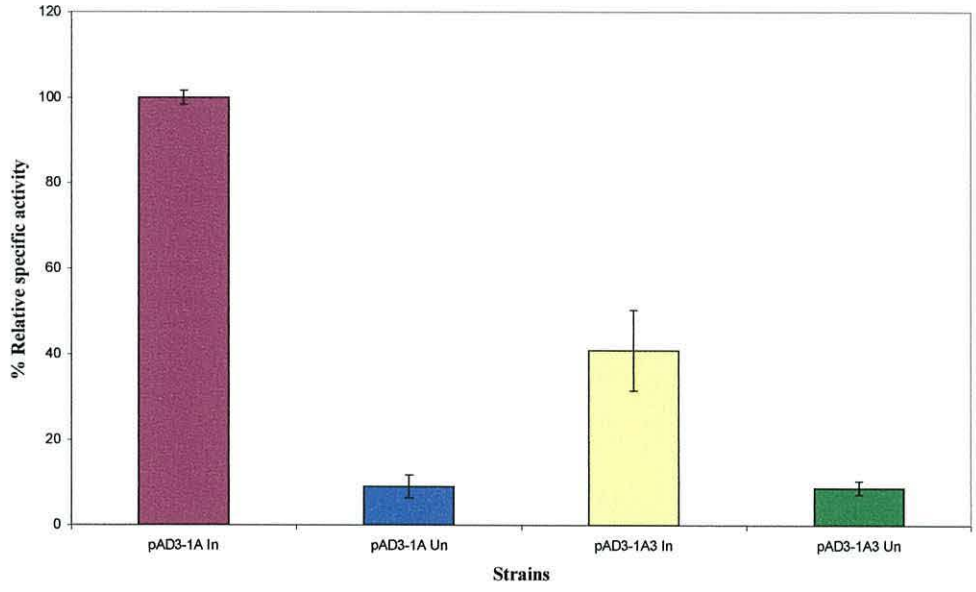


Fig.6.3.1.3. % Relative specific activity of pAD3-1A3 (*sgfI⁻ salK* deletion) compared with pAD3-1A. Bars show average of three replicates ± standard deviations (n = 3).

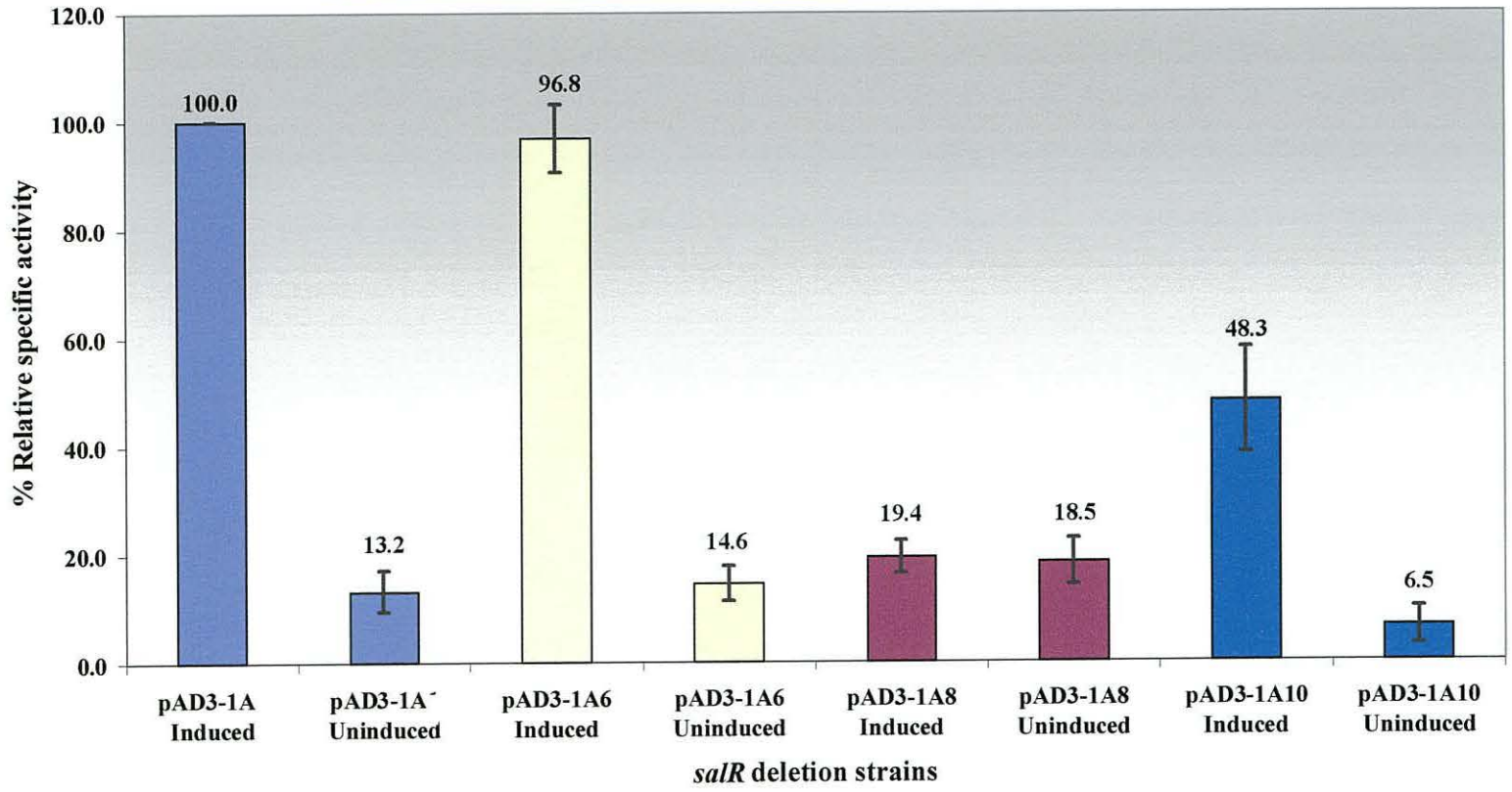


Fig.6.3.1.4 % Relative specific activity of the *salR* knockout constructs, pAD3-1A6 (*Acc*I⁻ loop frameshift deletion), pAD3-1A8 (*Hin*dIII⁻ upstream *salR* deletion - 381 bp) and pAD3-1A10 (*Hin*cII⁻ internal *salR* deletion - 27 bp) compared with the wild type unaltered pAD3-1A sequence. Bars show average of three replicates ± standard deviations (n = 3).

6.3.1.5 Effects of the *salR*⁻ mutations

The deletions designed to disrupt the AD3-1 *salR* were investigated using assays of enzyme activity in cell free extracts. The specific activity observed for each mutated SalR construct was compared against the wild type activity of pAD3-1A cell-free extracts and expressed as % relative specific activity (Fig.6.3.1.4).

6.3.1.5.1 pAD3-1A6

The *Acc*I frameshift deletion of pAD3-1A6 showed no significant difference in % relative specific activity compared to the wild type induced and uninduced standards (Fig.6.3.1.4). The *Acc*I frameshift mutation changed the sequence by only two bps which may have not been effective enough to disrupt the sequence and/or function of the hairpin loop structure. The *Acc*I site is located in the head of the loop structure, a region that does not base pair and probably had no great effect on the formation of the hairpin structure.

6.3.1.5.2 pAD3-1A8

The pAD3-1A8 *Hin*dIII deletion resulted in the complete disruption of SalA activity. The % relative specific activity of this construct was at the same level as that for the uninduced samples (Fig.6.3.1.4).

6.3.1.5.3 pAD3-1A10

Sequencing of pAD3-1A10 revealed the introduction of a 27 bp deletion at the *Hin*cII site within the coding region of *salR*. This deletion however did not completely disrupt SalR function. The specific activity was reduced by ~50% in this construct as compared with the wild type (Fig.6.3.1.4). The *Hin*cII region removed is between the DNA binding motif region (residues 1-66) and the next conserved region involved in co-inducer recognition / response (residues 95-173) (Fig.6.3.1.5). It is assumed that the introduced deletion occurs within a region non-essential for SalR activity creating a truncated protein that retains 50% of the wild type relative specific activity.

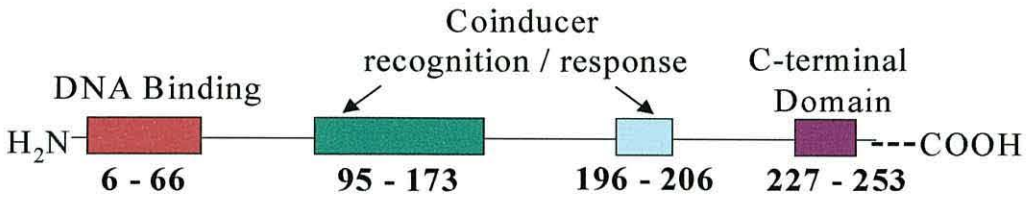


Fig.6.3.1.5. Schematic structure of a LysR-type transcriptional regulator protein showing the conserved regions responsible for protein function (Schell, 1993).

These results show that the expression of *salA* is under the regulation of SalR, a LysR type transcriptional regulator (See section 6.3.2.4). The 381 bp *Hin*dIII deletion between *salK* and *salR* completely destroyed SalA activity, this encompasses the region responsible for DNA binding. The 27 bp *Hin*cII deletion resulted in a decrease in salicylate hydroxylase activity of ~50% suggesting that the mutation produced a truncated regulator protein with partial regulatory activity as assessed by the effects on the specific activity of SalA.

Protein	Organism	% identity	Reference
AD3-1 - SalA	*Acinetobacter* sp. strain AD3-1	100	This study
FS50 - SalA	*Acinetobacter* sp. strain FS50	90	This study
BS6 - SalA	*Acinetobacter* sp. strain BS6	61	This study
ADP1 - SalA	*Acinetobacter baylyi* strain ADP1	60	Accession no.AAF04312 Accession no.YP_046111
AN10 - NahG	*Pseudomonas stutzeri* strain AN10	46	Accession no. AAD02146
ND6 -NAhG	*Pseudomonas* sp ND6 pND6-1	45	Accession no. NP_943122
PpG7 - NahG	*Pseudomonas putida* PpG7	45	Accession no. A39181
pTDG1 - NahG	*Pseudomonas putida* NCIB-9816-4 pTDG1	45	Accession no. NP_863106
P.fluorescens - NahG	*Pseudomonas fluorescens*	39	Accession no. AAL07273

Table 6.3.1. Table accompanying the SalA phylogenetic alignments of Fig.6.3.2.1.

6.3.2 The AD3-1 'sal' operon - organization and structure

6.3.2.1 SalA

The 1.296 kbp salicylate hydroxylase gene encodes a 413 residue protein with a molecular weight of 48.11 kDa. Downstream of the *salA* gene separated by a 45 bp intergenic region is the putative salicylate transporter gene *salK*. The AD3-1 SalA was aligned with other salicylate 1-hydroxylases from *Pseudomonas* (NahG) and *Acinetobacter* spp. (SalA). The NahG homologues from *Pseudomonas* spp. group together in one clade as do the SalA homologues from *Acinetobacter* spp.. The *Acinetobacter* SalAs also group in a pattern similar to that found for their 16S rRNA data (Fig.4.1a).

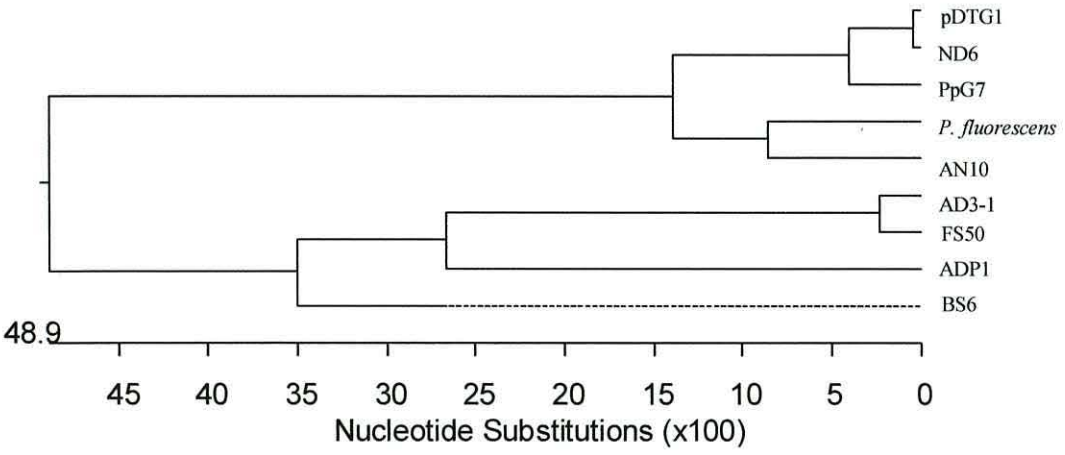


Fig.6.3.2.1. ClustalW generated phylogenetic tree showing the relationship between the AD3-1 SalA and other closely related salicylate 1-hydroxylases of *Pseudomonas* and *Acinetobacter* spp.. See Table 6.3.1 for protein classifications and Fig.6.3.2.2 for amino acid alignment.

```
                    *        20         *        40         *        60         *        80
pDTG1  : ---MKNNKLGLRIGIIGGGISGVALALELCRYSHLQVQLFECAPAFGEVGAGVSFGPNAVAAIVGLGLGEAYLQVADRTSEPWED :  82
ND6    : ---MKNNKPGLRIGIIGGGISGVALALELCRYSHLQVQLFECAPAFGEVGAGVSFGPNAVRAIVGLGLGEAYLQVADRTSEPWED :  82
PpG7   : ---MKNNKLGLRIGIVGGGISGVALALELCRYSHIQVQLFEAAPAFGEVGAGVSFGPNAVRAIVGLGLGEAYLQVADRTSEPWED :  82
Pfluor : MNDMNAKKPALRVAIVGGGISGLALALSLCKHSHLNVQLFEAAPALGEVGAGVSFGPNAVRAIVGLGLGQAYFQVADRTPQPWED :  85
AN10   : MNDMNAKKPALRVAIVGGGISGLALALSLCKHSHLNVQLFEAAPAFGEVGAGVSFGPNAVRAIVGLGLGQSYFQVADRTPQPWED :  85
AD3-1  : ---MLKMDKKIRVAVIGGGIAGLALTTQLVKNKHLDVQLFESASQFSEIGAGISFGANAVKAIQLLGLSQEYESIADQVKTPYTD :  82
FS50   : ---------------IGGGIAGLALTTQLVKNKHLDVHLFESASQFSEIGAGISFGANAVKAIQLLGLSQEYESIADQVKAPYTD :  70
BS6    : ---------------IGGGIAGLALASNLSKHAHLDVQMFESAPQFSEIGAGISFGANAVKAIELLGLANEYHAIADKVSAPFQD :  70
ADP1   : ------VGKKISIAIIGGGIAGVALAANLFKQPHLEVCLYEAAPQFSEIGAGISFGANAVRAIELLGLASQYTAIADQVSAPFQD :  79
                        6GGGI G6ALa  L 4  H6 Vq65E Ap f E6GAG6SFG NAV AI  LGL   Y  6AD    P5 D

              *       100         *       120         *       140         *       160         *
pDTG1  : VWFEWRRGRDASYLGATIAPGVGQSSVHRADFLDALVNHLPKGIAQFRKRATQVEQKGGEVQVLFADGTEHRCDLLIGADGIKSA : 167
ND6    : VWFEWRRGRDASYLGATIAPGVGQSSVHRADFLDALVNHLPKGIAQFRKRATQVEQKGGEVQVLFADGTEHRCDLLIGADGIKSA : 167
PpG7   : VWFEWRRGSDASYLGATIAPGVGQSSVHRADFIDALVTHLPEGIAQFGKRATQVEQQGGEVQVLFTDGTEYRCDLLIGADGIKSA : 167
Pfluor : IWFEWR--PVQSRLSRPHHCGVGQSSVHRADFLDALVKHLPEGIAQFRKRATQIEQQGDELQVLFADATEYRCDLLIGADAIKSA : 168
AN10   : IWFEWRRGSDASYLGATIAPGVGQSSVHRADFLDALVRHLPEGIAQFRKRATQIEQQGDELQVLFADGTEYRCDLLIGADGIKSA : 170
AD3-1  : IWFQWRNGYTDEYLSASIAPQVGQSSVHRADFLDRLIPLVPLTNVHFNKRVQEIEADEEQATIHFIDGQHITFDYVIGCDGIRSV : 167
FS50   : IWFQWRNGYTDEYLSASIAPQVGQSSVHRADFLDRLIPLVPLTNVHFNKRVQEIEADEEQATIHFIDGQHITFDYVIGCDGIRSV : 155
BS6    : VWFQWRNGYTDEYLSASVAAGVGQSSVHRADFLDAIIPHMPTQNVHFSKRLESIEEQEDQVILHFNDGSRHECDYLIGADGIRSV : 155
ADP1   : VWFQWRNGYNDDYLSSSISPQVGQSSVHRADFLDAILGNIPQHQCKFNKKLKSIQEYDTHIELSFEDGTCAEADYVIGADGIHSA : 164
         6WF2WR g    yL a  ap VGQSSVHRADF6Da66  6P     F K4   62       6 F Dg     D 6IGaDgI S

                180         *       200         *       220         *       240         *
pDTG1  : LRSHVLEGQGLAPQVPRFSGTCAYRGMVDSLHLREAYRAQGIDEHLVDVPQMYLGLDGHILTFPVRNGRLINVVAFISDRSEPKP : 252
ND6    : LRSHVLEGQGLAPQVPRFSGTCAYRGMVDSLHLREAYRAQGIDEHLVDVPQMYLGLDGHILTFPVRNGRLINVVAFISDRSEPKP : 252
PpG7   : LRSHVLEGQGLAPQVPRFSGTCAYRGMVDSLHLREAYRAHGIDEHLVDVPQMYLGLDGHILTFPVRNGGIINVVAFISDRSEPKP : 252
Pfluor : LRSYVLEGQGLDHLEPRFSGTCAYRGMVDSLQLRERYRINGIDEHLVDVPQMYLGLYGHILTFPVRKGRIVNVVAFTSDRSQPEP : 253
AN10   : LRSYVLEGQGLDHLEPRFSGTCAYRGMVDSLQLREAFRIKGIDEHLVDVPQMYLGLDGHILTFPVRKGRIVNVVAFTSDRSQPEP : 255
AD3-1  : VRNHVLDSNQLARVEPQFSGTWAYRGIIKFQDFKQAIAKLGNDAEIADVPQMLLGKDKHILTFPIRQGEEINIVAFKSDRTQT-- : 250
FS50   : VRNHVLDSNQLARVEPQFSGTWAYRGIIKFQDFKQAIAKLGNDAEIADVPQMLLGKDKHILTFPIRQGEEINIVAFKSDRTQT-- : 238
BS6    : ARQYVLATHHLLPVQPRFSGTWAYRGIISHASFKQAIVANG-DTDLADIPQMLLGKDKHILTFPIRKGEQINIVAFCSNREDT-- : 237
ADP1   : TRDYVLQTHQFAPVRPNFTGTWAYRGIIKAAEFRQAIVAAGLDVEIADVPQMFLGQNKHILTFPIRQGEDINIVAFKTNPEQR-- : 247
          R  VL    l    P F3GT AYRG66     42a    G D  6 D6PQM LG   HILTFP6R G  6N6VAF 31r
```

Continued on facing page….

```
           260         *       280         *       300         *       320         *       340
pDTG1   : NWPADAPWVRDVSQREMLDAFAGWGDAARTLLECIPTPTLWALHDLAELPGYVHG--RVVLIG-DAAHAMLPHQGAGAAQGLEDA : 334
ND6     : NWPADAPWVRDVSQREMLDAFAGWGDAARTLLECIPAPTLWALHDLAELPGYVHG--RVVLIG-DAAHAMLPHQGAGAAQGLEDA : 334
PpG7    : TWPADAPWVREASQREMLDAFAGWGDAARALLECIPAPTLWALHDLAELPGYVHG--RVVLIG-DAAHAMLPHQGAGAGQGLEDA : 334
Pfluor  : TWPADAPWVREASQPEMLDAFAGWGDR-RRLLECIPAAQLSGQCMTCQNCQATY---TVGLPDRRRSDACCRNQGAGAGQGLEDA : 334
AN10    : TWPADAPWVREASQREMLDAFAGWGDAPRALLGCIPAPTLWALHDLPELPGYVHG--RVALIG-DAAHAMLPHQGAGAGQGLEDA : 337
AD3-1   : VLAENTPWTSPVSKQTMLADFADWSESCKALLDLIDSPTIWALHEIQPLESYKNKSNNVILIG-DAAHAMLPHQGAGAGQGLEDA : 334
FS50    : VLAENTPWTSPVSKQTMLADFADWSESCKALLDLIDSPTIWALHEIQPLETYKNKSNNVILIG-DAAHAMLPHQGAGAGQGLEDA : 322
BS6     : VLPADTPWTRPVDKAQMLSDFSDWSESCQKLLGLIEQPTIWALHEIEELSTYQSASGHVILMG-DAAHAMLPHQGAGAGQGLEDA : 321
ADP1    : TLPEHTPWTRAVDKQEMLDDFQDWSESCRILLGLIERPTLWALHELAELPTYQSHSGRVILMG-DAAHAMLPHQGAGAGQGLEDA : 331
            p   PW r  s   ML  Fa W      LL  I  pt6walh    l  y      V L g daahAmlphQGAGAgQGLEDA

                    *       360         *       380         *       400         *       420
pDTG1   : YFLARLLGDTQVDADNLAELLEAYDDLRRPRACRVQRTSRETGELYEFRDPVVGANEHLLGENLATRFDWLWSHDLDADLAEARA : 419
ND6     : YFLARLLGDTQVDADNLAELLEAYDDLRRPRACRVQRTSRETGELYEFRDPVVGANEHLLGENLATRFDWLWSHDLDADLAEARA : 419
PpG7    : YFLARLLGDTQADAGNLAELLEAYDDLRRPRACRVQQTSWETGELYELRDPVVGANEQLLGENLATRFDWLWNHDLDTDLAEARA : 419
Pfluor  : YFLARLLGLVGPKQATSPSCLELTHDLRRPHACRVQRTTVETGELYELRDPIVGADEQLVGEILATRFDWLWNHDLDADVAEARL : 419
AN10    : YFLARLLGDSRTETGNLPELLGAYDDLRRPHACRVQRTTVETGELYELRDPIVGADEQLVGEILATRFDWLWNHDLDADVAEARL : 422
AD3-1   : LILAKLLENPELTSESIKAVSEIYEQVRLARAVKVQNTSRESGEIYELYSSKY-PDFESIGQHLAQRFDWLWQHSLEDDIQTAKA : 418
FS50    : LILAKLLENPELTSESIKAVSEIYEQVRLARAVKVQNTSRESGEIYELYSSKY-PDFESIGQHLAQRFDW--------------- : 391
BS6     : LILAALLATKNLHSDQLTDVSSIYEKLRVKRACRVQQTSRESGEIYECYSTQY-STFAEIGEHLEHRFDWLWQH----------- : 394
ADP1    : LTLKVLFEHTELTVEDLPRVSAIYEQIRKERACKVQRTSRESGQIYELNSALY-PSFEAVGAHLQNRFDWLWQHDLAQDMLAARA : 415
            La Ll                y  6R  rAc4VQ T3 E3G26YE            6G  La RFDWlw h l  d   a

              *       440
pDTG1   : RLGWENGSRGVLRQG : 434
ND6     : RLGWENGSRGVLRQG : 434
PpG7    : RLGWEHGGGGALRQG : 434
Pfluor  : RMGWEAHEQIALRQG : 434
AN10    : RMGWEAHEQIALRQG : 437
AD3-1   : ALAKQLQHTEMTT-- : 431
FS50    : --------------- :   -
BS6     : --------------- :   -
ADP1    : AIQPVATI------- : 423
```

Fig.6.3.2.2 Clustalx amino acid alignments of SalA homologues. See Table 6.2.2.1 for details

Amino acid alignments of salicylate 1-hydroxylases (Fig.6.3.2.2) reveal highly conserved sequences typical of mono-oxygenases, i.e the ADP binding site (residues 12-34) and the NADH (residues 160-173) binding site which is suggested by Bosch *et al.* (1999) to follow the consensus sequence DXXIXXDGX[K,R]SXXR. Residues Lys163 and Arg167 are important for the formation of salt bridges with the pyrophosphate moiety of NADH. Arg167 is conserved in all of the SalA homologues. In *Acinetobacter* spp., with the exception of strain BS6, Lys163 is replaced by valine. The motif daahAmlpHQGAGAgQLEDA (residues 320-340) is suggested by Bosch *et al.* (1999) to be the substrate catalytic binding site.

6.3.2.2 SalK

The 1.3 kbp ORF termed *salK* encodes a protein of 450 residues and has a molecular weight of 48.9 kDa. Immediately upstream of the *salK* AUG translational start codon a DNA sequence AGGGAGGTCTA occurs. These 11 bp also occur 9 bp downstream of the stop codon of *salK* forming part of an 18 bp GC-rich inverted repeat structure, AGGGAGGTCTACCTCCCT, that is capable of forming a hairpin loop. This loop is possibly involved in the attenuation of transcription of the downstream putative *salR* regulator gene.

A blastx amino acid database search using the putative salicylate permease as a probe identified Salk as a member of the MFS family of transport proteins. The MFS is a large and diverse superfamily currently consisting of 37 families that

Protein	Organism	Transport substrate	% identity	Reference
SalK	*Acinetobacter* sp. strain AD3-1	Salicylate	100	This study
BenK	*Acinetobacter baylyi* strain ADP1	Benzoate	55	Accession No. YP_046120 Accession No. AAC46425
AzoBenT	Azocarus sp. EbN1	Benzoate	49	Accession No. YP_160044
PpBenK	*Pseudomonas putida*	Benzoate	46	Accession No. AAN68773
BenKADP1 putative	*Acinetobacter baylyi* strain ADP1	putative Benzoate	46	Accession No. YP_046472
PcaK	*Acinetobacter baylyi* strain ADP1	Protocatechaute	31	Accession No. YP_046373 Accession No. Q43975
PpPcaK	*Pseudomonas putida*	Protocatechaute	31	Accession No. AAA85137
TfdK	*Ralstonia eutropha*	2,4-dichloro-phenoxyacetate	30	Accession No. YP_025390
VanK	*Acinetobacter baylyi* strain ADP1	Vanillate	26	Accession No. YP_045696
HppK	*Rhodococcus globerulus*	3HPPK	22	Accession No. AAB81315
MucK	*Acinetobacter baylyi* strain ADP1	*cis,cis*-muconate	22	Accession No. AAC27117
LacY	*Escherichia coli*	Lactose	11	Accession No. P02920

Table 6.3.2. Accompanying table to the ClustalW generated SalK phylogenetic tree of Fig.6.3.2.3.

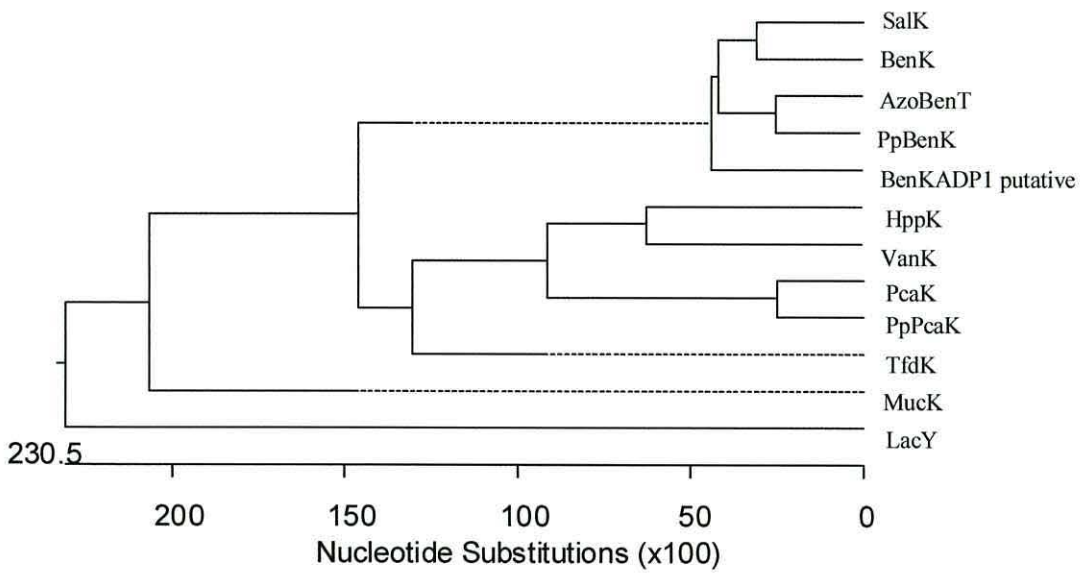


Fig.6.3.2.3. ClustalW phylogenetic alignment of the proposed evolutionary relationship between MFS permeases of aromatic acid/H^+ symporters belonging to family 15 of the MFS group and SalK. LacY is included as an outlier. See Table 6.3.2 for protein classifications. Amino acid alignments are included in Appendix B.

generally group according to substrate specificity. In the ClustalW phylogenetic alignments (Fig.6.3.2.3) SalK groups closely with the aromatic symporters and shows closest homology with benzoate transporters from *A. baylyi* strain ADP1, *Azocarus* sp. EbN1 and *P. putida* (Table 6.3.2).

```
                   *        20         *        40         *        60         *        80
LPR    : SREVYLPHSIEILQLIDLSLIKVFIIVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLFERDKFGVKPTKLAQSLYPSFRQAI :  87
ATGR   : ------------------------------------------------------------------------------------- :   -
TTGR   : --------------LIDLSLIKVFIIVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLFERDKFGVKPTKLAQSLYPSFRQAI :  73
ADP1SR : --------------NVDLSLIRIFICVYENKNISKAAEILNLSQPSVTYNLNRLRKHLNNPLFERTQYGVEATKLSHELYPVFKESI :  73
                        dlsli  fi v e knis aae lnlsqpsvtynlnrlrk ln plfer   gv  tkl   lyp f   i

           *       100         *       120         *       140         *       160         *
LPR    : TDIEMAITASKYFDPKTSTRIFRLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEDWLTEGFLDVAVFNSSYTVMPKIES : 174
ATGR   : ----MAITASKYFDPKTSTRIFRLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEDWLTEGFLDVAVFNSSYTVMPKIES :  83
TTGR   : TDIEMAITASKYFDPKTSTRIFRLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEDWLTEGFLDVAVFNSSYTVMPKIES : 160
ADP1SR : LKIEIAVDEALNFNPLTSNKTFRIGLSDIGEICLLPTLIEYLRAHAPKIKIEVEEIKIDQVEKWLIEGFIDVAVFNSTHLEFKHLEY : 160
           ie6A6taskyF1PkTSt4iFR6GLSD6GEICLLP3LI2YLskeAPfIK6E6EEIqshkVEdWLtEGF6DVAVFNS3ytvmpk6Es

            180         *       200         *       220         *       240         *       260
LPR    : RNLFEERYMCIAHKYHPRIRDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAK : 261
ATGR   : RNLFEERYMCIAHKYHPRIRDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAK : 170
TTGR   : RNLFEERYMCIAHKYHPRIRDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAK : 247
ADP1SR : ETLFLERYVALVNMNHPRIRSTLSFDAYLNESHVAIKSSTGHTQVDHVLKLMGHQRKIALEVPHFGVLQGVLDKTDLMVTLPSRAAQ : 247
         rnLFeERY6c6ahkyHPRIRdTLSleqYfqEnHaAIKSSTGHi26DqrLKe6GlkRKImLE6PHFsVLQdV6ts3eL6VTLPSRAAk

                  *       280         *       300         *
LPR    : VYARNGDVNIFELPFSVKPFNVSANWYNHRDDIEARTWFIQTLQELFQSL : 311
ATGR   : VYARNGDVNIFELPFSVKPFNVSANWYNHRDDIEARTWFIQTLQELFQSL : 220
TTGR   : VYARNGDVNIFELPFSVKPFNVSANWYNHRDDIEARTWFIQTLQELFQSL : 297
ADP1SR : QYLNQSHVRVLELPFQMSEFYVGLHWEAQTNEPLARIWLIQTCKKVISVL : 297
         vYarngdVn6fELPFs6kpFnVsanW5nhr1dieARtWfIQTlqe6fqsL
```

Fig.6.3.2.4. Alignment of AD3-1 SalR and ADP1 SalR from putative start codons. LPR - AD3-1 SalR putative leader-peptide sequence beginning 3 bp downstream from *salK* 3' end. ATGR - AD3-1 SalR truncated sequence translated from the 0.6 kbp ATG start sequence. TTGR - AD3-1R sequence translated from the TTG start codon. ADP1SR - ADP1 SalR sequence.

6.3.2.3 SalR

The exact start of the *salR* allele is unclear and this gene is suspected to have a rare TTG translational start codon. An ORF (663 bp) is present that has the common ATG start codon approx. 300 bp downstream from the 3' end of *salK,* whereas the TTG codon for leucine occurs 48 bp downstream from the *salK* stop codon. The ORF from this position is 894 bp, translation of this sequence encodes a protein of 34.5 kDa that is in good agreement with other LysR/SalR regulator proteins. Alignments of the translated gene from the putative start codons with the ADP1 sequence support this theory (Fig.6.3.2.4). The 663 bp sequence is truncated in these alignments. However a protein message can be translated from immediately downstream of *salK* which runs through the putative hairpin loop structure possibly forming a peptide leader/signal sequence of 14 residues beginning from a serine residue.

The proposed regulator of the '*sal*' operon closely aligns with NAH7 type LysR transcriptional regulators from *Pseudomonas* spp. known to be involved in regulation of the catabolic naphthalene and salicylate operons (Fig.6.3.2.5). It also groups closely with but is distanced from the *A. baylyi* strain ADP1 LysR-type transcriptional regulators CatM and BenM.

Protein	Organism	% Identity	References
AD3-1 SalR	*Acinetobacter* sp. strain AD3-1	100	This study
ADP1 SalR	*Acinetobacter baylyi* strain ADP1	54	Accession No. AAF04311 Accession No. YP_046112
AN10 NahR	*Pseudomonas stutzeri* AN10	31	Accession No. AAD02145
Pf NahR	*Pseudomonas fluorescens*	31	Accession No. AAM18544
pTDG1 NahR	*Pseudomonas putida* NCIB9816-4 pTDG1	30	Accession No. NP_863107
PpG7 NahR	*Pseudomonas putida* Nah7 PpG7	23	Accession No. AAA98184
ADP1 CatM	*Acinetobacter baylyi* strain ADP1	20	Accession No. P07774. Accession No. YP_046130.
ADP1 BenM	*Acinetobacter baylyi* strain ADP1	17	Accession No. AAC46441 Accession No. YP_046121
TfdR	*Burkholderia cepacia*	15	Accession No. AAK81677
VanR	*Acinetobacter baylyi* strain ADP1.	11	Accession No. AAC27105

Table. 6.3.3. Bacterial regulatory genes described in the phylogenetic alignment of Fig.6.2.2.5 Name, source, % identity and accession number.

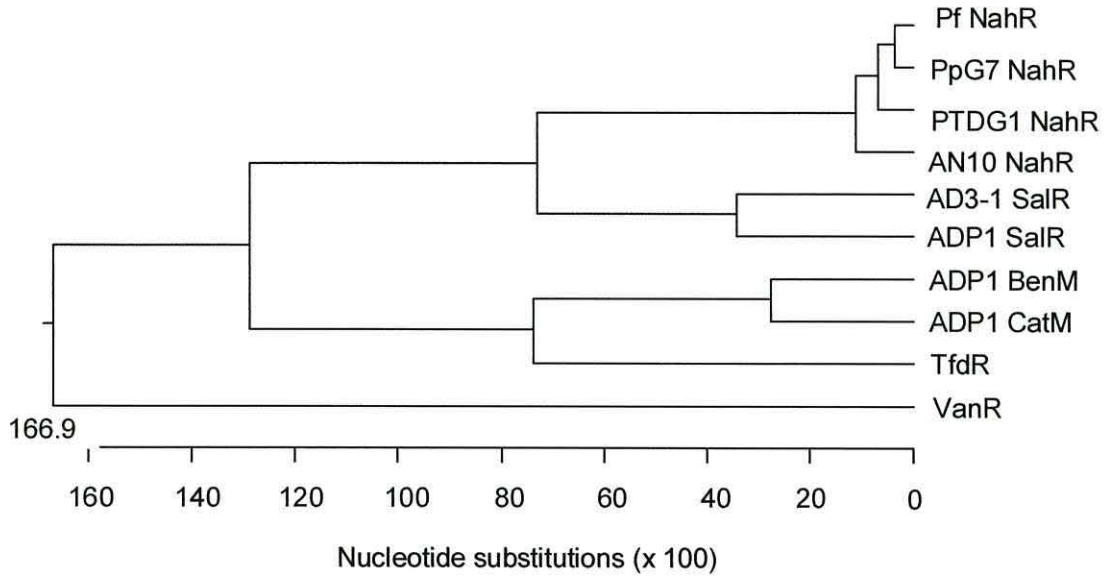


Fig.6.3.2.5. ClustalW generated phylogenetic tree showing the relationship between LysR-type transcriptional regulators involved in aromatic hydrocarbon degradation. Amino acid sequence alignments are included in Appendix B.

	DNA binding region (helix-turn helix motif) residues 6-66
LysR consensus	HpLRpLRxFxxhxpppphSxAApxLphSQPAhSxQhppLEpxLGxxLFxRxpRxhxxxTxA
AD3-1 SalR residues 1-66	HpL-pLRxFxxhxpppphSxAApxLphSQP-hTx-hppL-pxL-xxLFxRxp-xhxxx-x-
	Coinducer recognition / response region residues 98-150
LysR consensus	LxIGx8hLPx10Px5Lx11LpxxphDhh
AD3-1 SalR residues 95-150	LGL-x7hLPx10Px5Ix11LpxxhhDhh
	C-terminal domain residues 236-246
LysR consensus	VxxGxGxVLP
AD3-1 SalR residues 232-242	VxxSxLxTLP

Table 6.3.4. Comparison of the AD3-1 SalR amino acid sequence with the LysR type transcriptional regulator conserved motifs where h = hydrophilic residue, p = hydrophilic residue, x = any residue, - = missing residue, pink letters = mismatch or extra residue and blue letters = substituted residues. LysR consensus sequences taken from Schell (1993).

6.3.2.3.1 Conserved LysR motifs

LysR-type transcriptional regulators regulate transcription of structural genes and also regulate their own transcription (Chapter 1, 1.2.3.4). The conserved motifs of a LysR occur in three regions as described by Schell (1993) (Fig.6.3.1.4). These regions are conserved in the AD3-1 SalR protein and and the properties of the residues within these regions adhere strongly to the consensus pattern (Table 6.3.2.4).

6.3.2.3.2 Potential LysR promoter/regulatory binding sites

The mode of action typical of a LysR regulator is well documented in the literature (Schell, 1993). These proteins bind to specific sites upstream of the promoter region of the target structural genes in the absence and presence of inducing metabolite. Binding of the inducer promotes changes in protein conformation. This change in conformation results in a different binding pattern for the LysR, relieving DNA bending and exposing specific bases to the cell transcription machinery. The regulatory binding sites have a distinctive motif T-N_{11}-A that commonly comprises an interrupted inverted repeat dyad sequence (TCA-N_7-TGA). Examination of the upstream *salA* region revealed the possible presence of three LysR binding site motifs and potential -35 and -10 box sites (Fig.6.3.2.7). Pairwise alignments of the AD3-1 *sal* promoter region with that of ADP1 did not reveal any clear consensus in LysR binding sites and/or promoter/operator sequence. To shed more light on the putative LysR motifs the AD3-1 upstream region was aligned using data obtained from Lessner *et al.* (2003) in their study on a nitroarene-responsive LysR-type transcriptional activator.

```
                                                                                  T- N11-A
            *       100         *       120         *       140         *       160         *
U2     : ---------------------------------------------ATGCCTCAC---CATTATTCAT-GCT-GGTGATTT------T :  31
JS765  : ---------------------------------------------ATGCCTCAC---CATTATTCATTGCT-GGTGATTT------T :  32
R34    : ---------------------------------------------CACCCTCAC---CATTATTCAT-GCT-GGTGATTT------T :  31
Dnt    : ---------------------------------------------CACCCTCAC---CATTATTCAT-GCTTGGTGATTT------T :  32
G7SAL  : ---------------------------------------------GGCCGGGCG--CAATATTCAT-GTT-GATGATTTAT----T :  33
G7NAH  : ---------------------------------------------GACGATTCG--CAGTATTCAC-GCT-GGTGATAAA-----C :  32
Ad3-1  : TAAAATTCCCCTCATCCTAGCCTTTTCCCGTGGGGAGAAGGGATTTTCAATGCAAATATGATTCAC-GCTACATGATCACTTTTTC : 157
Adp1   : ATATGAAAACTTAAACCTATACCGGATAGGCAATTAAGAAAGAAATATGATATA---ATTGTTTTAAGAT-ATTGATTTTAATAAA : 168
                                                        c    c   cAttaTTca  GcT g TGATtt
                                  -35                       -10
            180         *       200         *       220         *       240         *       2
U2     : AACTATCAGACTTGATCTATAGC--GCTATACCGATCGACGCGCCAGAATG-CGCAGCCATTC----------------------- :  91
JS765  : AACTATCAGACTTGATGTATAGC--GCTATACCGATCGACGCGCCAGTAT--CGCAGCCATTC----------------------- :  91
R34    : AACTATCAGACTTGATCTATAGC--GCTATACCGATCGACGCGCTAGAAT--CGCAGCCATTC----------------------- :  90
Dnt    : AACTATCAGACTTGATCTATAGC--GCTATACCGATCGACGCGCTAGAAT--CGCAGCCATTC----------------------- :  91
G7SAL  : ATATATC-GAGTGGTGTTATTTA--TCAATATTGTTTGCTCCGTTATCGTT-ATTAACAAGT------------------------ :  91
G7NAH  : AAAT-TCAACTATGCTTTATTGA--CAAATAAAAGCACGCTCACCATCATCGCGAATACAA------------------------- :  90
Ad3-1  : ATTTATGAAAGA-GATTTATTTAAATAAATATCAAAAAAATAGATAGTATAAACCTAAAATTATAAATTGGATTTATTCTAAAAAA : 242
Adp1   : TTATTTAAAAAATTAATAACTGCTTTAAAAAACTTTAAATTGGATTTAATTGAGTTTGCAGACTAAAAATAAAACAATAAATCAAG : 254
         a  TaTca a  tgat tAt g      AtA c  t  a  cg  a  aT   g a  cA t
```

Fig.6.3.2.6. Alignment of AD3-1 *salA* promoter sequence with naphthalene and nitroarene gene promoters (the alignment of naphthalene and nitroarene gene promoters was taken from Lessner *et al.* (2003) and realigned with the AD3-1 and ADP1 sal gene promoter). U2 - *Ralstonia* sp. strain U2. JS765 - *Comamonas* sp. strain JS765. R34 - *Burkholderia cepacia* R34. DNT- *Burkholderia* sp. strain Dnt. G7sal - *P.putida* G7. G7nah - *P.putida* G7. AD3-1 - *Acinetobacter* sp. strain AD3-1. ADP1 - *A. baylyi* strain ADP1.

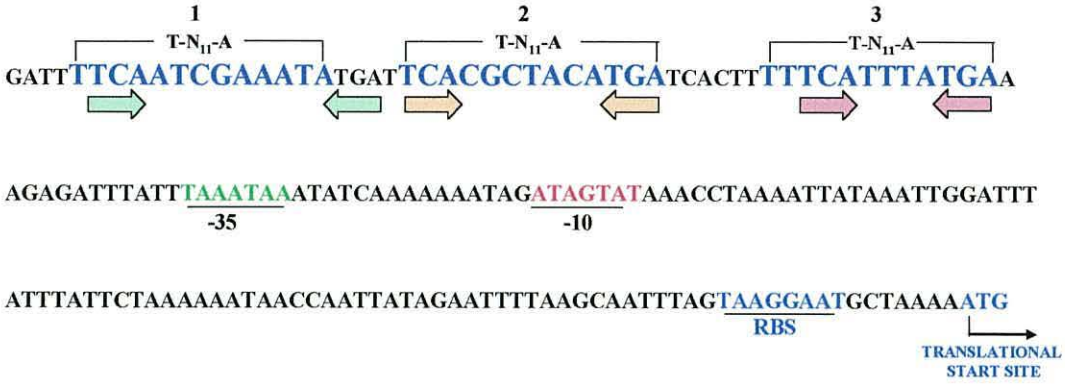


Fig. 6.3.2.7 AD3-1 *salA* promoter region. Upstream of the AD3-1 *salA* gene are three possible LysR T-N_{11}-A binding motifs containing the interrupted inverted repeat TCA-TGA. The LysR binding site (third T-N_{11}-A site), and the -35 and -10 promoter/operator sequences are as putatively identified from the promoter alignments of Fig.6.3.2.6. RBS is the putative ribosome binding site.

This alignment revealed significant homology between AD3-1 and the consensus determined for highly conserved naphthalene promoters. 14 out of 17 bps comprising the G7 NahR consensus-binding site TATTCA-CGCTGGTGATA were conserved (Fig.6.3.2.6). Less homology was seen at the putative -35 and -10 sites. Homology with the ADP1 promoter sequence can also be observed within the consensus but in this position the ADP1 sequence lacks the inverted repeat TCA-Nx-TGA. The binding site reported for ADP1 TCA-N_3-TGA occurs 56 bp upstream of this sequence (Jones *et al.*, 1999).

Putative siderophore proteins	Organism	% identity (over last 500 bps)	Reference
AD3-1 PFSR	*Acinetobacter* sp. strain AD3-1	100	This study
ADP1 PFSR	*Acinetobacter baylyi* strain ADP1	67	Accession No. YP_046425
P. syringae B728a TDSR	*Pseudomonas syringae* strain B728a	51	Accession No. AAY38800
P. syringae DC300 TDSR	*Pseudomonas syringae* pv. tomato strain DC3000	50	Accession No. NP_791435
X. campestris PSR	*Xanthomonas campestris* pv. *campestris* strain ATCC 33913	49	Accession No. NP_636066
B. cenocepacia TDSR	*Burkholderia cenocepacia* H12424	45	Accession No. ZP_00462061
B. parapertussis FSR	*Bordetella parapertussis* 12822	30	Accession No. NP_882442

Table 6.3.5. Accompanying table to Fig.6.3.2.8. Protein homologues of the AD3-1 siderophore. Protein classification, identity and accession numbers. Abbreviations: PFSR - putative ferric iron siderophore receptor. FSR - ferric iron siderophore receptor. TDSR - TonB dependent siderophore receptor. PSR - putative siderophore receptor.

6.3.2.4 The siderophore

The first ORF of the cloned sequence of pAD3-1A had homology with siderophore receptor proteins present in many bacterial species, including *Acinetobacter*, *Pseudomonas*, *Burkholderia* and *Xanthomonas* as identified from Blastx searches. In comparison with the identified homologous proteins the AD3-1 gene appears truncated and may actually begin transcription from a position further upstream in the full chromosomal sequence. The gene transcribed in the same direction as the downstream *salAKR* is small, only 498 bp in length and % amino acid identity values were taken from the last ~500 bp of the alignments with the homologous sequences. Deletion of 255 bp of the putative siderophore receptor protein had no effect on SalA activity.

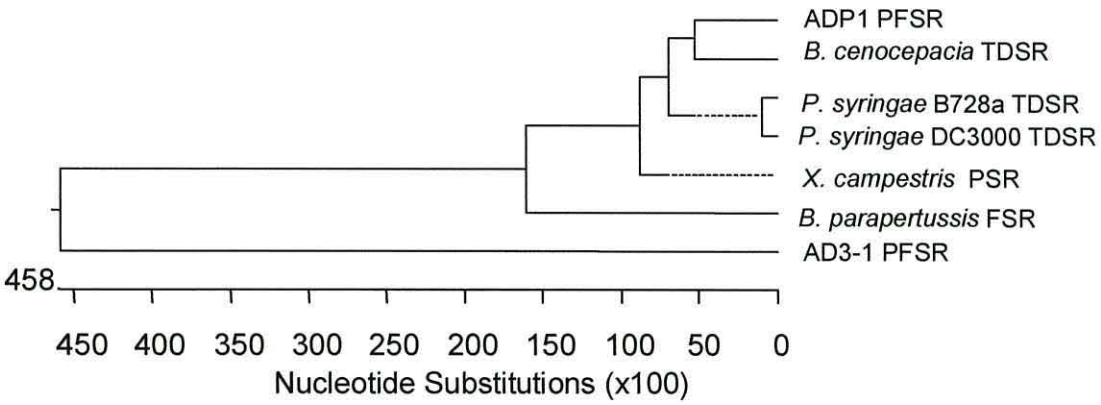


Fig.6.3.2.8. ClustalW phylogenetic relationship between the AD3-1 putative siderophore and homologous proteins identified by blastx searches. Alignments were performed using the whole proteins and the putative truncated AD3-1 siderophore. Amino acid sequence alignments are included in Appendix B.

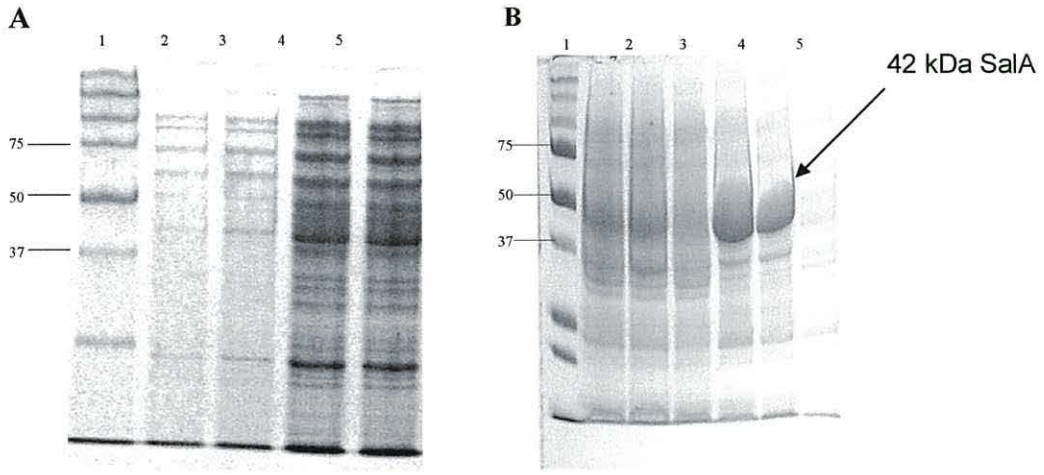


Fig.6.4.A. SDS-PAGE analysis of recombinant SalA purification.

A. Cell-free extract of overexpression culture of recombinant p196EXA incubated and induced at 37°C. **Lane 1** (for A and B). Molecuar weight standard. **Lane 2**. IPTG induced 1 μl CFE. **Lane 3**. Uninduced 1 μl CFE. **Lane 4**. IPTG induced 2 μl CFE. **Lane 5**. Uninduced 2 μl CFE.

Fig.6.4.B. Crude cell lysis of p196EXA cultures. **Lanes 2 - 4** induced and incubated at 37°C. **Lanes 5-7** induced and incubated at 20°C. **Lane 2.** Induced. **Lane 3.** Induced in presence of salicylate. **Lane 4**. Uninduced. **Lane 5**. Induced. **Lane 6**. Induced in the presence of salicylate. **Lane 7** Uninduced.

6.4 SalA purification

Initial over-expression experiments with *E. coli* Rosetta (p196EXA) were conducted at 37°C, induced (0.4 mM IPTG) when the culture reached an OD_{600} of 0.6 and incubated for a further 4 hours. Cell-free extracts (CFE) of this culture were prepared as described previously (Chapter 2 2.4.2). Protein expression was analysed by SDS-PAGE (Fig.6.4 A). No over-expression was detected in these samples. The formation of inclusion bodies during the expression of some recombinant proteins is known to occur. The simplest way to counteract this problem is growth at lower temperatures.

To determine if the lack of an induced protein band was due to the formation of inclusion bodies during growth a simple assay was conducted. Three 5 ml cultures of *E. coli* Rosetta (DE3) (p196EXA) were grown to 37°C to an OD_{600} of 0.6 and then induced with 0.4 mM IPTG, IPTG and salicylate (1 mM), or uninduced as a control. After induction the cultures were incubated for 4 h at 37°C. Three separate 5 ml cultures were grown at 37°C until OD_{600} 0.3 when the temperature was downshifted to 20°C. At OD_{600} 0.8 the cultures were induced as described above at the lower temperature. The cultures were pelleted in a microcentrifuge, the pellets resuspended in 200 µl of SDS-PAGE loading buffer and incubated at 95°C for 5 min to denature the proteins. Samples from the supernatants were analysed by SDS-PAGE (Fig.6.4 B). Protein bands of approx. 48 kDa were observed in those cultures incubated and induced at 20°C but not in those incubated at 37°C (Fig.6.4. B). As a result of this simple assay all subsequent recombinant protein induction procedures were performed at the reduced growth temperatures.

6.4.1 Purification of recombinant AD3-1 SalA

Following initial analyses of protein expression, *E. coli* Rosetta (DE3) (p196EXA) cultures were grown in a total volume of 2 L as described above and in Chapter 2 section 2.4.3. Purification was performed as detailed previously (Chapter 2 section 2.4.4). Upon loading the sample fraction onto the nickel column a bright yellow colour indicative of an oxidized flavoprotein was clearly visible. The protein eluted at a 150 mM imidazole concentration (Fig.6.4.1) and this fraction was also yellow.

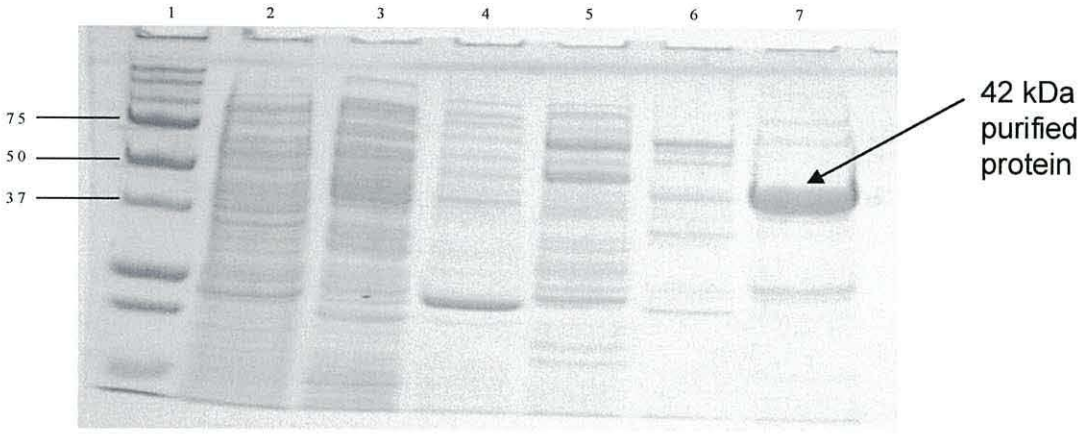


Fig.6.4.1. SDS-PAGE analysis of recombinant SalA purification. **Lane 1.** Molecular weight standard. **Lane 2**. Sample flow through. **Lane 3**. Wash flow through. **Lanes 4-7**. Imidazole gradient (10, 50, 100 and 150 mM respectively). 48 kDa protein eluted at 150 mM imidazole concentration in lane 7 estimated as 95% pure.

6.4.2 Activity profile and properties of purified AD3-1 SalA

Wavelength scans of the purified protein were conducted (Fig.6.3.2) using the previously described method. Protein activity was observed only when the protein was purified in the presence of β-mercaptoethanol (1 mM). The decrease in NADH observed at 340 nm occurs concomitantly with a decrease in salicylate and an increase in the catechol product. The purified enzyme was stable for up to four days when stored at 4°C; attempts to store the purified protein at -20°C resulted in a total loss of enzyme activity.

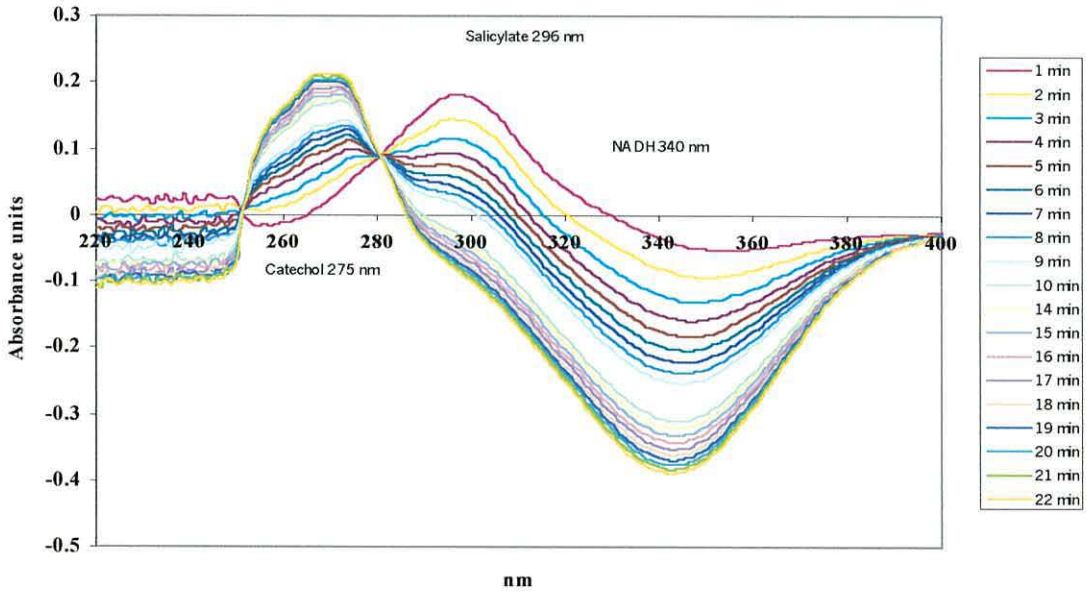


Fig.6.4.2. Wavelength scan of purified AD3-1 salicylate 1-hydroxylase activity. Assay conditions: In test cuvette; sodium phosphate buffer (50 mM), 100 μM NADH, 100 μM salicylate, enzyme concentration 24 μg/ml. The blank cuvette contained the same constituents but without the addition of salicylate.

The properties of the purified salicylate hydroxylase were investigated. The pH optimum was observed to lie within pH 7.0 - 8.0 with sodium phosphate buffer (50 mM)(Fig.6.4.2.1). Tris-HCl buffer over this pH range was observed to have an inhibitory effect (approx. 50% decrease) on enzyme activity.

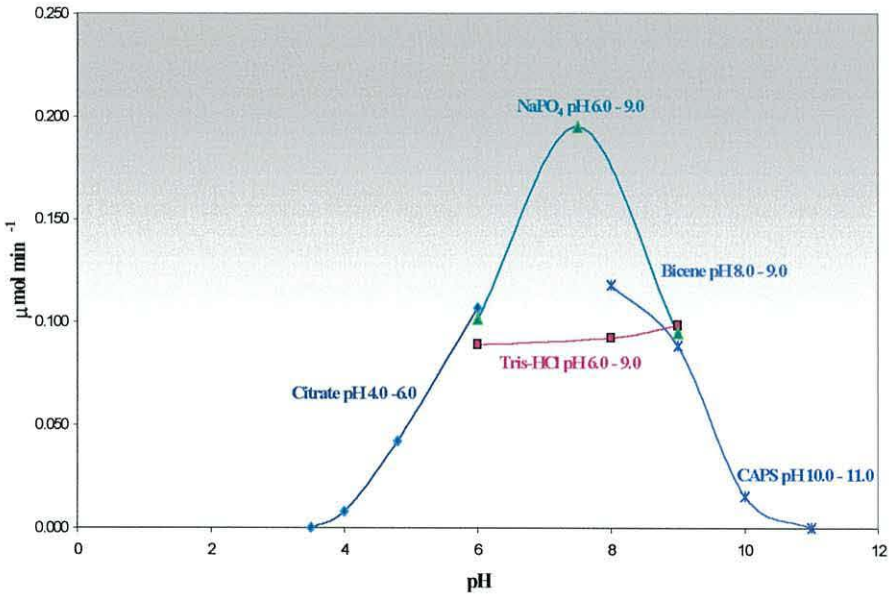


Fig.6.4.2.1. pH optimum of purified AD3-1 salicylate 1-hydroxylase. Optimum pH using sodium phosphate buffer (50 mM) lies within pH 7.0 - 8.0 range.

The enzyme was shown to have a 1:1 stoichiometry between salicylate and NADH (Fig.6.4.2.2) calculated as described (Chapter 2 2.4.5.3).

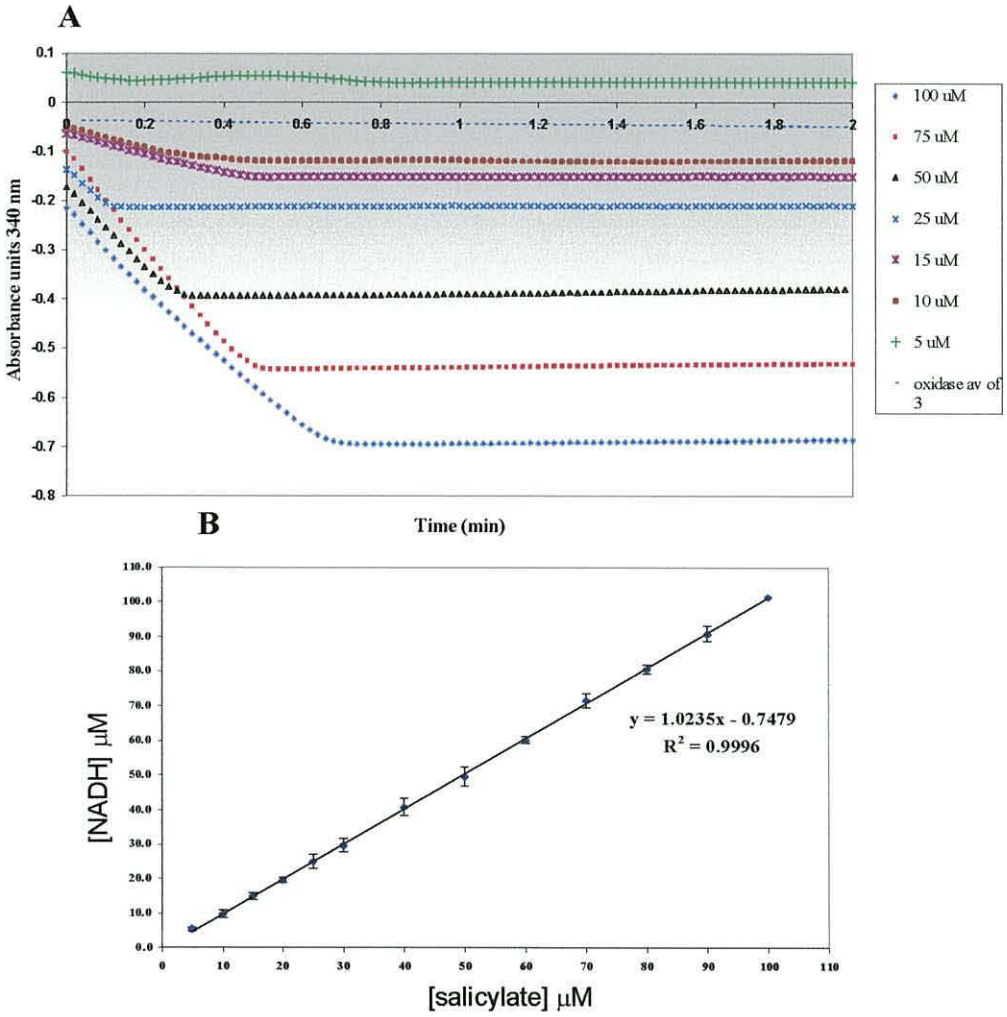


Fig.6.4.2.2. **A.** Stoichiometry traces: NADH oxidation as a function of the salicylate concentration. **B.** 1:1 stoichiometry: The relationship between the concentration of NADH oxidized vs the concentration of salicylate. Presented results are the mean of three separate experiments.

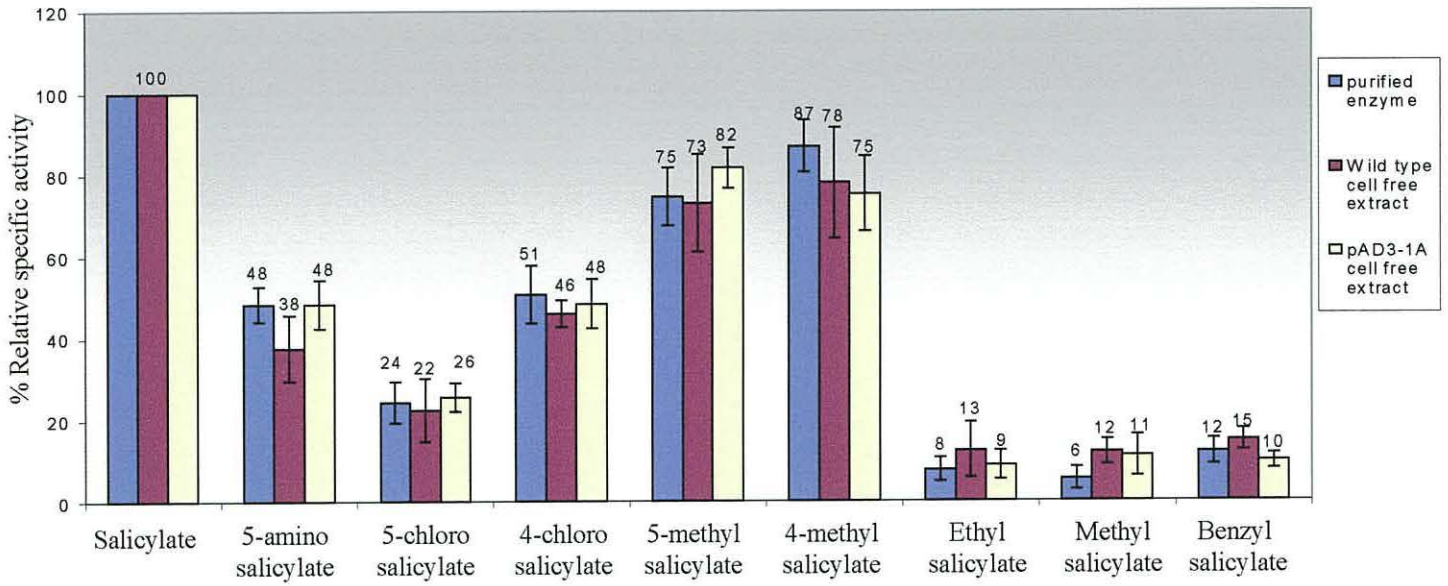


Fig.6.4.2.3. AD3-1 SalA % relative specific activity against substituted salicylates. Bars show average of three replicates ± standard deviation (n = 3).

The specific activity of SalA against substituted salicylates was measured using preparations of AD3-1 wild-type CFE, pAD3-1A1 CFE and purified enzyme (Fig.6.4.2.3). The enzyme was observed to possess relative specific activity of approx 70% with 5- and 4-methylsalicylates and approx 50% relative specific activity with 5-aminosalicylate and 4-chlorosalicylate. Limited activity (25 %) was observed with 5-chlorosalicylate as the substrate and no activity was detected with benzyl, methyl or ethyl salicylates.

6.4.3 Purified AD3-1 SalA kinetic properties

The kinetics of salicylate hydroxylase were investigated using the initial velocities of the two substrates NADH and salicylate. Concentrations of salicylate varied at three fixed concentrations of NADH produced a Michaelis-Menten kinetic plot (Fig.6.4.3) that showed hyperbolic curves.

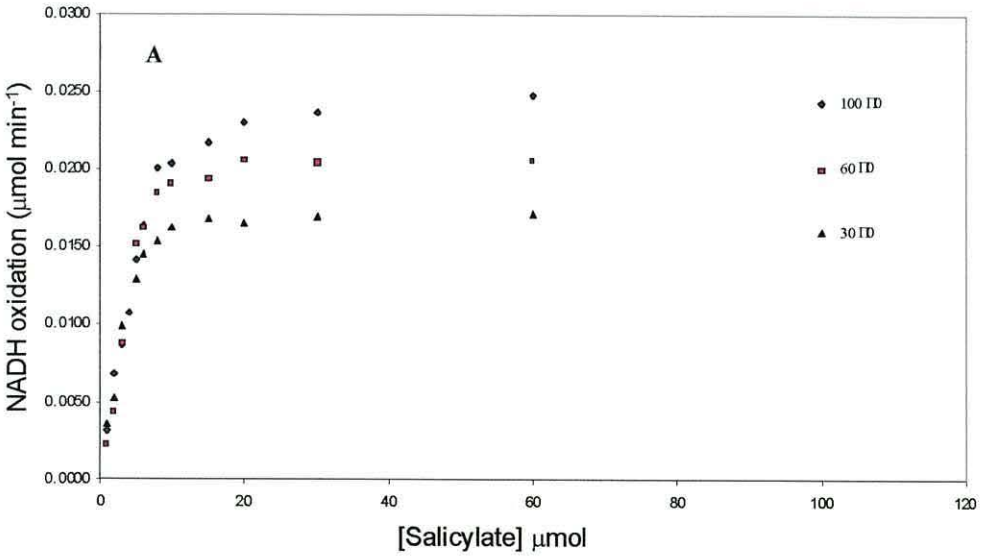


Fig.6.4.3. Michaelis-Menten kinetic plot of NADH oxidation measured as µmol min^{-1} for varied salicylate concentrations at fixed concentrations of NADH.

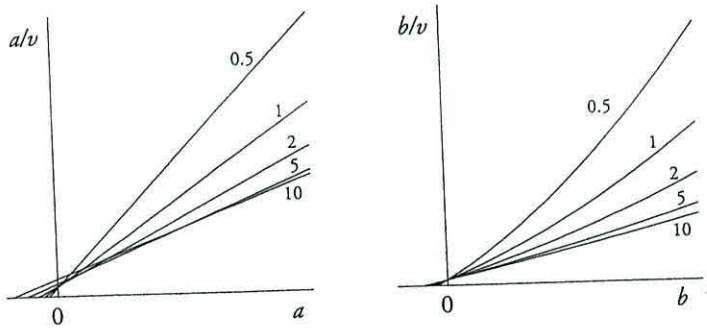


Figure 6.8. Effect of substrate inhibition by B (with $K_{siB} = 10K_{mB}$) on primary plots for substituted-enzyme mechanisms (cf. Figure 6.6).

Fig.6.4.3.1.a. Photocopied page from Cornish-Bowden (1976) showing the effects of substrate inhibition in a bi-substrate reaction with Hanes-Woolf primary plots. This type of inhibition can be useful as it allows an initial determination of the reaction mechanism, i.e. compulsory ternary complex formation and an indication as to which order the substrates A or B bind to the enzyme.

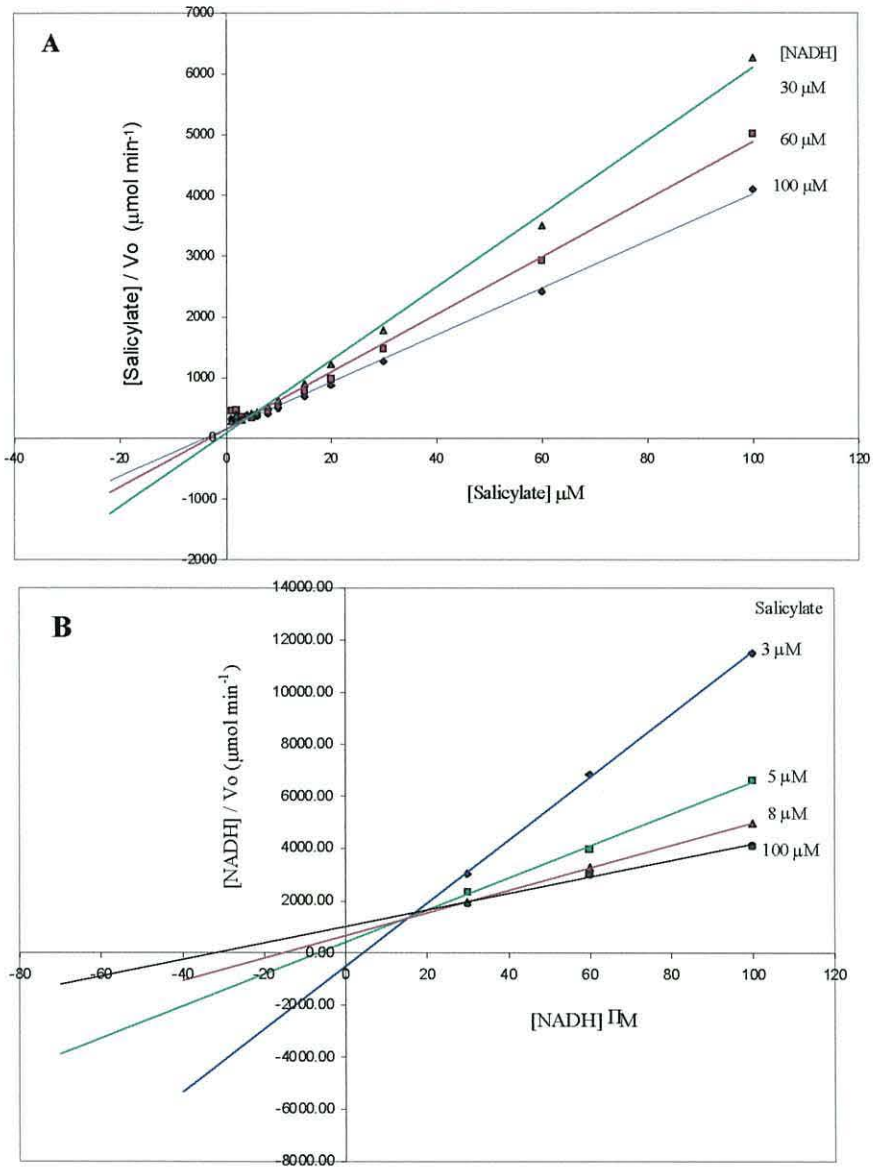


Fig.6.4.3.1. Hanes-Woolf primary plots. **(A)**. Salicylate vs salicylate / rate at three constant concentrations of NADH. **(B)**. Using the same data generated in **(A)** Hanes-Woolf plot of NADH vs NADH / rate at constant concentrations of salicylate.

The Michaelis-Menten data were analysed using Hanes-Woolf primary and secondary plots (Fig.6.4.3.1) to determine K_m and V_{max} and to analyse the basic enzyme mechanism. The Hanes-Woolf primary plot for salicylate (Fig.6.4.3.1 A) appears to follow a substituted (ping-pong) enzyme mechanism but the lines are slightly parabolic converging at the salicylate/V_{app} (apparent V_{max}) axis.

When the plot is reversed (Fig.6.4.3.1 B) and plotted as NADH/V_{app} the plot should be similar. However in plot B the lines are linear, have no common intersection point and converge to the left of the NADH/V_{app} axis at a positive value of NADH. These plots are identical to the effects described for substrate inhibition (Cornish-Bowden, 1976; see Fig.6.4.3.1.a). Substrate inhibition occurs by a reaction constituent binding to the wrong form of the enzyme forming a dead-end complex or possibly high concentrations of salicylate block access of NADH to the enzyme preventing the ordered binding of the second substrate, NADH. The direct reaction with salicylate as the varied substrate proceeded at a rapid rate at low substrate concentrations and was difficult to measure accurately, therefore initial rate measurements were performed with 5-methylsalicylate as the variable substrate. Michaelis-Menten hyperbolic curves were again observed (Fig.6.4.3.2)

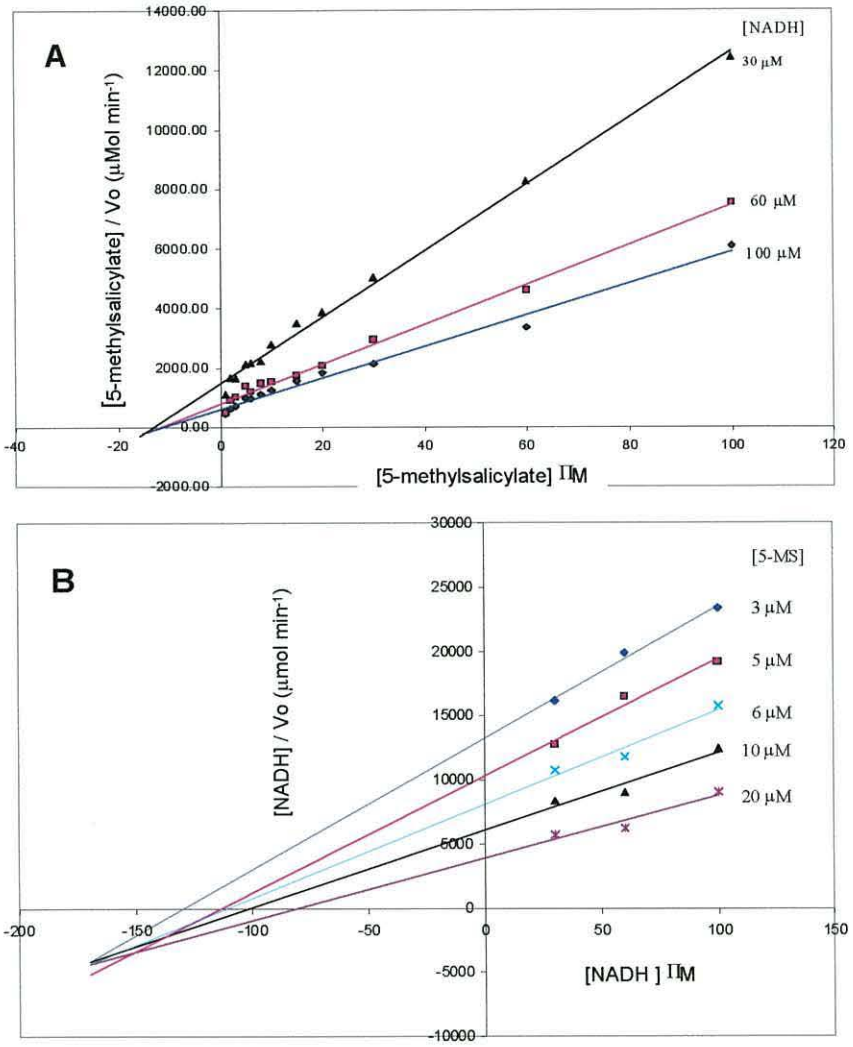


Fig.6.4.3.3. Hanes-Woolf plot for 5-methylsalicylate (5-MS). **(A)** Varied 5-MS at constant concentrations of NADH. **(B)** Using the same data as **(A)** varied NADH at constant concentrations of 5-MS.

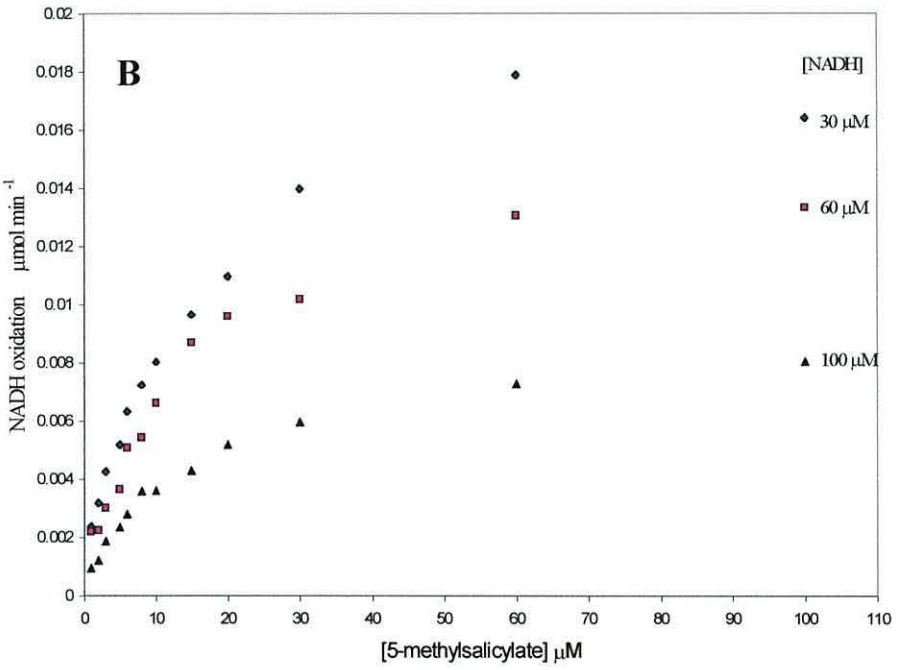


Fig.6.4.3.2. Michaelis-Menten kinetic plots for 5-methylsalicylate. The rate of oxidation of NADH measured as μmol min^{-1} at varied 5-methylsalicylate concentrations at three constant levels of NADH.

Hanes-Woolf primary plots of these data produced sets of convergent lines indicative of a ternary complex formation mechanism between enzyme, salicylate and NADH (Fig.6.4.3.3 A and B). (The linear lines converge at a point to the left of the substrate/velocity axis). V_{max} and K_m for both salicylate and 5-methylsalicylate were calculated by the direct linear method using Enzpack kinetic data modelling programme. The results are presented in Table 6.4.

	V_{max} (μmol min^{-1})	K_{cat} s^{-1}	K_m μM	68% confidence limits
Salicylate	0.03	6.1 x 10^7	5.56	0.0263 - 0.0302
$NADH_{(sal)}$	0.03	6.1 x 10^7	27.3	0.0267 - 0.0343
5-MS	0.017	3.4 x 10^7	9.53	0.0152 - 0.0184
$NADH_{(5\text{-}MS)}$	0.017	3.4 x 10^7	110	0.0117 - 0.0396

Table 6.4. Values for V_{max} and K_m calculated by direct linear method (Enzpack).

The results suggest that the enzyme follows a compulsory ordered ternary complex mechanism where the binding of salicylate to the free enzyme facilitates the binding of NADH. After the FAD reduction NAD^+ is released from the complex whereupon O_2 associates with the reduced enzyme salicylate complex. This reaction scheme has been well documented for salicylate 1-hydroxylase.

6.5 Discussion

6.5.1 SalA

The putative operon *salAKR* is involved in the catabolism of salicylate to catechol. Gene knockouts performed on the individual genes resulted in either total or partial loss of SalA activity. The SalA enzyme was shown directly to catalyze conversion of salicylate to catechol with stoichiometric consumption of NADH. Substrate specificity experiments indicate that the AD3-1 SalA has a broader substrate specificity for 4- and 5-methylsalicylates than the ADP1 enzyme with an increase in observed relative specific activity of 20 and 50% respectively. Similar activities for 4- and 5-chlorosalicylate were observed as reported for the ADP1 SalA (Jones *et al.*,1999). Kinetic studies performed on the enzyme suggest an ordered ternary-complex formation enzyme mechanism where salicylate forms a binary complex with the enzyme facilitating the ordered binding of NADH. The K_m values for salicylate and NADH also support this theory (5.5 and 27.3 μM respectively.), showing a greater affinity of the enzyme for salicylate. An enzyme mechanism of this order was described in early work characterizing the kinetic properties of the salicylate hydroxylase from *P. putida* (Takemori *et al.*, 1970).

Comparisons with salicylate 1-hydroxylases from the NCBI database reveal conserved regions associated with enzyme function i.e. FAD fingerprint, NADH binding sites and substrate catalytic sites (Eppink *et al.*, 1997; Bosch *et al.*, 1999). The phylogenetic relationships observed with the *Pseudomonas* spp. enzymes suggest a common ancestry of the salicylate genes with distinctive

Protein	Transmembrane loop 1	Transmembrane loop 2-3	Transmembrane loop 8-9
SalK (AD3-1)	DGYD	GSLSDKLESYGFSRK	GYLADRFNLAK
BenK (ADP1)	DGYD	GTIADKLEHLGVSRK	GYLADRYNVKF
MucK (ADP1)	DGAD	GWACDRFGRV	GFMADKLGRRF
PcaK (ADP1)	DGID	GPTADR-----FGRK	GWAMDRFNPNR
PcaK* (*P. putida*)	DGLD	GPLADR-----FGRK	GWAMDRYNPHK

Table 6.5. Conserved amino acid motifs occurring within aromatic acid/H^+ symporters of the MFS family (cluster 15).

* Motif sequences taken from Ditty and Harwood (2002).

evolutionary divergence occurring at species level hence the defined clusters denoting the separation between *Acinetobacter* and *Pseudomonas* spp..

6.5.2 SalK

The *salK* gene knockout experiment of pAD3-1A3 indicated that *salK* encodes a protein active salicylate transport. MFS transporters share conserved amino acid sequence motifs GXXXD[R/K]XGR[R/K] in the hydrophilic regions between the second and third transmembrane regions and although to a lesser extent the eighth and ninth transmembrane regions required for substrate accumulation. Aromatic acid/H^+ symporters possess these motifs and share a third conserved motif consisting of four charged amino acid residues, DGXD, in the first predicted transmembrane region (Ditty and Harwood 1999; 2002). From the amino acid alignments of SalK (Appendix 2) these conserved motifs are also present at i) residues 49-52 (DGYD), ii) residues 93-107 (GSLSDKLESYGFSRK) and iii) residues 336-344 (GYLADRFNLAK). Interestingly the SalK (AD3-1) and BenK (ADP1) homologues share five extra residues in the second and third transmembrane region motif not present in the other aromatic acid transporters and have a greater identity within the conserved motifs (Table 6.4).

6.5.3 SalR

The organisation of the AD3-1 salicylate genes (*salAKR)* is similar to that of ADP1, with the exception of *salK*, in that the regulatory gene is located downstream of *salA* and transcribed in the same direction (Jones *et al.*, 1999).

A very similar gene organisation to AD3-1 is inferred from the partial sequencing of the second shotgun clone pFS50-1 (Chapter 4 4.3.4.5), the last 1.8 kbp of which is ~98% identical to that of AD3-1 identifying the regulatory gene as non-divergently transcribed and downstream of the structural genes (Appendix B). It is possible that this gene order is prevalent in *Acinetobacter* spp.. This organization differs from that observed with most LysR-type activators where the regulatory protein is divergently transcribed from the structural genes. The divergent promoters overlap allowing dual regulation of structural genes and LysR activator (Schell, 1993). Despite these differences the AD3-1 *sal* genes appear to be expressed from a promoter region with significant similarity to that of naphthalene and nitroarene catabolic operons (Schell and Wender, 1986; Lessner *et al.*, 2003).

Chapter 7

The purpose of this research project was to study quantitatively the occurrence of a specific gene transfer by transformation between *A. baylyi* strain ADP1 and other *Acinetobacter* species using a non-antibiotic screening method. Detection of transformation was demonstrated by restoration of salicylate gene function i.e. the ability to utilize this compound as a sole carbon and energy source. Identification and isolation of salicylate utilizers provided us with an opportunity to study the diversity of gene structure and organization of *sal* genes present within closely related bacteria that also may be prevalent in the environment. The effects of the presence of other salicylate hydroxylase genes on the transformation frequencies of ADP1 within the soil / worm transformation microcosm study may be taken into account and may help to understand the potential of actual gene transfer events between representative soil microbial populations.

Of the 34 *Acinetobacter* isolates tested for the purpose of this study only eight strains appeared to catabolize salicylate during phenotypic screening. Putative *sal* operons were isolated from the genomes of strains AD3-1 and FS50 by shotgun cloning. Degenerate PCR performed on these and one other strain, *Acinetobacter* sp. strain BS6, resulted in the isolation of *salA* homologues. Southern hybridization also indicated *salA* gene homologues in two other strains, AD321 and FS30 although these genes were not isolated and the intensity of the chemiluminescent signal indicated low homologies. When sequenced, the *salA* homologues were shown to have approximately 60% identity at the nucleotide level with that of the ADP1 *salA*. The BS6 homologue also had ~60 % amino acid identity with AD3-1 and FS50. Amino acid alignments of these and the ADP1 SalA with *nahG*-encoded salicylate

hydroxylases of different *Pseudomonas* spp. revealed a close relationship, indicating a common ancestral evolution that has diverged over time. Strains AD3-1 and FS50 were isolated from the environment in two different geographical locations (USA and UK respectively). Sequencing of the putative *sal* operons of these strains (complete and partial respectively) revealed virtually identical *salA* and *salR* genes (~98% homology) and also suggested the presence of the AD3-1 putative salicylate transporter protein (*salK*) within FS50.

Quantitative transformation frequencies were assessed using just one donor strain, AD3-1 and various forms of donor DNA. Restoration of catabolic function was observed at low levels (10^{-7}) using heterologous donor DNA. In comparison, entirely homologous ADP1 DNA displayed varying transformation frequencies depending on the length of flanking sequence homology and gene copy number. Colony PCR on randomly selected heterologous DNA transformants suggested that recombination occurred solely within the *salA* gene itself, possibly within small regions of microhomology.

The transformation frequencies observed in mutant strain ADPW259 with homologous ADP1 genomic DNA and both plasmid and chromosomal heterologous AD3-1 DNA fell below the level of detection ($<10^{-8}$). However, transformation frequencies of 10^{-2} and 10^{-4} were observed with the homologous plasmid donors containing an insert of either the ADP1 *sal* operon (pADPW34) or the single *salA* gene (pADP1A) respectively. The reduced transformability of strain ADPW259 was anomalous compared to other results obtained in this thesis: all the other 4-bp deletion mutants were detectably transformed. This result and the inability to insert the *sacB*-Km cassette within a central region of *salA*, when it could be successfully inserted at the 3' termini to create strain

pADPW250

```
600        607        615        620            631        638 bp
|          |          |          |              |          |
CATTGTCGCAGCCGCCCTAGATGTAGAAATTGCCGACGT

CATTGTCGCAGCCGCGGATCCTGTAGAAATTGCCGACGT
                   BamHI site
```

ADPW259

```
600        607        615        620            631        638 bp
|          |          |          |              |          |
CATTGTCGCAGCCGCCCTAGATGTAGAAATTGCCGACGT

CATTGTCGCAGCCGGATCCGATGTAGAAATTGCCGACGT
                   BamHI site
```

Fig.7.1 Positions of the engineered *Bam*HI site within the *salA* in construct pADPW250 and strain ADPW259. Nucleotide numbers (bp) indicate the region of the sequence in question relevant to the *salA* translational start site (GTG). pADPW250 top line - wild type sequence, bottom line - containing the restriction site. ADPW259 as above. Blue highlight – the GC-rich 7 bp repeated sequences. Pink highlight - introduced *Bam*HI restriction site.

ADPW249, led to the examination of the DNA sequence immediately adjacent to the position of the mutation in strain ADPW259 (Fig. 7.1). By coincidence the engineered *Bam*HI site that marked the position of the 4 bp deletion was in an almost identical position as the constructs pADPW250 (*salA*$^-$ *Bam*HI) and pADPW251 (*salA*::*sacB*-Km cassette) originally designed to incorporate the positive selection cassette in the centre of the chromosomal allele (see Chapter 3 sections 3.2.2. - 3.2.3 and 3.3.1.1 - 3.3.1.2). However, with pADPW251 as the transforming DNA all attempts to insert the *sacB*-Km cassette into this chromosomal region resulted in anomalous transformants that were kanamycin resistant but also sucrose tolerant, indicating that the expected 'simple' insertion of the cassette had not occurred. Furthermore, colony PCR performed with *salA* specific primers on these transformants produced PCR fragments of wild type length (~1.2 kbp) inconsistent with the expected size indicating the presence of the *sacB*-Km cassette. A 7 bp GC-rich DNA sequence, GCAGCCG, was observed in the ADPW259 mutation region that was repeated 17 bp downstream in reverse, GCCGACG. These two 'motifs' were present on either side of the engineered mutation in ADPW259 (Fig.7.1). Whether these sites form a structure that has any relevance to the process of or inhibition of homologous recombination is probably unlikely and their presence at this deletion site may be just coincidental.

The homologous recombination cross-over events using heterologous DNA is suggested to have occurred within small nucleotide regions of micro-homology, as the acquisition of restriction sites from the heterologous donor DNA in the transformants was not detected (Chapter 5 section 5.3.4.3). Alignments of the ADP1, deletion strain, donor DNA and the primary amino acid sequences (Fig.7.3) have attempted to assess the possible effects on the protein structure of incorporation of heterologous donor DNA bordering the deletion site in

```
                 *        20         *        40         *        60         *        80
ADP1   : ---VGKKISIAIIGGGIAGVALAANLFKQPHLEVCLYEAAPQFSEIGAGISFGANAVRAIELLGLASQYTAIADQVSAPFQDVWFQWR :  85
AD3-1  : MLKMDKKIRVAVIGGGIAGLALTTQLVKNKHLDVQLFESASQFSEIGAGISFGANAVKAIQLLGLSQEYESIADQVKTPYTDIWFQWR :  88
FS50   : ------------IGGGIAGLALTTQLVKNKHLDVHLFESASQFSEIGAGISFGANAVKAIQLLGLSQEYESIADQVKAPYTDIWFQWR :  76
             kki  a IGGGIAG6ALttqLvKnkHLdV L5EsAsQFSEIGAGISFGANAV4AI2LLGLsq2YesIADQVkaP5tD6WFQWR

          *       100        *       120        *       140        *       160        *
ADP1   : NGYNDDYLSSSISPQVGQSSVHRADFLDAILGNIPQHQCKFNKKLKSIQEYDTHIELSFEDGTCAEADYVIGADGIHSATRDYVLQTH : 173
AD3-1  : NGYTDEYLSASIAPQVGQSSVHRADFLDRLIPLVPLTNVHFNKRVQEIEADEEQATIHFIDGQHITFDYVIGCDGIRSVVRNHVLDSN : 176
FS50   : NGYTDEYLSASIAPQVGQSSVHRADFLDRLIPLVPLTNVHFNKRVQEIEADEEQATIHFIDGQHITFDYVIGCDGIRSVVRNHVLDSN : 164
         NGYtDeYLSaSIaPQVGQSSVHRADFLDr66pl6PltnvhFNK46qeI2adeeqat6hFiDGqhitfDYVIGcDGIrSvvR1hVLd3n

          180        *       200        *       220        *       240        *       260
ADP1   : QFAPVRPNFTGTWAYRGIIKAAEFRQAIVAAGLDVEIADVPQMFLGQNKHILTFPIRQGEDINIVAFKTNPEQRTLPEHTPWTRAVDK : 261
AD3-1  : QLARVEPQFSGTWAYRGIIKFQDFKQAIAKLGNDAEIADVPQMLLGKDKHILTFPIRQGEEINIVAFKSDRTQTVLAENTPWTSPVSK : 264
FS50   : QLARVEPQFSGTWAYRGIIKFQDFKQAIAKLGNDAEIADVPQMLLGKDKHILTFPIRQGEEINIVAFKSDRTQTVLAENTPWTSPVSK : 252
         QlArVePqF3GTWAYRGIIKfqdF4QAIaklGnDaEIADVPQMlLGk1KHILTFPIRQGEeINIVAFK31rtQtvLaEnTPWTspVsK

             *       280        *       300        *       320        *       340        *
ADP1   : QEMLDDFQDWSESCRILLGLIERPTLWALHELAELPTYQSHSGRVILMGDAAHAMLPHQGAGAGQGLEDALTLKVLFEHTELTVEDLP : 349
AD3-1  : QTMLADFADWSESCKALLDLIDSPTIWALHEIQPLESYKNKSNNVILIGDAAHAMLPHQGAGAGQGLEDALILAKLLENPELTSESIK : 352
FS50   : QTMLADFADWSESCKALLDLIDSPTIWALHEIQPLETYKNKSNNVILIGDAAHAMLPHQGAGAGQGLEDALILAKLLENPELTSESIK : 340
         QtMLaDFaDWSESC4aLLdLIdsPT6WALHE6qpLe3YknkSnnVIL6GDAAHAMLPHQGAGAGQGLEDALiLakLlEnpELTsEs6k

             360        *       380        *       400        *       420        *
ADP1   : RVSAIYEQIRKERACKVQRTSRESGQIYELNSALYPSFEAVGAHLQNRFDWLWQHDLAQDMLAARAAIQPVATI----- : 423
AD3-1  : AVSEIYEQVRLARAVKVQNTSRESGEIYELYSSKYPDFESIGQHLAQRFDWLWQHSLEDDIQTAKAALAKQLQHTEMTT : 431
FS50   : AVSEIYEQVRLARAVKVQNTSRESGEIYELYSSKYPDFESIGQHLAQRFDW---------------------------- : 391
         aVSeIYEQ6RlaRAvKVQnTSRESG2IYELySskYPdFEs6GqHLaqRFDWlwqh l  d   a aa
```

Fig.7.2 Clustalx generated SalA amino-acid alignments from *Acinetobacter baylyi* sp. ADP1 and *Acinetobacter* sp. strains AD3-1 and FS50. The complete sequence of FS50 is truncated due to the position of the primers used.

strain ADPW259 in comparison with strains ADPW257, -258, -260 and -261. The amino acid alignments of ADP1, AD3-1 and FS50 are depicted in Fig.7.2 for full comparison and observation of conserved amino acid motifs and regions.

Highly conserved consensus sequences in the primary protein structure are observed in flavoprotein hydroxylases (Eppink *et al.*, 1997). It is possible that recombination events did occur with strain ADPW259 but had a deleterious effect on protein functionality and/or ability to form a correct tertiary or quaternary conformation resulting in a non-viable protein. This then would give the non-transformable phenotype observed with this strain in the salicylate specific selection procedure. The ADPW259 mutation occurs within a divergent DNA region bordered by blocks of homology (Fig.7.3A) and begins downstream from conserved residues FXQAIXXXG at ADP1 residue Leu209 corresponding to the AD3-1 residue Arg209 (Fig.7.2 and 7.3A). Five possible amino acid substitutions resulting from exchange with the heterologous donor DNA within this region could have occurred depending on the position of the cross-over event. Of these five possible substitutions, three involve conservative amino acids changes with two conferring amino acids with different properties (Fig.7.3A). A possible substitution at ADP1 SalA residue 206 changes alanine (Ala), a neutral, non-polar, hydrophobic residue to lysine (Lys), which is strongly basic, polar and hydrophilic. Substitution at the ADP1 residue 209 changes leucine (Leu), a non-polar, neutral, hydrophobic residue to asparagine (Asn), which is neutral, polar and hydrophilic. However, alterations in amino acid residue properties at other deletion sites did not apparently affect the transformation frequency or protein activity, although this may be due to the effect of a specific substitution, cross over event or structural residue interaction on the protein conformation.

A (ADPW259)

```
        600       *       620       *       640       *       660       *       680
ADP1  : GAATTTAGGCAAGCCATTGTCGCAGCCGGCCTAGATGTAGAAATTGCCGACGTACCACAAATGTTTTTAGGCCAAAACAAGCACA :  670
259   : GAATTTAGGCAAGCCATTGTCGCAGCCGGATCC----TAGAAATTGCCGACGTACCACAAATGTTTTTAGGCCAAAACAAGCACA :  648
AD3-1 : GATTTTAAGCAAGCTATTGCAAAACTGGGAAATGATGCCGAAATTGCCGATGTACCGCAAATGCTTTTAGGGAAAGATAAACATA :  679
FS50  : GATTTTAAGCAAGCCATTGCAAAACTGGGAAATGATGCCGAAATTGCCGATGTACCGCAAATGCTTTTAGGGAAAGATAAACATA :  644

        202aa                                                                   214
        GluAlaIleValAlaAlaGlyLeuAspValGluIleAla
ADP1    CAAGCCATTGTCGCAGCCGGCCTAGATGTAGAAATTGCC
        605bp                                                                   643
259     CAAGCCATTGTCGCAGCCGGATCC----TAGAAATTGCC

AD3-1   CAAGCTATTGCAAAACTGGGAAATGATGCCGAAATTGCC
        GluAlaIleAlaLysLeuGlyAsnAspAlaGluIleAla

FS50    CAAGCCATTGCAAAACTGGGAAATGATGCCGAAATTGCC
        GluAlaIleAlaLysLeuGlyAsnAspAlaGluIleAla
```

Fig.7.3A. Nucleotide alignment showing the DNA region surrounding the mutation site of ADPW259. Enlargement of the alignment includes the amino acid sequences, numbers above the enlarged sequence represent the amino acid (aa) residue directly below (relative to the alignment of Fig.7.2) and the base-pair (bp) region referred to (the first nucleotide to the last nucleotide of the enlarged sequence). The DNA sequence highlighted in grey show areas of DNA homology; in yellow followed by four consecutive dashes - the bases altered from the wild type sequence to introduce the *Bam*HI site and the position of the 4 bp deletion. Residues highlighted in blue indicate amino acid synonymous amino acid substitutions possible from recombination with donor DNA; highlighted in pink indicate amino acid substitutions possible that are non-synonymous.

B (ADPW260)

```
           860       *      880       *      900       *      920       *
ADP1  : TTGATTGAGCGTCCGACCTTGTGGGCACTGCACGAACTGGCCGAATTGCCGACTTATCAAAGCCACTCTGGCCGCGTCATATTAA :  925
260   : TTGATTGAGCGTCCGACCTTGTGGGCACGGATCC----GGCCGAATTGCCGACTTATCAAAGCCACTCTGGCCGCGTCATATTAA :  903
AD3-1 : CTGATCGACTCACCAACAATTTGGGCATTACATGAAATTCAACCTTTAGAAAGCTATAAAAACAAAAGCAATAATGTGATTTTGA :  934
FS50  : CTGATCGACTCACCAACAATTTGGGCATTACATGAAATTCAACCTTTAGAAACATATAAAAACAAAAGCAATAATGTGATTTTAA :  899
```

```
          289 aa                                        301
          ThrLeuTrpAlaLeuHisGluLeuAlaGluLeuProThr
ADP1      ACCTTGTGGGCACTGCACGAACTGGCCGAATTGCCGACT
          862bp                                        904
260       ACCTTGTGGGCACGGATCC----GGCCGAATTGCCGACT

AD3-1     ACAATTTGGGCATTACATGAAATTCAACCTTTAGAAAGC
          ThrIleTrpAlaLeuHisGluIleGlnProLeuGluSer

FS50      ACAATTTGGGCATTACATGAAATTCAACCTTTAGAAACA
          ThrIleTrpAlaLeuHisGluIleGlnProLeuGluThr
```

Fig.7.3B. Nucleotide alignment showing the DNA region surrounding the mutation site of ADPW260. Enlarged region shows the ADPW260 mutation region and protein sequence. Highlighted regions and representation of the deletion site are as described for Fig.7.3A.

C (ADPW258)

```
                 *       340         *       360         *       380
ADP1  : TTCGGTGCATCGTGCCGATTTCTTAGATGCTATTTTAGGGAATATTCCACAGCACCAGTGTA-A :  374
258   : TTCGGTGCATCGTGCGGATCC----GATGCTATTTTAGGGAATATTCCACAGCACCAGTGTA-A :  352
AD3-1 : TTCGGTACATCGTGCAGATTTTCTCGATCGCTTAATTCCGT-TAGTGCCTTTAACAAATGTACA :  383
FS50  : TTCGGTACATCGTGCAGATTTTCTCGATCGTTTAATTCCGT-TGGTGCCTTTAACAAATGTACA :  348

          111aa                                        123
          ArgAlaAspPheLeuAspAlaIleLeuGlyAsnIlePro
ADP1      CGTGCCGATTTCTTAGATGCTATTTTAGGGAATATTCCA
          331bp                                        369
258       CGTGCGGATCC- - - - GATGCTATTTTAGGGAATATTCCA

AD3-1     CGTGCAGATTTTCTCGATCGCTTAATTCCGT- TAGTGCC
          ArgAlaAspPheLeuAspArgLeuIlePro- LeuValPro

FS50      CGTGCAGATTTTCTCGATCGTTTAATTCCGT- TGGTGCC
          ArgAlaAspPheLeuAspArgLeuIlePro- LeuValPro
```

Fig.7.3C. Nucleotide alignment showing the DNA region surrounding the mutation site of ADPW258. Enlarged region shows the ADPW258 mutation region and protein sequence. Highlighted regions and representation of the deletion site are as described for Fig.7.3A.

D (ADPW261)

```
           *     1120       *     1140       *     1160       *     1180       *
ADP1  : GTTCAGCGTACTTCGCGTGAGTCTGGGCAAATATATGAACTTAACTCAGCACTGTATCCAAGCTTTGAAGCAGTGGGTGCACATT : 1180
261   : GTTCAGCGTACTTCGCGTGAGTCTGGGCAAAGGATCC----TAACTCAGCACTGTATCCAAGCTTTGAAGCAGTGGGTGCACATT : 1158
AD3-1 : GTTCAAAATACCTCGCGTGAATCAGGTGAAATCTATGAGCTATATTCATCGAAATATCCTGATTTTGAATCAATAGGCCAACATT : 1189
FS50  : GTTCAAAATACCTCGCGTGAATCAGGTGAAATCTATGAGCTATATTCATCGAAATATCCTGATTTTGAATCAATAGGCCAACATT : 1154

          375aa                                   397
          GluSerGlyGlnIleTyrGluLeuAsnSerAlaLeuTyr
ADP1      GAGTCTGGGCAAATATATGAACTTAACTCAGCACTGTAT
          1124bp                                  1162
261       GAGTCTGGGCAAAGGATCC----TAACTCAGCACTGTAT

AD3-1     GAATCAGGTGAAATCTATGAGCTATATTCATCGAAATAT
          GluSerGlyGlnIleTyrGluLeuTyrSerSerLysTyr

FS50      GAATCAGGTGAAATCTATGAGCTATATTCATCGAAATAT
          GluSerGlyGlnIleTyrGluLeuTyrSerSerLysTyr
```

Fig.7.3D. Nucleotide alignment showing the DNA region surrounding the mutation site of ADPW261. Enlarged region shows the ADPW261 mutation region and protein sequence. Highlighted regions and representation of the deletion site are as described for Fig.7.3A.

E (ADPW257)

```
          *      100       *      120       *      140       *      160       *
ADP1  : AACCACATTTAGAGGTTTGCTTGTATGAAGCAGCACCTCAATTTTCTGAAATTGGTGCCGGAATTTCATTTGGCGCGAATGCGGT :  161
257   : AACCACATTTAGAGGTTTGCTTGTATGAAGCAGCACCTCGGATCC----AATTGGTGCCGGAATTTCATTTGGCGCGAATGCGGT :  139
AD3-1 : ATAAACATTTAGATGTACAGCTTTTTGAGTCTGCATCACAGTTTTCTGAAATTGGTGCAGGGATTTCGTTTGGTGCCAATGCAGT :  170
FS50  : ATAAACATTTAGATGTACATCTCTTTGAGTCTGCATCACAGTTTTCTGAAATTGGTGCAGGGATTTCGTTTGGTGCGAATGCAGT :  135

          33aa                                                    45
          GluValCysLeuTyrGluAlaAlaProGlnPheSerGlu
ADP1      GAGGTTTGCTTGTATGAAGCAGCACCTCAATTTTCTGAA
          97bp                                                   135
257       GAGGTTTGCTTGTATGAAGCAGCACCTCGGATCC----A

AD3-1     GATGTACAGCTTTTTGAGTCTGCATCACAGTTTTCTGAA
          AspValGlnLeuPheGluSerAlaSerGlnPheSerGlu

FS50      GATGTACATCTCTTTGAGTCTGCATCACAGTTTTCTGAA
          AspValHisLeuPheGluSerAlaSerGlnPheSerGlu
```

Fig.7.3E. Nucleotide alignment showing the DNA region surrounding the mutation site of ADPW257. Enlarged region shows the ADPW257 mutation region and protein sequence. Highlighted regions and representation of the deletion site are as described for Fig.7.3A.

The ADPW260 deletion (Fig.7.3B) occurs in a region with less surrounding microhomology than in strain ADPW259 and six amino acid substitutions are possible. Three of these, ADP1 Ala210 to glutamine (Gln), glutamate211 (Glu) to proline (Pro) and Pro213 to Glu involve non-synonymous amino acid substitutions. The remaining substitutions confer conservative sequence changes. However, the transformability of this strain was not affected. In strain ADPW258 (Fig.7.3C) one possible non-conservative residue alteration may occur at ADP1 residue Ala117 to arginine (Arg). In ADPW261 (Fig.7.3D) Asn383 could be replaced with the AD3-1 tyrosine (Tyr) and Leu398 also could be changed to a lysine. For ADPW257, the deletion was incorporated within a highly conserved DNA region and substitutions that could occur in this region lie upstream of the deletion site (Fig.7.3E) involving three synonymous substitutions and two non-synonymous changes, Ala39 to serine (Ser) and Pro41 to Ser. With ADPW259, the direct substitution of leucine to asparagine occurs within the *Bam*HI site of the mutation, and so if recombination did occur in this strain it would have to incorporate this residue. With the exception of strain ADPW260 (Fig.7.3B), where a conservative substitution from leucine to isoleucine (Ile) would have to occur to repair the 4 bp deletion region, alterations in residue properties of the other strains, conservative or otherwise, do not occur directly within the mutation site but within the surrounding sequence. If the recombination events were restricted to very small regions of microhomology as suggested by the similar transformation frequencies with any of the three sources of heterologous AD3-1 donor DNA, then these other residue substitutions may not have been incorporated at all.

Eppink *et al.* (1997) identified two highly conserved flavoprotein hydroxylase FAD-binding domains based on three-dimensional crystal structure work and site-directed mutagenesis studies of the *P. fluorescens* prototype FAD-dependent hydroxylase *p*-hydroxybenzoate (PHBH), and alignments of 50

flavoprotein hydroxylases. The first conserved motif is the β-α-β-fold, containing the GXGXXG sequence (Fig.7.2 residues 14 - 19) important for binding the ADP moiety of FAD. This dinucleotide binding fold was first reported by Wierenga *et al.* (1986). The second motif occurs within PHBH at residues 153-166 and contains a highly conserved DG sequence. The conserved aspartate (D) residue of this sequence contacts the ribose moiety of FAD. This second FAD-binding motif occurs near the cleft leading toward the active site of PHBH. Residues 154-157 of PHBH at this conserved region are completely buried and multiple contacts are made with residues of the FAD-binding domain and a long excursion of the substrate-binding domain located downstream. The long excursion together with the FAD and interface domain may form one large globular domain and the contacts of residues 154-157 with residues in the region preceding the substrate binding domain may be important for maintaining the integrity of this catalytically important domain. In the ADP1 SalA, the conserved aspartate residue corresponding to the second FAD-binding motif occurs at residue 163 (relative to the alignment of Fig.7.2). Considering the highly conserved nature of FAD-dependent hydroxylases it is possible that the *Acinetobacter* SalAs form a similar conformation to the PHBH prototype. In which case, the ADPW259 deletion may occur within a similar long excursion domain (Fig.7.2 residues163 - 313) leading to the putative substrate binding region of SalA (Fig.7.2 residues 313 – 335). Deleterious amino acid substitutions resulting from homologous recombination at this position may affect residue contacts important in the formation of the catalytic domain region. However, the deletion of ADPW260, which interrupts the protein sequence at residue 295 (Fig.7.3B) would then also occur within the long excursion domain and possible substitutions in this region did not have a similar affect on the transformation frequency. Therefore, the leucine to asparagine substitution in ADPW259 may have occurred in a region important

for maintaining domain integrity or catalytic function leading to an inactive protein structure. However, the possible effects on protein structure and folding as an explanation of the ADPW259 transformation frequency levels does not account for the anomalous result observed with ADP1 genomic DNA which also resulted in a null transformation phenotype. This anomaly may indicate some other factor coming into play such as a coldspot for recombination.

Barriers to HGT by natural transformation between 'foreign' DNA and the recipient chromosome are provided by the homologous recombination system. Recombination events occur via DNA pairing, strand invasion and exchange at DNA regions that possess significant amounts of homology. Increasing sequence divergence between donor and recipient increases the likelihood of allelic replacements that are detrimental to protein expression (Thomas and Nielsen, 2005). A sequence divergence of ~40% is present between the *Acinetobacter salA*s used in these transformation studies and low frequencies of ADP1 transformation were observed with heterologous DNA from species classified as belonging to the same genus. Thus, the salicylate genes present within *Acinetobacter* have diverged sufficiently to prevent a high frequency of horizontal gene transfer but are still able to facilitate low level HGT by recombination events at localized homologous interactions within conserved DNA regions encoding the amino acids that constitute conserved protein domains.

The following are suggested further areas of work to follow up this study.

- Examine and quantify the ADPW257 – ADPW261 transformation frequencies resulting from natural transformation with other salicylate hydroxylase gene homologues from *Acinetobacter* spp. or another genus such as *Pseudomonas* spp. using the assay method outlined in this thesis.

- Perform conformational studies of the ADP1 SalA using point mutations to alter amino acid residues localized around the position of the ADPW259 deletion. This would then determine if the non-transformable phenotype of strain ADPW259 was due to the disruption of residue contacts essential for establishing the correct protein conformation.

- Investigate further the suggestion that the central deletion in *A. baylyi* strain ADPW259 occurs at or near a cold-spot of recombination. In a recent bioinformatic study of bacterial protein evolution indels (insertion/deletions events) were most frequently observed to occur at the N- or C- terminus (Pasek et al., 2006) which suggests that these regions are more amenable to recombination events. This could be achieved by adapting the *salA*$^-$ transformation assays to create ADP1 mutants with small deletions within the central regions of relevant genes and observe the effect on transformation frequencies using homologous and heterologous gene homologues as the donor DNA. These assays could be contrasted with the transformation frequencies of a second deletion inserted into the same genes at a different location (e.g. either at the 5'- or 3'- terminus) to determine whether the central regions of

genes are potentially 'cold' to recombination resulting in a non-transformable phenotype.

- In the transformation assays using donor DNA gene homologues that had 60% DNA identity to the wild type appeared to have undergone homologous recombination at small regions of micro-homology restoring *salA* gene function without replacing a significant amount of the wild-type DNA sequence. If these 'rescued' *salA* genes were disrupted in a second position to create a second-generation Δ*salA* mutant and re-transformed with the same heterologous DNA, what would be the transformation frequency? Would there be a significant increase in the frequency upon repeated repair and re-mutation events? Does this screen, albeit accelerated by the use of high quantities of a specific donor DNA, represent lateral gene transfer events that could occur in nature? The evolution of bacterial proteins is considered to result from elementary events that create new protein domain architectures categorized into three classes (Pasek *et al.*, 2006): domain insertion or deletion (indel), exchange and repetition. Could the exchange of small regions of DNA between homologues create a platform for more extensive genetic exchange in subsequent transformation assays and be a causative factor in the evolution and diversification of bacterial genomes?

- Investigate the mechanism of regulation of the *Acinetobacter* sp. strain AD3-1 salicylate genes. Determine if the genes *salA, -K,* and *-R* are co-transcribed as single RNA message or if *salR* is separately transcribed. LysR-type regulatory genes are generally divergently transcribed and upstream from the structural genes they control. Therefore binding to

promoter region enables them to regulate their own transcription in addition to that of the structural genes. As the AD3-1 *salR* lies downstream of the salicylate structural genes it would be interesting to determine whether SalR binds to the promoter region upstream of *salA* promoter region at a consensus LysR-type transcriptional regulator binding site and also to the intergenic region between *salK* and *salR*. Primer extension analysis to could be performed to determine the transcriptional start site of both *salA* and *salR*. Studies on SalR binding regions could be performed using gel-retardation DNA-protein binding assays and DNA foot-print analysis to determine which bases in the SalR binding regions are protected in DNA methylation studies. The presence of a stem-loop structure downstream from the *salK* gene may affect transcription attenuation to reduce the intramolecular levels of the regulator protein which may be toxic to the cell in high concentrations. To examine this more fully, the stem of the stem-loop structure could be disrupted by altering specific bases and observing the effect this has on either *salA* expression or *salR* transcription.

FS50 transformation frequencies

pFS50-1 donor DNA	5' position of the 4bp deletion [a]	Transformation frequency [b]	Log_{10} [c]	Fold reduction of transformation frequency
ADPW257	121-124	6.3 (± 0.9) x 10^{-7}	-6.2	0.45
ADPW258	332-335	1.4 (± 0.2) x 10^{-6}	-5.85	1
ADPW259	618-621	< 1.0 x 10^{-8}	<-8.0	>0.007
ADPW260	874-877	7.3 (± 0.87) x 10^{-7}	-6.14	0.5
ADPW261	1132-1135	1.65 (± 0.10) x 10^{-8}	-7.78	0.012

Table 8.1 Transformation frequencies of strains ADPW257-261 using donor DNA isolated from *Acinetobacter* sp. strain FS50.

[a] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[b] Mean transformation frequency ± Standard deviation (n=3).

[c] Log_{10} mean transformation frequency

[d] Fold reduction in pFS50-1 donor DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW258.

pFS50-pA donor DNA	5' position of the 4bp deletion [(a)]	Transformation frequency [(b)]	Log_{10} [(c)]	Fold reduction in transformation frequency
ADPW257	121-124	$4.0 (\pm 1.2) \times 10^{-7}$	-6.4	0.3
ADPW258	332-335	$1.5 (\pm 0.2) \times 10^{-6}$	-5.82	1
ADPW259	618-621	$< 1.0 \times 10^{-8}$	<-8.0	>0.007
ADPW260	874-877	$4.0 (\pm 0.5) \times 10^{-7}$	-6.39	0.3
ADPW261	1132-1135	$1.0 (\pm 0.08) \times 10^{-7}$	-6.97	0.07

Table 8.1.1. Transformation frequencies of strains ADPW257-261 using donor DNA pFS50-pA isolated from *Acinetobacter* sp. strain FS50.

[(a)] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[(b)] Mean transformation frequency (± Standard deviation) (n=3).

[(c)] Log_{10} mean transformation frequency.

[(d)] Fold reduction in donor pFS50-pA DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW258.

FS50 genomic donor DNA	5' position of the 4bp deletion [a]	Transformation frequency [b]	Log_{10} [c]	Fold reduction in transformation frequency
ADPW257	121-124	1.1 (± 0.06) x 10^{-7}	-6.95	0.26
ADPW258	332-335	4.3 (± 0.45) x 10^{-7}	-6.37	1
ADPW259	618-621	< 1.0 x 10^{-8}	<-8.0	>0.02
ADPW260	874-877	2.2 (± 0.2) x 10^{-7}	-6.66	0.5
ADPW261	1132-1135	1.0 (± 0.01) x 10^{-7}	-6.98	0.23

Table 8.1.2 Transformation frequencies of strains ADPW257-261 using genomic donor DNA isolated from *Acinetobacter* sp. strain FS50.

[a] The position of the deletion is given in bp relative to the GTG translational start site of the complete ADP1 *salA* gene.

[b] Mean transformation frequency (± Standard deviation) (n=3).

[c] Log_{10} mean transformation frequency.

[d] Fold reduction in FS50 genomic donor DNA transformation frequencies between the Δ*salA* mutant strains compared with strain ADPW258.

Amino acid alignments

MFS transport proteins

```
                        *        20         *        40         *        60         *        80
SalK        : ---------------MQF-INANEIIDHAKLKAIHWRVILLS-ALIIIFDGYDLVIYGVALPKLMAEWKIDSITAGFLGSVA :  65
BenKADP1    : ---------------MSRE-INVNQMIDDSKLTPFHWRVIILS-TLIIIFDGYDLVIYGVALPLLMKEWAIDPVTAGFIGSIA :  66
AzoBenT     : ---------------MRK-IDVHELIDKAKFNRFHWQVLFWC-ALIIIFDGYDLVIYGVVLPVLMKEWDLTPLQAGTLGSYA :  65
PpBenK      : ---------------MRT-LDVHPIIDNARFTPFHWMVMAWC-GLLLIFDGYDLFIYGVVLPVIMKEWGLTPLQAGALGSYA :  65
BenKADP1pu  : -------------MSTMS-LNINDLLNQAKFSKFHLSVVFWC-TLINICDGYDLVIYGVVLPRLMEQWSLTSVQAGMLASTA :  67
PcaKADP1    : MPKEANMASQDYATQRSSLDAQALINDAPLSRYQWLIAIVC-FLIVFVDGIITAAMGFIAPALAQDWGVDRSQLGPVMSAA :  80
PpPcaK      : ----MNQAQNSVG--KS-LDVQSFINQQPLSRYQWRVVLLC-FLIVFLDGLLTAAMGFIAPALSQEWGIDRASLGPVMSAA :  73
TfdK        : ---------------MNTTIDITDVIEKQSPNRTVGTLLVLS-FLIMLCDGYDLQAIAFASPTIIASWGIEKASFGFIFSAG :  66
HppK        : -------------------MDVREAIRTQPMRRYQISTIAIC-LILSMIDGFEILVMAFVAPHLGKSWDISSVEIGYLLSAG :  62
VanKADP1    : ---------------MQR--DILSEINQNNMSRYQWFVIVVC-VCLNIIDGFDVMVMAFTAPSVSAEWSLSGAQIGLLLSAG :  64
MucK        : ---------------------MYSNNQRSRIGSHTWKIAFLFAFLALLVDGADLMLLSYSLNSIKAEFNLSTVEAGMLGSFT :  61
LacY        : ---------------------MYYLKNTNF----WMFGLFF-FFYFFIMGAYFPFFPIWLHDIN---HISKSDTGIIFAAI :  52
                                                                dG d         p 6     6     G 6 s

                        *       100         *       120         *       140         *       160
SalK        : LFGMMFGALIFGSISDKLESYGFSRKKLIIICICLFSGFTLLCVYANNPQSFAIFRFLAGLGIGGVMPNVIAIMTEYAPKK : 146
BenKADP1    : LFGMMFGALIFGTIADKLEHLGVSKKKVIAVCIILFSLCTVLCGFSETTTQFSIFRFLAGVGIGGVMPNVIAIVSEYAPKK : 147
AzoBenT     : LFGMMFGALIFGPISDRFG-----RKRVIAICVMLFSFFTFINGFATTPTEFAICRFIAGLGIGGVMPNVVAIMTEYSPKK : 141
PpBenK      : LFGMMFGALAFGSLADRIG-----RKKGIAICFALFSGATILNGFASNPSEFGIYRFIAGLGCGGLMPNAVAIMNEYAPKR : 141
BenKADP1pu  : MCGMLVGALSCGILADRFG-----RKRVILICVFFLSTFTLLGSFASTPTEFAILRFLAGLGIGGAMPNVVAITSEYSPKK : 143
PcaKADP1    : LGGMIIGALVSGPTADRFG-----RKIVISISMLVFGGFTLACAYSTNLDSLVIFRFLTGIGLGAAMPNATTLFSEYCPAR : 156
PpPcaK      : LIGMVFGALGSGPIADRFG-----RKGVIVGAVLVFGGFSLASAYATNVDQLLVLRFLTGLGLGAGMPNATTLLSEYTPER : 149
TfdK        : LLGVMLGGFLFGYLADKIG-----RRPALLLGTFIFSAFTLAAVLATNVTELAALRFFAGLGIGGITPICFALNVEYVPKR : 142
HppK        : IIGTALGAIFISPLADKIG-----RRRLAIWCLVGITVGMALSACAVNVPQLIAFRAFAGLGIGGLVANLNVLVSEFSSDK : 138
VanKADP1    : LFGMAAGSIFLAPLADKIG-----RRLLILICLAIAGISMLACGLVHSHSMLAALRFVTGIGVGGILASSNVLASEYANLR : 140
MucK        : LAGMAIGGIFGGWACDRFG-----RVRIVVISILTFSILTCGLGLTQSFIQFGVLRFFASLGLGSLYIACNTLMAEYVPTK : 137
LacY        : SLFSLLFQPLFGLLSDKLG----LRKYLLWIITGMLVMFAPFFIFIFGP--LLQYNILVGSIVGGIY--LGFCFNAGAPAV : 125
                g   g    g   D4 g     R                                rf  g g G        l  e  p

                        *       180         *       200         *       220         *       240
SalK        : LRSTLVSLMFSGYAVGGMCSALLGMWLVPQFGWKIMFILGG-TPLLLIPIIWKLLPESIDYLVRRDK-TDKAIEIIKQIDG : 225
BenKADP1    : FKSFFVTLMFSGYAIGGMTAAFLGSILVPLYGWKIMFMIAG-IPLVLLLPIMEVLPESIDILVRKKK-DETVRFIMTKMVP : 226
AzoBenT     : LRTTLVAIMFSGYSVGGMLSAGLGMYLIPNFGWPSVFFVAG-VPLLLLPVILYALPESIGFLIHAKR-DAEAGVLIGRVEP : 220
PpBenK      : LRSTLVAIMFSGYSLGGMLSAGVGIFMIPRFGWESMFFAAA-VPLLLLPVILYYLPESIGFLVRQGR-IEEARTLIKRLDP : 220
BenKADP1pu  : SRVTMVSIMSCGYAVGAIISALVGSILVERFGWQIMFILAG-IPLLFIPIIWKYLPESLVPNVKKNN-TEKAHALIAKVDP : 222
PcaKADP1    : IRSLLVTCMFCGYNLGMAIGGFISSWLIPAFGWHSLFLLGGWAPLILMLVIFFLPESYRFLIVKGKNTKKVRQIISRIAP : 237
PpPcaK      : LKSLLVTSMFCGFNLGMAGGGFISAKMIPAYGWHSLLVIGGVLPLLLALVIMVWLPESARFLVVRNRGTDKIRKTISPIAP : 230
TfdK        : FHASVVGVVMIGYVIGISVGGLFAAWLVPVFGWQVLFYIGGLAPLVLLPILAIALPESIKYLVLAGRGGDTIARIVNAIDP : 223
HppK        : RRGTALGIYGVGYPLGGAVGGAISGVLISVFDWRSAFIFGTIITAIMLVVVIKTLPESIEFLIEKRP--VGALERYNEIAD : 217
VanKADP1    : WRSLAVSLMSTGYGIGATLGGILSLLLIEHLGWRSIFLAGGVTTVVMLLVSLWLLPESLDYLLAKKP--KQALEKINLTVQ : 219
MucK        : YRTTVLGTLQAGWTVGYIVATLLAGWLIPDHGWRVLFYVAI-IPVLMAVLMHFFVPE---------P----AAWQQSRIAP : 204
LacY        : EAFIEKVSRRSNPEFGRARMFGCVGWALCASIVGIMFTINN--QFVFWLGSGCALILAVLLFFAKTD--APSSATVANAVG : 202
                          g5  G           6   gw   f         6         6pes
```

```
                    *        260         *        280         *        300         *        320
SalK        : NITYTSQTQIGLHQDNQSS----SKTPIKDLFAENRGMVTLLFWGSVF-MALVLVYALGNWLPKLMV-EAGYDLSTSLVFL : 300
BenKADP1    : SYQYQPDHVFVLNSSNQNQ----AQAPVKMTFQEQRAFSTMMFWCSIF-MTLIMVYALGNWLPKLMI-EAGYNLSKSLIFL : 301
AzoBenT     : NFAARTGDRYEMPT--GKT----SHVSLVALFKDGRALSTLMFWLAFF-MCLLMVYALASWLPKLMT-QAGYGLGSSLSFL : 293
PpBenK      : DCDVQADDVLQAAD--RKG----SGASVLELFRNGLAIRTLALWVAFF-CCLLMVYALSSWLPKLMA-NAGYSLGSSLSFL : 293
BenKADP1pu  : TVLVSQHTKLVLNES-LEK----AYGSRRVLFQEGRAAGTLKFWIAFF-MCLLVVYALSSWLPKLMY-NAGYPLGKSILFL : 296
PcaKADP1    : -QKVQGVTEFHVPEEKVEAGT--KKGVFGMLFSAKYVKGTVLLWVTYF-MGLVMIYLLTSWLPTLMR-ETGASLERAAFLG : 313
PpPcaK      : -QVVAEAGSFSVPEQKAVA----ARSVFAVLFSGTYGLGTMLLWLTYF-MGLVIVYLLTSWLPTLMR-DSGASMEQAAFIG : 304
TfdK        : SAQANAQTRFVLEAEEARAKVGGWKESVKALFRGRLSRITPVLWLAFI-ANSVTVFFLASWIPVLTV-GTGRPATLAAIGL : 302
HppK        : RLGYEQSSELPVPIRKTER-----VV-RDGLFKGIMLRRTVFLWLGYS-CLIASFYFANTWTPKLLS-DATGDSNMGVTAG : 290
VanKADP1    : RIGLVRLNALPQHEKNTQQ----HSSDIAKLFQGKQGFQTLCLWFAFF-LVMFGFYFVMSWTPKLLI-AMGMSPSQGVTTG : 294
MucK        : --SKQTETVKTSAF---------------KLLFQDKRNRNMFILWALTAGFLQFGYYGVNNWMPSYLESELGMKFKEMTAYM : 269
LacY        : ---AN-HSAFSLKL--------------ALELFRQPKLWFLSLYVIGVSCTYDVFDQQFANLFTSFFA--TGEQGTR--VFG : 262
                                            6F        t   w                  5 p        g

                    *        340         *        360         *        380         *        400
SalK        : LALNIGGMLGA---ICGGYLADRFNLAKVLCSLFLSGALALILLSYNLP-TVILYFCIGVAGAASIGGQILLLAYMSQFYS : 377
BenKADP1    : FSLNVGGMIGS---ILGGYLADRYNVKFVTMGLLLLGALSLSLLSFQFS-SVILYILIACAGAASIGAQIMLLAYMAKFYA : 378
AzoBenT     : LVLNFGAIFGA---IGGGFLADRLHLKTVLMVFFVIAAFSISLLGFKSP-TAVLYLLVAAAGATTIGSQILLYAYVAQFYP : 370
PpBenK      : LALNFGGMAGA---ILGGWLGDRYNLVKVKVAFFIAAALSISLLGVNSP-MPVLYLLIFIAGATTIGTQILLYAGAAQMYG : 370
BenKADP1pu  : LAMNVGGIIGG---TTSGMLADKFKIKPVVISLFCIAALSLIALGFNTS-AIILYTLVAVAGASIQPASILLYSYVAQFYP : 373
PcaKADP1    : GLFQFGGVLSA---LFIGWAMDRFNPNRIIAGFYLAAGTFAVIVGQSLSNPTLLALFILCAGIAVNGAQSSMPVLSARFYP : 391
PpPcaK      : ALFQFGGVLSA---VGVGWAMDRYNPHKVIGIFYLLAGVFAYAVGQSLGNITVLATLVLIAGMCVNGAQSAMPSLAARFYP : 382
TfdK        : TLFSAGGAFGG---LFASRLIDRFGVG-AMGAVPLVAAVFVVAVGSMAFSDTGFLSMLFIVGFFVLGGHVGLMGSMGLFYA : 379
HppK        : VLVNVGGVIGS---IIFAGISIVLRPRVVTVLLMFGGVFAFIGYASGFHIVALALTLAVLVGIFANGGITGFYAISPPIYP : 368
VanKADP1    : ILISIGGVFGA---AIIGLLASRIKIFHALSLFLGLTAVCVWLFVAVSSQVSIALMVGLLLGTLINGCVAGLYSISPTIYD : 372
MucK        : VGTYTAMILGK---ILAGFMADKLGRRFTYAFGAIGTAIFLPLIVFYNS-PDNILYLLVIFGFLYGIPYGVNATYMTESFP : 346
LacY        : YVTTMGEILNASIMFFAPLIINRIGGKNALLLAGTIMSVRIIGSSFATS-ALEVVILKTLHMFEVPFLLVGCFKYITSQFE : 342
                g                                                                g    g        5

                    *        420         *        440         *        460         *        480
SalK        : SNIRATG-LGMALGVGRLGAILGPILCGWLLSLSLPINYNFIALSIPCITAVISVSMIHLQLKKTKLAVKALQI------- : 450
BenKADP1    : PNVRSTG-IGWGLGMGRVGAILGPILTGWLLSLQLPHFYNFLALSIPAVLGIVTVFLINDRRMYQPEPISPIANQNDTTTV : 458
AzoBenT     : MAIRSTG-IGWASGVGRSGAIVGPILGGTLLGMAMPLEYNFMFFAIPGALAAFAVSL-VARSAAHAPSEELPLPFAAETGA : 449
PpBenK      : LSVRSTG-IGWASGIGRNGAIVGPLLGGALMGINLPLQLNFIAFAIPGALAALAMAVHLASGRRHAQT--L----AAQA-- : 442
BenKADP1pu  : LVANSTG-IGWASAIGRTGAIMGPILAGMLLVLNLPHLMNFVLISVPALIAAISIGFLTTSKKKVEQPVALVTQQSLE--- : 450
PcaKADP1    : TQCRATG-VAWMSGIGRFGAVFGAWIGAVLIGNNWSFTMILSMLIIPARAAAIAIFVKSLVAHTDAT--------------- : 457
PpPcaK      : TQGRATG-VSWMLGIGRFGAIIGAWSGATLLGLGWNFEQVLTALLVPAALATVGVIVKGLVSHADAT--------------- : 448
TfdK        : NANRANG-VGWAISIGKLGSIIGPAIAGILIAANASIGTLFMAAAAPLLIAGAGVVLLGRLQKASDLAHATTAGHVADPLV : 459
HppK        : AAARGTG-VGWMIGCGRTVSIIAPILTGYLLDGGWTPKSVYMMFGGVLVIAGIATLLLDATYRGRSENPETPEARREEDLA : 448
VanKADP1    : ADIRSRG-VGYAIGFGRIGAILSPTVACIFLDKGLAPSTLYAYYGIVFILAIFLILSLSKAFYRQKEGLGSSIKTVA---- : 448
MucK        : TAIRGTA-IGGAYNVGRLGAIIAPATIG-FLASGGSIGLGFVVMGAAYFICGVIPALFIKEKQYDP--------------- : 410
LacY        : VRFSATIYLVCFCFFKQLAMIFMSVLAGNMYESIGFQGAYLVLGLVALGFTLISVFTLSGPGPLSLLRRQVNEVA------ : 417
                r tg 6        g  g        g

                *
SalK        : -------- :   -
BenKADP1    : KVNEAVSH : 466
AzoBenT     : ISD----- : 452
PpBenK      : -------- :   -
BenKADP1pu  : -------- :   -
PcaKADP1    : -------- :   -
PpPcaK      : -------- :   -
TfdK        : S------- : 460
HppK        : AAATA--- : 453
VanKADP1    : -------- :   -
MucK        : -------- :   -
LacY        : -------- :   -
```

LysR-type transcriptional regulators

```
                    *        20         *        40         *        60         *        80
AD3-1R   : ----LIDLSLIKVFIIVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLEERDKFGVKPTKLAQSLYPSFRQAITDIEMAL :  80
ADP1R    : -----VDLSLIRIFICVYENKNISKAAEILNLSQPSVTYNLNRLRKHLNNPLEERTQYGVEATKLSHELYPVFKESILKIEIAV :  79
PFlnahR  : MELHDLDLNLLVVFNQLLVDRRVSVTAENLGLTQPAVSNALKRLRTSLQDPLFVRTHQGMEPTPYAAHLAEPVTSAMHALRNAL :  84
PpG7NR   : MELRDLDLNLLVVFNQLLVDRRVSITAENLGLTQPAVSNALKRLRTSLQDPLFVRTHQGMEPTPYAAHLAEPVTSAMHALRNAL :  84
pTDG1R   : MELRDLDLNLLVVFNELLVCRRVSIVAENLGLTQPAVSNALKRLRTALQDQLFVRTYQGMEPTPYAANLAEPVALAMHALREAL :  84
AN10R    : MELHDLDLNLLVVFNQLMVDKRVSIVAQSLGLTQPAVSNALKRLRTALQDELFVRTHQGMEPTPYAAHLAEPIAHAMHSLREAL :  84
BenM     : -----MELRHLRYFVAVVEEQSFTKAADKLCIAQPPLSRQIQNLEEELGIQLLERGSRPVKTTPEGHFFYQYAIKLLSNVDQMV :  79
CatM     : -----MELRHLRYFVTVVEEQSISKAAEKLCIAQPPLSRQIQKLEEELGIQLFERGFRPAKVTEAGMFFYQHAVQILTHTAQAS :  79
tfdR     : -----MEFRQLRYFVAAAEEGNVGAAARRLHISQPPVTRQIQALEQHLGVLLEERGPRGVRLTPAGAAFLEDARRMLELGRTSV :  79
VanR     : ------------------------------MSSG--HQVLIKLRKMIIDGELEGGSR-IAEIPTAELLGVSRQPIRMAFRLLE :  50
                 l     f            a  l 6 qp     6  L   6   lf r       t

                *        100        *        120        *        140        *        160
AD3-1R   : TASKYFDPKTSTRIFRLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEDWLTEGFLDVAVFNSSYTVMPKIESRNLF : 164
ADP1R    : DEALNFNPLTSNKTFRIGLSDIGEICLLPTLIEYLRAHAPKIKIEVEEIKIDQVEKWLIEGFIDVAVFNSTHLEFKHLEYETLF : 163
PFlnahR  : QHHKSFDPLTSERTFTLAMTDIGEIYFMPRLMDVLAVQAPNCVISTVRDSSMSLMQALQNGTVDLAVGLLPNLQTGFFQRR-LL : 167
PpG7NR   : QHHESFDPLTSERTFTLAMTDIGEIYFMPRLMDVLAHQAPNCVISTVRDSSMSLMQALQNGTVDLAVGLLPNLQTGFFQRR-LL : 167
pTDG1R   : QHHERFDPLTSERTFTLAMTDIGEIYFMPRLMDAITLQAPNCAISTVRDSSMSLMQALQNGTVDLAVGLLPNLQTGFFQRR-LL : 167
AN10R    : HHEERFDPLTSERTFTLAMTDIGEIYFMPRLIDALTRKAPHCTISTVRGSSVSLGQSLLDGTVDLAVGLLPNLQAGFLQRR-LL : 167
BenM     : SMTKRIA--SVERTIRIGFVGSLLFGLLPRIIHLYRQAHPNLRIELYEMGTKAQTEALKEGRIDAGFGRLK-ISDPAIKRSLLR : 160
CatM     : SMAKRIA--TVSQTLRIGYVSSLLYGLLPEIIYLFRQQNPEIHIELIECGTKDQINALKQGKIDIGFGRLK-ITDPAIRRIVLH : 160
tfdR     : DRSRAAS-RGEIGQLDVGYLGTAIYQTVPALLHAFTQAVPGATLSLAQMSKVRQMEALRAGTIHVGVGRFY-PQETGITVEHLH : 161
VanR     : QEGLLIK--NPTRGYVVREISD---ELVHDALEVRGVLEGLAAKTIAEQGLTEQQKNILHGCIEETEKLFN-GRNEFGDEELEK : 128
                          6          6p  6       p                 6  G 6d                  l

                 *       180        *       200        *       220        *       240        *
AD3-1R   : EERYMCIAHKY--HPRIRDT---LSLEQYFQENHA---AIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVT : 240
ADP1R    : LERYVALVNMN--HPRIRST---LSFDAYLNESHV---AIKSSTGHTQVDHVLKLMGHQRKIALEVPHFGVLQGVLDKTDLMVT : 239
PFlnahR  : QNHYVCLCRKD--HPVTREP---LTLERFCSYGHVR--VIAAGTGHGEVDTYMTRVGIRRDIRLEVPHFAAVGHILQRTDLLAT : 244
PpG7NR   : QNHYVCLCRKD--HPVTREP---LTLERFCSYGHVR--VIAAGTGHGEVDTYMTRVGIRRDIRLEVPHFAAVGHILQRTDLLAT : 244
pTDG1R   : HNHYVCLCRKD--HPATREP---LTLERFCSYGHVR--VIAAGTGHGEVDTYLTKAGIRRDIRLEVPHFVAVGHILQRTELLAT : 244
AN10R    : HNRYVCLCRKN--HPATREP---LTLERFCAYSHIR--VIAASTGHGEVDSHMARAGIRRDIRLEVPHFVAVGHILQHTELLAT : 244
BenM     : NERLMVAVHAS--HPLNQMKDKGVHLNDLIDEKILLYPSSPKPNFSTHVMNIFSDHGLEPTKINEVREVQLALGLVAAGEGISL : 242
CatM     : KEQLKLAIHKH--HHLNQFAATGVHLSQIIDEPMLLYPVSQKPNFATFIQSLFTELGLVPSKLTEIREIQLALGLVAAGEGVCI : 242
tfdR     : YERLYIAASSSVARELSREP----TLLRLKSESLVLFPQEGRPSFADEVIALMRRAGVEPRVTAIVEDVNAALGLVAAGAGVTL : 241
VanR     : YHHYNVIFHDT--IIEGAKN---IALMQALAKNNQ---LPLASAQAITYDQNRALSEYRRLHYAHLQHCSIYNALVNRQAGR-- : 202
                        h           l                                g       e6        66
```

```
                 260         *       280         *       300         *       320         *
AD3-1R   : LPSRAAKVYARNGDVNIFELPFSVKPFNVSANWYNHRDDI------EARTWFIQTLQELFQSL---------------------- : 297
ADP1R    : LPSRAAQQYLNQSHVRVLELPFQMSEFYVGLHWFAQTNEP------LARIWLIQTCKKVISVL---------------------- : 296
PFlnahR  : VPIRLADCCVEPFGLSALPHPVVLPEIAINMFWHAKYHKD------LANIWLRQLMFELFSD----------------------- : 300
PpG7NR   : VPIRLADCCVEPFGLSALPHPVVLPEIAINMFWHAKVPQGPSQYLVAGNRCLTCLRIEKKSEPINVMGSDLIQPACALAGNGGS : 328
pTDG1R   : VPIRFADCCVEPFGLSVLPHPVALPEIAINMFWHAKYHQD------LANIWLRQLMFDLFSD----------------------- : 300
AN10R    : VPERFADCCVEPFGLSVLPIPIDLPEIPINMFWHAKYHKD------LANIWLRQLMFELFSD----------------------- : 300
BenM     : VPASTQSIQLFNLSYVPLLDPDAITPIYIAVRNMEESTYIYSLYETIRQIYAYEGFTEPPNW----------------------- : 304
CatM     : VPASAMDIGVKNLLYIPLLDDDAYSPISLAVRNMDHSNYIPKILACVQEVFATHHIRPLIE------------------------ : 303
tfdR     : VPASVAMLRRPFVRMIEMTAASAQVPVSLTYLTDSRVPMLRAFLDAARRLKPGAQAQRARPLG---------------------- : 304
VanR     : AENLMREHSSVFVTRRDCTSDFSGCPSKKLLYMGLMYRRK------------TELLSKIND------------------------ : 251
            p

            340         *       360         *       380
AD3-1R   : ---------------------------------------------- :   -
ADP1R    : ---------------------------------------------- :   -
PFlnahR  : ---------------------------------------------- :   -
PpG7NR   : SSTGSSRRRCCAAHQPSRRRRSARPKMTAPSVELFWFFAGSQDREN : 374
pTDG1R   : ---------------------------------------------- :   -
AN10R    : ---------------------------------------------- :   -
BenM     : ---------------------------------------------- :   -
CatM     : ---------------------------------------------- :   -
tfdR     : ---------------------------------------------- :   -
VanR     : ---------------------------------------------- :   -
```

Siderophore receptor proteins

```
                    *        20         *        40         *        60         *        80
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : ---MKLNMFCTTRLYRCVQLILGTGVTTFAVVAHATEQQTQQLSTITLEAEH-----VDSLY----NRKNANLDGFGVQDI :  69
PsyringaeB  : ------MRTPYRPAMRHAWRLLPLGMVVASPLLAAEVLSLDAVNVTGFSEP-----ESTGDY----KAERASMFGFDQASL :  66
PsyringaeD  : ------MSTPYRPAMHHAWRLLPLGMAVASPLLAAEPLNLDAVNVTGFSEQ-----ESASDY----RAGRASIYGLDQALL :  66
Xcampestri  : ------MYAPCPSLLSRRPLALALALCITLPALAQQQPDGAAVQLDAVN--------VRGER----ADAATTLSGFGATSL :  63
Bcenocepac  : MTVNRTTAAPARRRRAKVGLGVGLGVTLFAAQAAVAQGDTPAPGTEAAGNATLPAITVSGERGQPLRAREASVAGLDDAPL :  81
Bparapertu  : --------MPALFMLALGAVAAGARAQSAAAGVPDTQGVAQMPAVTVNAAP------VDDTLEH--LEAPVDTGALGRRTQ :  65

                    *       100         *       120         *       140         *       160
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : AKIDASISVISTERDADQAKVLTDVIKNDASLGDGYAAIGYDPN-PMSRGFGDALGSSYLMNGF-VIRGEQNDALDNKDR : 148
PsyringaeB  : LDTDASVSVFNAALDKDRAKLLSEVLRNDASVGDGYAPVGYDEN-FVVRGFSDNAANSYRINGR-SIAGEQNDALDNKQQ : 145
PsyringaeD  : LDTDASVSVFSEALDKDRAKLLSDVLRNDASVGEGYAPVGYDEN-FVVRGFSDNAANSYRINGR-SIAGEQNDALDNKQQ : 145
Xcampestri  : LDADASIAVIERQQLDRAARVLSEVLRADASAGDAYAPIGYDEN-FVLRGYSDNAANSYRINGL-SAVGEQSDALDNKQQ : 142
Bcenocepac  : RDTDASVNVITRALDDDQAKRLSDVVRNDASVVNDYAPVGYDEG-FAIRGFPDDLASAIRIDGL-TVSGEQNDPLDNKER : 160
Bparapertu  : LETDFSTTVVTARDDEERAVNKLGDVFALDASVTDNSDSYGADASYLTVRGLPDDWQNSYRIDGRPFLSYVTTDPFHFEQ : 146
                p s  v        q  l  v  das    a  g        rg           g           e

                    *       180         *       200         *       220         *       240
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : VEILKGISAIQSGMSTPGGVVNYVTKRP--EDIHAVSMGVNSEGGASIGLDMGGFAGSDQRLGYGVDLAQDQIHPDVEHAN : 227
PsyringaeB  : VELLKGLSGLQSGVAEPGGVINYQTKRA--QDVRSVTVSTNEHGERYIATDVGGWFGSEQQFGLGADLAHDDIRSDVEHAD : 224
PsyringaeD  : VELLKGLSGLQSGVSEPGGVINYQTKRA--QDVRSVTVSSNEHGERYIATDVGGWFGSEQQFGLGADLAHDDIRSDVEHAD : 224
Xcampestri  : VQILKGLAGLQSGVNEPAGLIDYRTKRP--EHVRSVTLGTDEQGSRYIAADLGDWFDADRTLGVGVDAARDDFRSDVEHAD : 221
Bcenocepac  : VEILKGLAGIDSGVVARGGVINPVTKRS--ANVASVTGGVDSRGSTSAAIDLGRRFGPDHQFGFGIDAAKDNMHSDIDGTN : 239
Bparapertu  : DDLLKGASGFMYGFGSPGGLVNYVTKKPTDEAVRSVELGYVSKGLLREHVDLGRVGESGAFGYGLDATHDEGNTDNGGS- : 226
                lkg      g   p g    tk         v      g     d g        g r n   e   y

                    *       260         *       280         *       300         *       320
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : GKGLFGEVALDLQLDEKSQIQFDIESGRQKQRSVPGYQLLDGTTVPTDVTWDGLLGYQSWSKPVTNTSLN--SSLKYNDQI : 306
PsyringaeB  : GQGDFASLAFDWNISERALLQLDVEYGSKEQRSVPGYQLLGGTTLPHDASPGGLLGYQNWSNPVGIDSLN--INGRFEDRF : 303
PsyringaeD  : GQGDFASLAFDWNISERALLQLDVEYGNKEQRSVPGYQLLGGTTLPHDASPRGLLGYQNWSNPVGIESLN--MSGRFEDRF : 303
Xcampestri  : GYGNFLSLAGDWKPTAQSLLQVDVEYGHRQQRSVPGYQLLGGTQVPRNIDVHGLLGYQPWSRPVEMDSLN--AQLRYTDQF : 300
Bcenocepac  : GRGTFGSIAADWDISPRASLQLNAEFGQWIQRSAPGYQLLGGTVVPSVKTTSGALGTQPWAKPVTTDALN--LNARFDDQF : 318
Bparapertu  : LYGDSVSLALDARLSDRLTWDFQSIYGDRKAIGQEPTIYAGTMAGSELPSPVGNDNDRLVGQGPYADNAFRYYSTGLKDQL : 307
                r   s a d              q                                                    y

                    *       340         *       360         *       380         *       400
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : NSDGKASLGAAYGHVKIDDYIAFPGGSY--KGLKPIEN---TFDENGNYDIYDYRNPDDIRITTQFKSTLVGQLQGEDVH : 382
PsyringaeB  : NDNGKGSLSASRGRVVIDDYSAFAGGCYGSASCAGQAV-PNHFSAEGGYDIYDFRSPDDTRRNDEVEAAMSGTFTPGSDGH : 383
PsyringaeD  : TDSGKGSLSASRGRVVIDDYSAFAGGCYGSASCATEAV-PNHFSAEGGYDIYDFRSPDDTRRNDEVEAAMSGTFAPGADEH : 383
Xcampestri  : NEVGRASVEASRGRSTINDFSAFAGGCYGAASCADIAT-PNFFSAQGEYDVYDYRSPDDTRVHDQLRASLDGRVTPGADDH : 380
Bcenocepac  : NDDGKAYVAVGRGRTMIDDNSAFAGGCSYAASCAAGATSPFFFGANGDYDVYDFRSPGEYRRNDDLRAVTTGKFAPGPDRH : 399
Bparapertu  : ADEGTLSTNYSYGSTRTRRNESVLGLRD--------------QAG--DYDDYRSDYGEAYGYN-QWQAMLEGKFAPGPDKH : 371
                w        s                                  yd y                  g   tg l h

                    *       420         *       440         *       460         *       480
AD3-1       : ------------------------------------------------------------------------------ :  -
ADP1        : HGNFELAKTLKSQKRYVEVNIPGTGNIDQETNDFDPAAEDLGNRYKPLESDGVSFALSGRIEFNPQGSTLLGGWIHLNE : 463
PsyringaeB  : EGTFGSSAYRRTVDTRGTFNEFGSGNIDEAPEQVAPSTQALPHTERRLDSRGYGLFATGLDSLNEHGQTVLGGGVRLDE : 464
PsyringaeD  : QGTFGTSAYRRTVDTRGTFNEFGSGNINESPEPVAPSSLPLTHTERRLDSRGYGLFVSGRISFNDRGQAVLGGGEVRLDE : 464
Xcampestri  : QGSIGADWLRRTIDRYGSVNEFGSGSIDRDPEVFAQTDVALDPKQRRLDSRGRALVVAGRDGFGSQGELSLGAGYVRLDE : 461
Bcenocepac  : EGTLGVSVQRRVVHMADAVYDYGSENIYGPDLTFPPSPNSPGPSYPRLDAWGYGVFGLGRDSIGEHGQVLAGGEVLLRQ : 480
Bparapertu  : HGVAGASWQKQKNDYSANGVYQGQGTGNLRARNTNTYYSEGQLHLYRAAEITGKALFASGTGDLTGGGSVLGGLRYTNYAQ : 452
                                                                  q     d       w   g
```

```
                 *       500         *       520         *       540         *       560
AD3-1       : ------------------------------------------------------------------------------ :   -
ADP1        : RAYDAESI-ENRHTQIDKFLEQLALMYSPFESTNVVVSYAKGLADEATAPWYTSNSNEVLAEVNSTRYEICVKQQYKN-AL : 542
PsyringaeB  : QAENEDGS-DARHTERYVFLEQAALIYKPVDNVSLVTSYSKGLSLEGTAAWFTTNASEIDAETVSRALEAGIRYDWQR-MS : 543
PsyringaeD  : RAENEDGS-DARHTERYVFLEQAALIYKPLDNISLVTSYSKGLSLEGTAAWFTTNASEIDAETVSRALEAGIRYDWQR-MS : 543
Xcampestri  : RAEDRDGV-LERRTRKDALLEQAALLEKPQQQVTVVASYAKSLAAEGTAPWFADVADEIDTETNAYALETEAKWQLQG-VR : 540
Bcenocepac  : RSTSSLDG-ETTHTDRSVFLEQVALVYKPVNVLSLVASYSKALSLEDQAPVRATVAYAFDPEVESHRIEVCARYDWLDRLS : 560
Bparapertu  : QGEDATGARTSRYDKNGVLTETFALMYKLTPRTMAVASYIESLEPESSVGATYAVFGALDDLKSKRYELCIKTEQDG-WA : 532
                              p  al          y sy   l  g        n    l p    q e g k

                 *       580         *       600         *       620         *       640
AD3-1       : -------MAQPLMLIKP----------IIIQQMVRKAVSQIQP-VCKPLDMTSSLALTRSRLKEIQVADEQGHCTQNVESV :  63
ADP1        : FTLALPDLKQDNQYVRPNDDGT-STFVSGEKQHNQGIELNVQGRITPHLDLISSLALLRSQWEDSGVDIEENHCTQNVEKL : 622
PsyringaeB  : LTAAVPQARQAYQYSRPNDDGT-FTYVQQEEQKNTGLELGASGWVTDRLQMSASAAAARARVEDSGTEAEDDHCALNVERY : 623
PsyringaeD  : LTAALPQARQAYQYSRPNDDGT-FTYVQQEEQKNTGLELGARGWVTDRLQMSASAAAARARVEGSGTEAEEDHCALNVEKY : 623
Xcampestri  : LGAAVPDIRQAYQFTQPQADGG-LLYVQQERLHNRGLELSADGAVTGQLQLFASVAARRARAEDTGIAAEEGHCAINVEKL : 620
Bcenocepac  : LTAAVPSISKPFQPADPDASGTSYTFVQRETQRHQGIELGAAGRVTERLGLTASVAARRARAVDSGSPAEEGHCIINVEAL : 641
Bparapertu  : ATAALPRIEKKAEVANAAN-----ELVQDEKTLYQGLELGASTRLARDWNVGGSLMLVDSEYKKG--SDETGNRVAGAEKF : 606
                 a f          p         v  g     g el     6   l 6  s a            5  hq  nvP

                 *       660         *       680         *       700         *       720
AD3-1       : PFASHLSERVIQVDGERLLAGIQYSSSFYANKTGTVKVSGYSVFDVGTAANFRAYGYDNTLRFNVDNVFNKTVRDAGSPL : 144
ADP1        : PFATHASQVAQLEGERLLGGIQYSASFYANKEGTVKVAGYWVFDLGAAARFNLYGYDTTMRLSVDNLFNKTVRDVGGPM : 703
PsyringaeB  : PGTLQADESLI-VPGLALLGGVQYSASFYADREGTVQVNDYAFFNLGSRVSTRLSGYDTVLRLTVDNLFDKTVRDAGEML : 703
PsyringaeD  : PATLQADESLI-IRGLALLGGIQYSASFYADRTGSVQVNDYAFFNLGSRVSTKLGGYDTVLRLTVDNLFDKTVRDAGEML : 703
Xcampestri  : PASVQADERVLGEDGLALLAGVQYSGRNFADRLGSASVPAYAFANLGTRVATTPGGLATTWRLNVDNVFDKTVRDAGEVQ : 701
Bcenocepac  : PASLYADEAVLGWAGLNVLGGVEYSAARNANEEGTARVPSVFVFNLGARVTTKPGGHRTVLRVSVDNLFNKTVRDAGEQQ : 722
Bparapertu  : VAAAQLAESVIQVPGLKLRADQKYGGNTMLGASNRVQVDDYAHVVFGATLDTQPHGYEATFTAGFNNVANKTVR----LVQ : 683
              r      Y 6p 6 GL 61 g6 Y3    a   g   V  5 6 16G  Y     G     r  61N6f1K YWrd g

                 *       740         *
AD3-1       : GDDYLELGAPRTAQLSWTVGF : 165
ADP1        : GDDYLELGAPRTAQFSINFKF : 724
PsyringaeB  : GDDYLFMGAPRTARLSATVNF : 724
PsyringaeD  : GDDYLFMGAPRTASLSASVNF : 724
Xcampestri  : GDAYLFPGAPRTARLDVQVDF : 722
Bcenocepac  : GDAYLFPGAPRTARVSLTYDF : 743
Bparapertu  : SSDYVKAEDPRTYGLDASLKF : 704
              gd Y6f GaPRTa  3    F
```

SalR and upstream sequence homology between pFS50-1 and pAD3-1A

```
                    *        20         *        40         *        60         *        80
pAD3-1A : RLGAILGPILCGWLLSLSLPINYNFIALSIPCIIAVISVSMIHLQLKKTKLAVKALQIKSREVYLPHSIEILQLIDLSLIKVFI :  84
pFS50-1 : --------------------------ALSIPCIIAVMSVSMIHLRLKKTKLAVKAVQIKSREVYLPHLIEILQLIDLSLIKVFI :  58
                                     ALSIPCIIAV6SVSMIHL LKKTKLAVKA6QIKSREVYLPH IEILQLIDLSLIKVFI

                 *       100         *       120         *       140         *       160
pAD3-1A : IVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLFERDKFGVKPTKLAQSLYPSFRQAITDIEMAITASKYFDPKTSTRIF : 168
pFS50-1 : IVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLFERDKFGVKPTKLAQSLYPSFRQAITDIEMAITASKYFDPKTSTRIF : 142
          IVFETKNISHAAERLNLSQPSVTYNLNRLRKFLNDPLFERDKFGVKPTKLAQSLYPSFRQAITDIEMAITASKYFDPKTSTRIF

           *       180         *       200         *       220         *       240         *
pAD3-1A : RLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEDWLTEGFLDVAVFNSSYTVMPKIESRNLFEERYMCIAHKYHPRI : 252
pFS50-1 : RLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVEEWLTEGFLDVAVFNSSYTVMPKIESRNLFEERYMCIAHKYHPRI : 226
          RLGLSDLGEICLLPSLIQYLSKEAPFIKVEIEEIQSHKVE WLTEGFLDVAVFNSSYTVMPKIESRNLFEERYMCIAHKYHPRI

                260         *       280         *       300         *       320         *
pAD3-1A : RDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAKVYARNGDVNIFELPFS : 336
pFS50-1 : RDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAKVYARNGDVNIFELPFS : 310
          RDTLSLEQYFQENHAAIKSSTGHIEIDQRLKELGLKRKIMLELPHFSVLQDVVTSSELLVTLPSRAAKVYARNGDVNIFELPFS

           340         *       360         *       380         *       400         *       420
pAD3-1A : VKPFNVSANWYNHRDDIEARTWFIQTLQELFQSLRRIFICNVLKKQIEPLITALYLNENFKSALRIKNLFILSKIEYHSNHRIS : 420
pFS50-1 : VKPFNVSANWYNHRDDIEARTWFIQTLQELFQSLHSILISNAKKTVN------------------------------------ : 357
          VKPFNVSANWYNHRDDIEARTWFIQTLQELFQSL  I I N  K

                    *       440         *       460         *       480         *       500
pAD3-1A : DLFFFLISELHEITARIRSWHFGTITKRLYLGDWRYCRSYWIKYSMLATHSNGRRRHHSRSCFIRSKTKYQCDCHGFYNSTTCR : 504
pFS50-1 : ------------------------------------------------------------------------------------ :   -

                 *       520         *       540         *
pAD3-1A : MDKISQSYPYNRSYRSVSNGWRDLFVKSRCTRCAVQSLKNNSNFWSYCEL : 554
pFS50-1 : ------------------------------------------------- :   -
```

The pAD3-1A SalK terminates at residue Ile58 leading directly into the putative leader sequence of SalR which show virtually identical sequence with pFS50-1. The pAD3-1A SalR terminates at residue Leu370 in an identical position to pFS50-1.

Aber, W., 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol. Rev. 24: 1-7.

Aldrich, T. L., and A. M. Chakrabarty, 1987. Cloning and complete nucleotide sequence determination of the *cat*B gene encoding *cis, cis*-muconate lactonizing enzyme. Gene 52:165-169.

Allen, C. C. R., D. R. Boyd, M. J. Larkin, K. A. Reid, N. D. Sharma, and K. Wilson, 1997. Metabolism of naphthalene, 1-naphthol, indene, and indole by *Rhodococcus* sp. strain NCIMB 12038. Appl. Environ. Microbiol. 63: 151-155.

Ausubel, S. F., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, (Ed), 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N. Y.

Averhoff, B., 2004. DNA transport and natural transformation in mesophilic and thermophilic bacteria. J. Bioenerg. Biomem. 36: 25-33.

Bale, M. J., M. J. Day, and J. C. Fry, 1988. Novel method for studying plasmid transfer in undisturbed river epilithon. Appl. Environ. Microbiol. 54: 2756-2758.

Barbe, V., D. Vallenet, N. Fonknechten, A. Kreimeyer, S. Oztas, L. Labarre, S. Cruveiller, C. Robert, S. Duprat, P. Wincker, L. N. Ornston, J. Weissenbach, P. Marliere, G. N. Cohen, and C. Medigue, 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. Nucleic. Acids Res. 32: 5766-5779.

Baumann, P.,1968. Isolation of *Acinetobacter* from soil and water. J. Bacteriol. 95: 520-541.

Baur, B., K. Hanelmann, W. Sclimme, and B. Jenni, 1996. Genetic transformation in freshwater: *Escherichia coli* is able to develop natural competence. Appl. Environ. Microbiol. 62: 3673-3678.

Bayly, R. C., and M. C. Barbour, 1984. The degradation of aromatic compounds by the meta and gentisate pathways. Biochemistry and regulation. In: Microbial degradation of aromatic hydrocarbons, Ch. 8 p.253-294, Microbiology series Vol 13. Ed. Gibson, D.T.

Bergmans, H., 1993. Acceptibility of the use of antibiotic resistance genes as marker genes in transgenic plants. In: OECD report on the scientific approaches for the assessment of research trials with genetically modified plants. p.106-108, April 6-7 1992. Jouy-en-Josas.

Bertolla, F., F. van Gijsegem, X. Nesme, and P. Simonet, 1997. Conditions for natural transformation of *Rastonia solanacearum*. Appl. Environ. Microbiol. 63: 4965-4968.

Bianco, P. R., and S. C. Kowalczykowski, 1997. The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as a single strand of DNA containing the sequence 5'-GCTGGTGG-3'. Proc. Natl. Acad. Sci. USA 94: 6706-11.

Biswas, G. D., T. Sox, and P. F. Sparling, 1977. Factors affecting genetic transformation of *Neisseria gonorrhoeae*. J. Bacteriol. 129: 983-992.

Biswas I., E. Maguin, S. D. Ehrlich, and A. Gruss, 1995. A 7-base-pair sequence protects DNA from exonucleolytic degradation in *Lactococcus lactis*. Proc. Natl. Acad. Sci. USA 92: 2244-2248.

Bjorklof, K., A. Suoniemi, K. Haahtela, and M. Romantschuk, 1995. High frequency of conjugation versus plasmid segregation of RP1 in epiphytic *Pseudomonas syringae* populations. Microbiol. 141: 2719-2727.

Boonchan, S., M. L. Britz, and G. A. Stanley, 1998. Surfactant-enhanced biodegradation of high molecular weight polycyclic aromatic hydrocarbons by *Stenotrophomonas maltophilia*. Biotechnol. Bioeng. 59: 482-494.

Bosch, R., E. R. B. Moore, E. Garcia-Valdes, and D. H. Pieper, 1999. NahW, a novel, inducible salicylate hydroxylase involved in mineralization of naphthalene by *Pseudomonas stutzeri* strain AN10. J. Bacteriol. 181: 2315-2322.

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

Bundy, B. M., L. S. Collier, T. R. Hoover, and E. L. Neidle, 2002. Synergistic transcriptional activation by one regulatory protein in response to two metabolites. Proc. Natl. Acad. Sci. USA 99: 7693-7698.

Busch, S., C. Rosenplanter, 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. Environ. Microbiol. 65: 4568-4574.

Cane, P. A. and P. A. Williams, 1982. The plasmid-encoded metabolism of naphthalene and 2-methylnaphthalene in *Pseudomonas* strain: phenotypic changes correlated with structural modification of the plasmid pWW60-1. J. Gen. Microbiol. 128: 2281-2290.

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.

Carr, E. L., P. Kampfer, B.K. Patel, V. Gurtler, and R.J. Seviour, 2003. Seven novel species of *Acinetobacter* isolated from activated sludge. Int. J. Syst. Evol. Microbiol. 53: 953-963.

Ceccherini, T., J. Pote, E. Kay, V. T. Van, J. Marechal, G. Pietramellara, P. Nannipieri, T. M. Vogel, and P. Simonet, 2003. Degradation and transformability of DNA from transgenic leaves. Appl. Environ. Microbiol. 69: 673-678.

Cerniglia, C. E., 1992. Biodegradation of polyaromatic hydrocarbons. Biodegradation 3: 351-368.

Chamier, B., M. G. Lorenz, and W. Wackernagel, 1993. Natural transformation of *Acinetobacter calcoaceticus* by plasmid DNA adsorbed onto sand and groundwater aquifer material. Appl. Environ. Microbiol. 59: 1662-1667.

Chaki, R. V. J., C. P. Whitman, and J. W. Kozarich, 1987. Absolute stereochemical course of muconolactone D-isomerase and of 4-carboxymuconolactone decarboxylase: a ^{1}H NMR "richochet" analysis. J. Amer. Chem. Soc. 109: 5514-19.

Chedin, F., P. Noirot, V. Biaudet, and S. D. Ehrlich, 1998. A five-nucleotide sequence protects DNA from exonucleolytic degradation by AddAB, the RecBCD analogue of *Bacillus subtilis*. Mol. Microbiol. 29: 1369-1377.

Clark, T. J., C. Momany, and E. L. Neidle, 2002. The *benPK* operon, proposed to play a role in transport, is part of a regulon for benzoate catabolism in *Acinetobacter* sp. strain ADP1. Microbiol. 148: 1213-1223.

Clark, T. J., R. S. Philips, B. M. Bundy, C. Momany, and E. L. Neidle, 2004. Benzoate decreases the binding of *cis*, *cis*-muconate to the BenM regulator despite the synergistic effect of both compounds on transcriptional activation. J. Bacteriol. 186: 1200-1204.

Claverys, J-P., M. Prudhomme, I. Mortier-Barriere, and B. Martin, 2000. Adaptation to the environment: *Streptococcus pneumoniae,* a paradigm for recombination-mediated genetic plasticity? Mol. Microbiol. 35: 251-259.

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

Collier, L. S., G. S. 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.

Copeland, A., S. Lucas, A. Lapidus, K. Barry, C. Detter, T. Glavina, N. Hammon, S. Israni, S. Pitluck, and P. Richardson, 2005. Sequencing of the draft genome assembly of *Burkholderia cenocepacia* HI2424. Genbank. Unpublished.

Cornish-Bowden, A., 1976. Principles of enzyme kinetics. London; Boston: Butterworths. ISBN 0408707216.

Cosper, N. J., L. S. Collier, T. J. Clark, R. A. Scott, and E. L. Neidle, 2000. Mutations in *cat*B, the gene encoding muconate cycloisomersae, activate transcription of the distal *ben* genes and contribute to a complex regulatory circuit in *Acinetobacter* sp. strain ADP1. J. Bacteriol. 182: 7044-7052.

Courthesy-Thelulaz, I. E., G. E. Bergonzelli, H. Henry, D. Bachmann, D. D. Schorderet, A. L. Blum, and L. N. Ornston, 1997. Cloning and characterisation of *Helicobacter pylori* succinyl-CoA: acetoacetate CoA transferase, a novel prokaryotic member of the CoA transferase family. J. Biol. Chem. 272: 25659-25667.

Daane, L. L., J. A. E. Molina, E. C. Berry, and M. J. Sadowsky, 1996. Influence of earthworm activity on gene transfer from *Pseudomonas fluorescens* to indigenous soil bacteria. Appl. Environ. Microbiol. 62: 515-521.

D'Argenio, D. A., M. W. Vetting, D. H. Ohlendorf, and L. N. Ornston, 1999. Substitution, insertion, deletion, suppression, and altered substrate specificity in functional protocatechuate 3,4-dioxygenases. J. Bacteriol. 181: 6478-6487.

Davidson, J., 1999. Review: Genetic exchange between bacteria in the environment. Plasmid 42: 73-91.

Davies, J. I. and W. C. Evans, 1964. Oxidative metabolism of naphthalene by soil pseudomonads. J. Bacteriol. 91: 251-261.

De Baere, T., L. Dijkshoorn, A. Nemec, M. Vaneechoutte, L. N. Ornston, and D. M. Young, 2004. The naturally hyper-transformable strain BD4 belongs to the newly described species *Acinetobacter baylyi*. 6th International symposium on the biology of *Acinetobacter*. 15-17 September, Dublin, Ireland.

Demaneche, S., E. Kay, F. Gourbiere, and P. Simonet, 2001b. Natural transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in soil. Appl. Environ. Microbiol. 67: 2617-2621.

Demaneche, S., L. Jocteur-Monrozier, H. Quiquampoix, and P. Simonet, 2001a. Evaluation of biological and physical protection against nuclease degradation of clay-bound plasmid DNA. Appl. Environ. Microbiol. 67: 293-299.

Demaneche, S., C. Meyer, J. Micoud, M. Louwagie, J. C. Willison, and Y. Jouanneau, 2004. Identification and functional analysis of two aromatic-ring-hydroxylating dioxygenases from a *Sphingomonas* strain that degrades various polycyclic aromatic hydrocarbons. Appl. Environ. Microbiol. 70: 6714-6725.

Denome, S. A., D. C. Stanley, E.S. Olson, and K. D. Young, 1993. Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. J. Bacteriol. 175: 6890-6901.

de Vries, .J., and W. Wackernagel, 1998. Detection of *nptII* (kanamycin resistance) genes in genomes of transgenic plants by marker-rescue transformation. Mol. Gen. Genet. 257: 606-613.

de Vries, J., P. Meier, and W. Wackernagel, 2001. The natural transformation of the soil bacteria *Pseudomonas stuzeri* and *Acinetobacter* sp. by transgenic plant DNA strictly depends on homologous sequences in the recipient cells. FEMS Microbiol. Letts. 195: 211-215.

de Vries, J., and W. Wackernagel, 2002. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. Proc. Natl. Acad. Sci. USA 99: 2094-2099.

de Vries, J., M. Heine, K. Harms, and W. Wackernagel, 2003. Spread of recombinant DNA by roots and pollen of transgenic tobacco plants, identified by highly specific biomonitoring using natural transformation of an *Acinetobacter* sp. Appl. Environ. Microbiol. 69: 4455-4462.

de Vries, J., T. Herzfeld, and W. Wackernagel, 2004. Transfer of plastid tobacco to the soil bacterium *Acinetobacter* sp. by natural transformation. Mol. Microbiol. 53: 323-334.

de Vries, J., and W. Wackernagel, 2005. Microbial horizontal gene transfer and the DNA release from transgenic crop plants. Plant Soil 266: 91-104.

DiMarco, A. A., B. Averhoff, and L. N. Ornston, 1993. Identification of the transcriptional activator, *pob*R and characterisation of its role in the expression of *pob*A, the structural gene for p-hydroxybenzoate hydroxylase in *Acinetobacter calcoaceticus.* J. Bacteriol. 175: 4499-4506.

Dodender, R., 1966. Levansucrase from *Bacillus subtilis*. Methods Enzymol. 8: 500-505.

Dorn, E. and H. J. Knackmuss, 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Two catechol-1,2-dioxygenases from a 3-chlorobenzoate-grown pseudomonad. Biochem. J. 174: 73-84.

Doten, R. C., L. A. Gregg-Jolly, and L. N. Ornston, 1987a. Influence of the *cat*BCE sequence on the phenotypic reversion of a *pca*E mutation in *Acinetobacter calcoaceticus.* J. Bacteriol. 169: 3175-3180.

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

Droge, M., A. Puhler, and W. Selbitschka, 1998 Horizontal gene transfer as a biosafety issue: A natural phenomenon of public concern. J. Biotechnol. 64: 75-90.

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

Dubnau, D., 1999. DNA uptake in bacteria. Annu. Rev. Microbiol. 53: 217-244.

Duggan, P. S., P. A. Chambers, J. Heritage, and J. M. Forbes, 2000. Survival of free DNA encoding antibiotic resistance from transgenic maize and the transformation activity of DNA in ovine saliva, ovine rumen fluid and silage effluent. FEMS Microbiol. Lett. 191: 71-77.

Eaton, R. W. and P. J. Chapman, 1992. Bacterial metabolism of naphthalene: construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions. J. Bacteriol. 174: 7542-7554.

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

Eppink, M. H. M., H. A. Schreuder, and W. J. H. van Berkel, 1997. Identification of a novel conserved sequence motif in flavoprotein hydroxylases with a putative dual function in FAD/NAD(P)H binding. Protein Sci. 6: 2454-2458.

Feng, Y., H. E. Khoo, and C. L. Poh, 1999. Purification and characterization of gentisate 1,2-dioxygenases from *Pseudomonas alcaligenes* NCIB 9867 and *Pseudomonas putida* NCIB 9869. Appl. Environ. Microbiol. 65: 946-950.

Finkel, S. E., and R. Kolter, 2001. DNA as a nutrient: novel role for bacterial competence gene homologues. J. Bacteriol. 183: 6288–6293.

Friedrich, A., T. Hartsch, and B. Averhoff, 2001. Natural transformation in mesophilic and thermophilic bacteria: identification and characterisation of novel, closely related competence genes in *Acinetobacter* sp. strain BD413 and *Thermus thermophilus* HB27. Appl. Environ. Microbiol. 67: 3140-3148.

Fuenmayor, S. L., M. Wild, A. L. Boyes, and P. A. Williams, 1999. A gene cluster encoding steps in conversion of naphthalene to gentisate in *Pseudomonas* sp. strain U2. J. Bacteriol. 180: 2522-2530.

Fukuyama, K., T. Hase, S. Matsumoto, T. Tsukihara, and Y. Katsube, 1980. Structure of *S. platensis* [2Fe-2S] ferredoxin and evolution of the chloroplast type ferredoxins. Nature 286: 522-524.

Gay, P., D. Le Coq, M. Steinmetz, T. Berkelman, and C. I. Kado, 1985. Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. J. Bacteriol. 164: 918-921.

Gebhard, F., and K. Smalla, 1998. Transformation of *Acinetobacter* sp. strain BD413 by transgenic sugar beet DNA. Appl. Environ. Microbiol. 64: 1550-1554.

Gebhard, F., and K. Smalla, 1999. Monitoring field releases of genetically modified sugar beets for persistence of transgenic plant DNA and horizontal gene trasnfer. FEMS Microbiol. Ecol. 28: 261-272.

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

Ghisla, S., and V. Massey, 1989. Mechanisms of flavoprotein-catalyzed reactions. Eur. J. Biochem. 181: 1-17.

Gibson, D. T., and V. Subramanian, 1984, In: Microbial degradation of Organic compounds,In: Microbial degradation of aromatic hydrocarbons, Ch. 7 p. 181-252, Microbiology series Vol 13. Ed. Gibson, D.T.

Gjuracic, K., E. Pivetti, and C. V. Bruschi, 2004. Targeted DNA integration within different functional gene domains in yeast reveals ORF sequences as recombinational cold-spots. Mol. genet. genomics 271 (4): 437-446.

Gogarten, J. P., and J. P. Townsend, 2005. Horizontal gene transfer, genomic innovation and evolution. Nature Revs. Microbiol. 3: 679-687.

Goodgal, S. H., 1982. DNA uptake in *Haemophilus* transformation. Annu. Rev. Genet. 16: 169-192.

Goodman, S. D., and J. J. Socca, 1988. Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. Proc. Natl. Acad. Sci. USA 85: 6982-6986.

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. Microbiol. 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.

Grund, E., B. Denecke, and R. Eichenlaub, 1992. Naphthalene degradation via salicylate and gentisate by *Rhodococcus* sp. Strain B4. Appl. Environ. Microbiol. 58: 1874-1877.

Gubriel, R. J., C. J. Batie, M. Sivaraja, A. E. True, and J. A. Fee, 1989. Electron nuclear double resonance spectroscopy of N-15 enriched phthalate dioxygenase from *Pseudomonas cepacia* proves that 2 histidines are coordinated to the [2Fe-2S] Rieske-type clusters. Biochemistry 28: 4861-4871.

Hagedorn, S. R., G. Bradley, and P. J. Chapman, 1985. Glutathione-independent isomerization of maleylpyruvate by *Bacillus megaterium* and other Gram-positive bacteria. J. Bacteriol. 163: 640-647.

Hahn, J., M. Albano, and D. Dunau, 1987. Isolation and characterisation of Tn*917lac* generated competence mutants of *Bacillus subtilis*. J. Bacteriol. 169: 3104-3109.

Haigler, B. E. and D. T. Gibson, 1990. Purification and properties of ferrsdoxin$_{NAP}$, a component of naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816. J. Bacteriol. 172: 465-468.

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

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

Harayama, S. and M. Kok, and E. L. Niedle, 1992. Functional and evolutionary relationships among diverse oxygenases. Annu. Rev. Microbiol. 46: 565-601.

Hartnett, C., E. L. Neidle, K-L. Ngai, and L. N. Ornston, 1990. DNA sequences of genes encoding *Acinetobacter calcoaceticus* protocatechuate 3,4-dioxygenase: evidence indicating shuffling of genes and of DNA sequences within genes during their evolutionary divergence. J. Bacteriol. 172: 956-966.

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

Hendrickx, L., M. Hausner, and S. Wuertz, 2003. Natural genetic trasnformation in monoculture *Acinetobacter* sp. strain BD413 biofilms. Appl. Environ. Microbiol. 1721-1727.

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

Herrick, J. B., K. G. Stuart-Keil, W. C. Ghiorse, and E. L. Madsen, 1997. Natural horizontal transfer of a naphthalene dioxygenase gene between bacteria native to a coal tar-contaminated field site. Appl. Environ. Microbiol. 63: 2330-2337.

Herzberg, C., A. Friedrich, and B. Averhoff, 2000. *com*B a novel competence gene required for natural transformation of *Acinetobacter* sp. strain BD413: identification, characterisation, and analysis of growth-phase-dependent regulation. Arch. Microbiol.

Hickey, W. J., G. Sabat, A. S. Yuroff, A. R. Arment, and J. Perez-Lesher, 2001. Cloning, nucleotide sequencing and functional analysis of a novel, mobile cluster of biodegradation genes from *Pseudomonas aeruginosa* strain JB2. Appl. Environ. Microbiol. 67: 4603-4609

Hobbs, M., and J. S. Mattick, 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein secretion apparatus: a general system for the formation of surface associated protein complexes. Mol. Microbiol. 10: 233-243.

Hughes, E. J., M. K. Shapiro, J. E. Houghton, and L. N. Ornston, 1988. Cloning and expression of *pca* genes from *Pseudomonas putida* in *Esherichia coli.* J. Gen. Microbiol. 134: 2877-2887.

Igual J. C., C. Gonzalez-Bosch, J. Dopazo, and J. E. Perez-Ortin, 1992. Phylogenetic analysis of the thiolase family. Implications for the evolutionary origin of peroxisomes. J. Mol. Evol. 35: 147-155.

Inamine, G. S., and D. Dubnau, 1995. ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. J. Bacteriol. 177: 3045-3051.

Ishiyama, D., D. Vujaklija, and J. Davies, 2004. Novel pathway of salicylate degradation by *Streptomyces* sp. strain WA46. Appl. Environ. Microbiol. 70: 1297-1306.

Israel, D. A., A. S. Lou, and M. J. Blaser, 2000. Characteristics of *Helicobacter pylori* natural transformation. FEMS Microbiol. Lett. 186: 275-280.

Kay, E., T. M. Vogel, F. Bertolla, R. Nalin, and P. Simonet, 2002. In situ transfer of antibiotic resistance genes from transgenic bacteria (transplastomic) tobacco plants to bacteria. Appl. Environ. Microbiol. 68: 3345-3351.

Johnson, B. G., and R. Y. Stainer, 1971. Dissimilation of aromatic compounds by *Alcaligenes eutrophus*. J. Bacteriol. 107: 468-475.

Jones, D. C., and R. A. Cooper, 1990. Catabolism of 3-hydroxybenzoate by the gentisate pathway in *Klebsiella pneumoniae* M5al. Arch. Microbiol. 154: 489-495.

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

Jones, R. 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.

Jones, R. M., and P. A. Williams, 2001. *are*CBA is an operon in *Acinetobacter* sp. strain ADP1 and is controlled by AreR, a σ^{54} -dependent regulator. J. Bacteriol. 183: 405-409.

Jones, R. M., B. Britt-Compton, and P. A. Williams, 2003. The naphthalene catabolic (*nag*) genes of *Ralstonia* sp. strain U2 are an operon that is regulated by NahR, a LysR-type transcriptional regulator. J. Bacteriol. 185: 5847-5853.

Jones, R. M., and P. A. Williams, 2003a. Mutational analysis of the critical bases involved in activation of the AreR-regulated σ^{54}-dependent promoter in *Acinetobacter* sp. strain ADP1. Appl. Environ. Microbiol. 69: 5627-5635.

Juhasz, A. L., M. L. Britz, and G. A. Stanley, 1997. Degradation of fluoranthene, pyrene, benz[a]anthracene and dibenz[a,h]anthracene by *Burkholderia cepacia*. J. Appl. Microbiol. 83: 189-198.

Juni, E., and A. Janik, 1969. Transformation of *Acinetobacter calcoaceticus* (*Bacterium anitratum*). J. Bacteriol. 98: 281-288.

Juni, E., 1972. Interspecies transformation of Acinetobacter: genetic evidence for a ubiquitous genus. J Bacteriol. 112: 917-31.

Juni, E., 1978. Genetics and physiology of *Acinetobacter*. Annu. Rev. Microbiol. 32: 349-71.

Karudapuram, S., and G. J. Barcak, 1997. The *Haemophilis influenzae dpr*ABC genes constitute a competence inducible operon that requires the product of the *tfo*X (*sxy*) gene for transcriptional activation. J. Bacteriol. 179: 4815-4820.

Kaschabek, S. R., B. Kuhn, D. Muller, E. Schmidt, and W. Reineke, 2002. Degradation of aromatics and chloroaromatics by *Pseudomonas* sp. Strain B13: Purification and Characterisation of 3-oxoadipate: succinyl-Coenzyme A (CoA) transferase and 3-oxoadipyl-CoA thiolase. J. Bacteriol. 184: 207-215.

Katagiri, M., S. Takemori, K. Suzuki, and H. Yasuda, 1966. Mechanism of the salicylate hydroxylase reaction. J. Biol. Chem. 241: 5675-5677.

Katti, S. K., B. A. Katz, and H. W. Wyckoff, 1989. Crystal structure of muconolactone isomerase at 3.3 Å resolution. J. Mol. Biol. 205: 557-571.

Kernsten, P. J., S. Dagly, J. W. Whittaker, D. M. Arciero, and J. D. Lipscomb, 1982. 2-Pyrone-4,6-dicarboxylic acid, a catabolite of gallic acids in *Pseudomonas* species. J. Bacteriol. 152: 1154-1162.

Khan, M. E., G. Maul, and S. H. Goodgal, 1982. Possible mechanism for donor DNA binding and transport in *Haemophilus*. Proc. Natl. Acad. Sci. USA. 79: 6370-6374.

Khan, M. E., and H. O. Smith, 1984. Transformation in Haemophilus: a problem in membrane biology. J. Membr. Biol. 81: 89-103.

Kiyohara, H., S. Torigoe, N. Kaida, T. Asaki, T. Iida, H. Hayashi, and N. Takizawa, 1994. Cloning and characterisation of a chromosomal gene cluster, *pah*, that encodes the upper pathway for phenanthrene and naphthalene utilisation by *Pseudomonas putida* OUS82. J. Bacteriol. 176: 2439-2443.

Kloos, D-U., M. Stratz, A. Guttler, R. J. Steffan, and K. N. Timmis, 1994. Inducible cell lysis system for the study of natural transformation and environmental fate of DNA released by cell death. J. Bacteriol. 176: 7352-7361.

Kok, R. G., D. D'Argenio, and L. N. Ornston, 1998. Mutation analysis of PobR and PcaU, closely related transcriptional activators in *Acinetobacter*. J. Bacteriol. 180: 5058-5069.

Kok, R. G., D. M. Young, and L. N. Ornston, 1999. Phenotypic expression of PCR-generated random mutation in a *Pseudomonas putida* gene after its introduction into an *Acinetobacter* chromosome by natural transformation. Appl. Environ. Microbiol. 65: 1675-1680.

Koonin, E. V., K. S. Makarova, and L. Aravind, 2001. Horizontal gene transfer in prokaryotes: Quantification and classification. Annu. Rev. Microbiol. 55: 709-742.

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 the *Acinetobacter alcoaceticus pca* operon. Gene 146: 23-30.

Koyama, Y., T. Hoshino, N. Tomizuka, and K. Furukawa, 1986. Genetic transformation of the extreme thermophilic *Thermus thermophilus* and of other *Thermus* spp. J. Bacteriol. 166: 338-340.

Lack, L., 1959. The enzymic oxidation of gentisic acid. Biochim. Biophys. Acta 34: 117-123

Lee, H., N. Vázquez-Laslop, K. A. Klyachko, and A. A. Neyfakh, 2003. Isolation of antibiotic hypersusceptible mutants of *Acinetobacter* spp. by selection for DNA release. Antimicrob. Agents Chemother. 47: 1267-1274.

Li, W., J. Shi, X. Wang, Y. Han, W. Tong, L. Ma, B. Liu, and B. Cai, 2004. Complete nucleotide sequence and organization of the naphthalene catabolic plasmid pND6-1 from *Pseudomonas* sp. strain ND6. Gene 336: 231-240.

Lilley, A. K., J. C. Fry, M. J. Day, and M. J. Bailey, 1994. *In situ* transfer of an exogenously isolated plasmid between *Pseudomonas* spp. in sugar beet rhizosphere. Microbiol. 140: 27-33.

Link, C., S. Eickernjager, D. Porstendorfer, and B. Averhoff, 1998. Identification and characterisation of a novel competence gene, *com*C, required for DNA binding and uptake in *Acinetobacter* sp. strain BD413. J. Bacteriol. 180: 1592-1595.

Lorenz, M. G., B. W. Aardema, and W. Wackernagel, 1988. Highly efficient transformation of *Bacillus subtilis* attached t sand grains. J.Gen Microbiol. 134: 107-112.

Lorenz, M. G., and W. Wackernagel, 1990. Natural genetic transformation of *Pseudomonas stutzeri* by sand adsorbed DNA. Arch. Microbiol. 154: 380-385.

Lorenz, M. G., K. Reipschlager, and W. Wackernagel, 1992. Plasmid transformation of naturally competent *Acinetobacter calcoaceticus* in non-sterile soil extract and groundwater. Arch. Microbiol. 157: 355-360.

Lorenz, M. G., and W. Wackernagel, 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. 58: 563-602.

Martin, P.R.., M. Hobbs, P. D. Free, Y. Jeske, and J. S. Mattick, 1993. Characterisation of *pil*Q, a new gene required for the biogenesis of type IV fimbriae in *Pseudomonas aeruginosa*. Mol. Microbiol. 9: 857-868.

Masai, E., K. Momose, H. Hara, S. Nishikawa, H. Katayama, and M. Fukuda, 2000. Genetic and biochemical characterisation of 4-carboxy-2-hydroxy-muconate-6-semialdehyde dehydrogenase and its role in the protocatechuate 4,5-cleavage pathway in *Sphingomonas paucimobilis*. J. Bacteriol. 182: 6651-6658.

Mason, J. R., and R. Cammack, 1992. The electron-transport proteins of hydroxylating bacterial dioxygenases. Annu. Rev. Microbiol. 46: 277- 305.

Meier, P., and W. Wackernagel, 2003. Mechanisms of homology-facilitated illegitimate recombination for foreign DNA acquisition in transformable *Pseudomonas stutzeri*. Mol. Microbiol. 48: 1107-1118.

Melnikov, A., and P. J. Youngman, 1999. Random mutagenesis by recombinational capture of PCR products in *Bacillus subtilis* and *Acinetobacter calcoaceticus*. Nucleic Acids Res. 27: 1056-1062.

Menn, F. -M., B. M. Applegate, and G. S. Sayler, 1993. NAH plasmid-mediated catabolism of anthracene and phenanthrene to naphthoic acids. Appl. Environ. Microbiol. 59: 1938-1942.

Metzgar, D., J. M. Bacher, V. Pezo, J. Reader, V. Doring, P. Schimmel, P. Marliere, and V. de Crecy-Lagard, 2004. *Acinetobacter* sp. ADP1: an ideal model organism for genetic analysis and genome engineering. Nucleic Acids Res. 32: 5780-5790.

Miller, M. A., and J. D. Lipscomb, 1996. Homoprotocatechuate 2,3-Di-oxygenase from *Brevibacterium fuscum*. J. Biol. Chem. 271: 5524 - 5535.

Muller, R., 1992. Bacterial degradation of xenobiotics, Ch.3. p.35-59. In: Microbial control of pollution, Society of General Microbiology symposium 48. (Eds. Fry, J.C., Gadd, G.M., Herbert, R.A., Jones, C.W. and I. A. Watson-Craik).

Nadalig, T., N. Raymond, N. M. Gilewicz, H. Budzinski, and J. C. Bertrand, 2002. Degradation of phenanthrene, methylphenanthrenes and dibenzothiophene by a *Sphingomonas* strain 2mpII. Appl. Environ. Microbiol. 59: 79-85.

Nakai, C., H. Uyeyama, H. Kagamiyama, T. Nakazawa, and S. Inouye, 1995. Cloning, DNA sequencing, and amino acid sequencing of catechol 1,2-dioxygenases (pyrocatechase) from *Pseudomonas putida* mt-2 and *Pseudomonas arvilla* C-1. Arch. Biochem. Biophys. 321: 353-362.

Neidle, E. L., and L. N. Ornston, 1986. Cloning and expression of *Acinetobacter calcoaceticus*catechol 1,2-Dioxygenase structural gene *cat*A in *Escherichia coli*. J. Bacteriol. 168: 815-820.

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, S. Bointz, and L. N. Ornston, 1988. DNA sequence of the *Acinetobacter calcoaceticus* catechol-1,2-dioxygenase I structural gene *cat*A: evidence for evolutionary divergence of intradiol dioxygenases by acquistion of DNA repetitions. J. Bacteriol., 170: 4874-4880.

Nielsen, J. W., K. L. Joesphson, I. L. Pepper, R. B. Arnold, G. D. DiGiovanni, and N. A. Sinclair, 1994. Frequency of horizontal gene transfer of a large catabolic plasmid (pJP4) in soil. Appl. Environ. Microbiol. 60: 4053-4058.

Nielsen, K. M., A. M. Bones, and J. D. van Elsas, 1997. Induced natural transformation of *Acinetobacter calcoaceticus* in soil microcosms. Appl. Environ. Microbiol. 63: 3972-3977.

Nielsen, K. M., A. M. Bones, K. Smalla, and J. D. van Elsas, 1998. Horizontal gene transfer from transgenic plants to bacteria - a rare event? FEMS Microbiol. Rev. 22: 79-103.

Nielsen, K. M., K. Smalla, and J. D. van Elsas, 2000a. Natural transformation of *Acinetobacter* sp. strain BD413 with lysates of *Acinetobatcer* sp., *Pseudomonas fluorescens,* and *burkholderia cepacia* in soil microcosms. Appl. Environ. Microbiol. 66: 206-212.

Nielsen, K. M., J. D. van Elsas and K. Smalla, 2000b. Transformation of *Acinetobacter* sp. strain BD413 (pFG4 *npt*II) with transgenic plant DNA in soil microcosms and effects of kanamycin on selection of transformants. Appl. Environ. Microbiol. 66: 1237-1242.

Nielsen, K. M., and J. D. van Elsas, 2001. Stimulatory effects of compounds present in the rhizosphere on natural transformation of *Acinetobacter* sp. BD413 in soil. Soil Biol. Biochem. 33: 345-357.

Noda, Y., S. Nishikawa, K-I. Shiozuka, H. Kadokura, H. Nakajima, K. Yoda, Y. Katayama, N. Morohoshi, T. Haraguchi, and M. Yamasaki, 1990. Molecular cloning of the protocatechuate 4,5-dioxygenase genes of *Pseudomonas paucimobilis.* J. Bacteriol. 172: 2704-2709.

Ochman, H., J. G. Lawrence, and E. A. Groisman, 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405: 299-304.

Ogram, A., G. S. Sayler, D. Gustin, and R. J. Lewis, 1988. DNA adsorption to soils and sediments. Environ. Sci. Technol. 22: 982-984.

Palmen, R., B. Vosman, P. Buijsman, C. K. D. Breek, and K. J. Hellingwerf, 1993. Physiological characterisation of natural transformation in *Acinetobacter calcoaceticus*. J. Gen. Microbiol. 139: 295-305.

Palmen, R., and K. J. Hellingwerf, 1997. Uptake and processing of DNA by *Acinetobacter calcoaceticus* - a review. Gene 192: 179-190.

Parales, R. E., S. Lee, S. M. Resnick, H. Jiang, D. J. Lessner, and D. T. Gibson, 2000. Substrate specificity of naphthalene dioxygenase: Effect of specific amino acids at the site of the enzyme. J. Bacteriol. 182: 1641-1649.

Parke, D., D. 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.

Parke, D., M.A.Garcia, and L. N. Ornston, 2001. Cloning and genetic characterization of *dca* genes required for β-oxidation of straight-chain Dicarboxylic acids in *Acinetobacter* sp. strain ADP1. Appl. Eniviron. Microbiol. 67: 4817-4827.

Pasek, S., J-L. Risler, P. Brezellec, 2006. Gene fusion / fission is a major contributor to evolution of multi domain bacterial proteins. Bioinformatics Advance Access. In press.

Paul, J. H., Frischer, M. E., and Thurmond, J. M., 1991. Gene transfer in marine water column and sediment microcosms by natural plasmid transformation. Appl. Environ. Microbiol. 57: 1509-1515.

Pinyakong, O., H. Habe, and T. Omori, 2003. The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons (PAHs). J. Gen. Appl. Microbiol. 49: 1-19.

Poh, C. L., and R. C. Bayly 1988. Regulation of isofunctional enzymes in *Pseudomonas alcaligenes* mutants defective in the gentisate pathway. J. Appl. Bacteriol. 64:451-458.

Popp, R., T. Kohl, P. Patz, G. Trautwein, and U. Gerischer, 2002. Differential DNA binding of transcriptional regulator PcaU from *Acinetobacter* sp. strain ADP1. J. Bacteriol. 184: 1988-1997.

Porstendorfer, D., U. Drotschmann, and B. Averhoff, 1997. A novel competence gene, *com*P, is essential for natural transformation of *Acinetobacter* sp. strain BD413. App. Environ. Microbiol. 63: 4150-4157.

Providenti, M. A., J. Mampel, S. MacSween, A. M. Cook, and R. C. Wyndham 2001. *Comamonas testosteroni* BR6020 possesses a single genetic locus for extradiol cleavage of protocatechuate. Microbiol. 147: 2157-2167.

Que, L. Jr., 1985. Spectroscopic studies of the catechol dioxygenases. J. Chem. Educ. 62: 938-943.

Rauch, P. J. G., R. Palmen, A. A. Burds, L. A. Gregg-Jolly, J. R. van der Zee, and K. J. Hellingwerf, 1996. The expression of the *Acinetobacter calcoaceticus recA* gene increases in response to DNA damage independently of RecA and of development of competence for natural transformation. Microbiol. 142: 1025-1032.

Reams, A. B., and E. L. Neidle, 2003. Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. Mol. Microbiol. 47: 1291-1304.

Ribbons, D. W., P. Keyser, D. A. Kunz, and B. F. Taylor, 1984. Microbial degradation of phthalates. In: Gibson, D. T. (Ed.), Microbial degradation of organic compounds. Marcel Dekker, New York.

Richaume, A., E. Smit, G. Faurie, and J. D. van Elsas, 1992. Influence of soil type on the transfer of plasmid RP4 from *Pseudomonas fluorescens* to introduced recipient and indigenous bacteria. FEMS Microbiol. Ecol. 10: 281-292.

Romanowski, G., M. G. Lorenz, and W. Wackernagel, 1993. Plasmid DNA in a groundwater aquifer microcosm-adsorption Dnase resistance and natural genetic transformation of *Bacillus subtilis*. Mol. Ecol. 2: 171-181.

Romero-Arroyo, C. E., M. A. Schell, G. L. Gaines III, E. L. Neidle, 1995. *cat*M encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. J. Bacteriol. 177: 5891-5898.

Sala-Trepat, J-M., K. Murray, and P. A. Williams, 1972. The metabolic divergence in the *meta* cleavage of catechols by *Pseudomonas putida* NCIB 10015. Physiological significance and evolutionary implications. Eur. J. Biochem. 28: 347-356.

Sandaa, R-A., and O. Enger, 1994. Transfer in marine sediments of the naturally occurring plasmid pRSA1 encoding multiple antibiotic resistance. Appl. Environ. Microbiol. 60: 4234-4238.

Saye, D. J., O. A. Ogunseitan, G. S. Sayler, and R. V. Miller, 1990. Transduction of linked chromosomal genes between *Pseudomonas aeruginosa* strains during incubation *in situ* in a freshwater habitat. Appl. Environ. Microbiol. 56: 140-145.

Schell, M. A., and P. Wender, 1986. Identification of the *nahR* gene product and nucleotide sequences required for its activation of the *sal* operon. J. Bacteriol. 166: 9.14.

Schell, M. A., 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. 47: 597-626.

Schlomann, M., 1994. Evolution of chlorocatechol catabolic pathways. Biodegradation 5: 301-321.

Schneider, J., R. Grosser, K. Jayasimhulu, W. Xue, and D. Warshawsky, 1996. Degradation of pyrene, benz[a]anthracene, and benzo[a]pyrene by *Mycobacterium sp.* strain RJGII-135, isolated from a former coal gasification site. Appl. Environ. Microbiol. 62: 13-19.

Segura, A., P. V. Bunz, D. A. D'Argenio, and L. N. Ornston, 1999. Genetic analysis of a chromosomal region containing *van*A and *van*B, genes required for conversion of either ferulate or vanillate to protocatechuate in *Acinetobacter*. J. Bacteriol. 181: 3494-3504.

Sexton, J. A., and J. P. Vogel, 2004. Regulation of Hypercompetence in *Legionella pneumophila*. J. Bacteriol. 186: 3814-3825.

Shanley, M. S., E. L. Neidle, R. E. Parales, and L. N. Ornston, 1986. Cloning and expression of *Acinetobacter calcoaceticus cat*BCDE genes in *Pseudomonas putida* and *Escherichia coli*. J. Bacteriol. 165: 557-563.

Smalla, K., S. Borin, H. Heuer, F. Gebhard, J. D. van Elsas, and K. M. Nielsen, 2000. Horizontal transfer of antibiotic resistance genes from transgenic plants to bactria - are there data to fuel the debate? In: The proceedings of the 6th international symposium on The biosafety of genetically modified organisms. P: 146-154. Saskatoo, Fairbairn, C., Scoles, G., McHughen, A., University extension press, Saskatchewan, ISBNO-88880-412-1.

Smith, M. A., V. B. Weaver, D. M. Young, and L. N. Ornston, 2003. Genes for chlorogenate and hydroxycinnamate catabolism (*hca*) are linked to functionally related genes in the *dca-pca-qui-pob-hca* chromosomal cluster of *Acinetobacter* sp. strain ADP1. Appl. Environ. Microbiol. 69: 524-532.

Sourice, S., V. Biaudet, M. El Karoui, S. D. Ehrlich, and A. Gruss, 1998. Identification of the Chi site of *Haemophilus influenzae* as several sequences related to the Escherichia coli Chi site. Mol. Microbiol. 27: 1021-1029.

Stanier, R.Y., and L. N. Ornston, 1973. The β-ketoadipate pathway. Adv. Microbiol. Physiol. 9: 89-151.

Stone, B. J. and Y. A. Kwaik, 1999. Natural competence for DNA transformation by *Leigionella pneumophila* and its association with expression of type IV pili. J. Bacteriol. 181: 1395-1402.

Takemori, S., M. Nakamura, K. Suzuki, and M. Katagiri, 1970. The kinetics of salicylate hydroxylase reaction. FEBS Lett. 6: 305-308.

Tanaka, H., Sugiyama, S., Yano, K. and Arima, K., 1957. Isolation of fumarylpyruvic acid as an intermediate of the gentisic acid oxidation by *Pseudomonas ovalis* var. S-5. Agr. Chem. Soc. Jpn. Bull. 21: 67-68.

Tennstedt, T., R. Szczepanowski, S. Braun, A. Puhler, and A. Schluter, 2003. Occurrence of integron-associated resistance gene cassettes located on antibiotic resistance plasmids isolated from a wastewater treatment plant. FEMS Microbiol. Ecol. 45: 239-252.

Thomas, C. M., and K. M. Nielsen, 2005. Mechanisms of, and barriers to, horizontal gene trasnfer between bacteria. Nature Rev. Microbiol. 3: 711-721.

Trautwein, G., and U. Gerischer, 2001. Effects exerted by transcriptional regulator PcaU from *Acinetobacter* sp. strain ADP1. J. Bacteriol. 183: 873-881.

van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos, 1991. Sequence analysis of the *Pseudomonas* sp. Strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol-1,2-dioxygenases for chlorinated substrates. J. Bacteriol. 173: 2425-2434.

van der Meer, J. R., W. M. De Vos, S. Harayama, and A. J. B Zehnder, 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol. Rev. 56: 677-694.

Vila, J., Z. Lopez, J. Sabate, C. Minguillon, A. M. Solanas, and M. Grifoll, 2001. Identification of a novel metabolite in the degradation of pyrene by *Mycobacterium* sp. strain AP1: actions of the isolate on two- and three-ring polycyclic aromatic hydrocarbons. Appl. Environ. Microbiol. 67: 5497-5505.

Vosman, B., P. J. G. Rauch, H. V. Westerhoff, and K. J. Hellingwerf, 1993. Regulation of the expression of *Pseudomonas stutzeri recA* gene. Antonie van Leeuwenhoek 63: 55-62.

Wackett, L. P., and C. D. Hershberger, 2001. Biocatalysis and biodegradation microbial transformation of organic compounds. ASM press Washington DC.

Wang, L-H., and S-C. Tu, 1984. The kinetic mechanism of salicylate hydroxylase as studied by initial rate measurement, rapid reaction kinetics, and isotope effects. J. Biol. Chem. 259: 10682-10688.

Wheelis, M. L., N. J. Palleroni, and R. Y. Stainer, 1967. The metabolism of aromatic acids by *Pseudomonas testosteroni* and *P. acidovorans*. Arch. Mikrobiol. 59: 302-14.

White-Stevens, R. H., and H. Kamin, 1972. Studies of a flavoprotein, salicylate hydroxylase. I. Preparation, properties, and the uncoupling of oxygen reduction from hydroxylation. J. Biol. Chem. 247: 2358-2370.

White-Stevens, R. H., and H. Kamin, 1972a. Studies of a flavoprotein, salicylate hydroxylase. I. Enzyme mechanism. J. Biol. Chem. 247: 2371-2381.

Wigmore, G. J., R. C. Bayly, and D. Di Berardino, 1974. *Pseudomonas putida* mutants defective in the metabolism of *meta*-fission product of catechol and its methyl analogues. J. Bacteriol. 120: 31-37.

Williams, P. A., and K. Murray, 1974. Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (*arvilla*) mt-2: evidence for the existence of a TOL plasmid. J. Bacteriol. 120: 416-423.

Williams, P. A., R. M. Jones, and E. L. Shaw, 2002. A third transposable element, IS*PpU12*, from the toluene-xylene catabolic plasmid pWWO of *Pseudomonas putida* mt-2. J. Bacteriol. 184: 6572-6580.

Wirenga, R.K., P. Terpstra, and W. G. J. Hol, 1986. Prediction of the occurrence of the ADP-binding βαβ fold in proteins, using an amino acid sequence fingerprint. J. Mol. Biol. 187: 101-107.

Wolgel, S. A., J. E. Dege, P. E. Perkins-Olson, C. H. Jaurez-Garcia, R. L. Crawford, E. Munck, and J. D. Lipscomb, 1993. Purification and characterization of protocatechuate 2,3-dioxygenase from *Bacillus macerans*: a new extradiol catecholic dioxygenase. J. Bacteriol. 175: 4414-4426.

Yeh, W-K., D. R. Durham, P. Fletcher, and L. N. Ornston, 1981. Evolutionary relationships among γ-carboxymuconolactone decarboxylases. J. Bacteriol. 146: 233-238.

Yamamoto, S., M. Katagiri, H. Maeno, and O. Hayaishi, 1965. Salicylate hydroxylase, a monooxygenase requiring flavin adenine dinucleotide. I. Purification and general properties. J. Biol. Chem. 240: 3408-3413.

Yang, Y., R. F. Chen, and M. P. Shiaris, 1994. Metabolism of naphthalene, fluorene, and phenanthrene: preliminary characterization of a cloned gene cluster from *Pseudomonas putida* NCIB 9816. J. Bacteriol. 176: 2158-2164.

Yen, K. M., and I. C. Gunsalus, 1985. Regulation of naphthalene catabolic plasmid NAH7. J. Bacteriol. 162: 1008-1013.

Yen, K. M. and C. M. Serdar, 1988. Genetics of naphthalene catabolism in pseudomonads. Crit. Rev. Microbiol. 15: 247-268.

You, I. S., D. Ghosal, and I. C. Gunsalus, 1991. Nucleotide sequence analysis of the *Pseudomonas putida* PpG7 salicylate hydroxylase gene (*nah*G) and its 3'-flanking region. Biochem. 30: 1635-1641.

Young, D. M., and L. N. Ornston, 2001. Functions of the mismatch repair gene *mutS* from *Acinetobacter* sp. strain ADP1. J. Bacteriol. 183: 6822-6831.

Young, D. M., D. Parke, and L. N. Ornston, 2005. Opportunities for genetic investigation afforded by *Acinetobacter baylyi*, a nutritionally versatile bacterial species that is highly competent for natural trasnformation. Annu. Rev. Microbiol. 59: 519-551.

Zhou, N-Y., S. L. Fuenmayor, and P. A. Williams, 2001. *nag* genes of *Ralstonia* (formerly *Pseudomonas*) sp. Strain U2 encoding enzymes for gentisate catabolism. J. Bacteriol. 183: 700-708.

Zhou, N-Y., J. Al-Dulayymi, M. Baird, and P. A. Williams, 2002. Salicylate hydroxylase from *Ralstonia* sp. strain U2: a monooxygenase with close relationships to and shared electron transport proteins with naphthalene dioxygenase. J. Bacteriol. 184: 1547-1555.