THE β-KETOADIPATE PATHWAY IN *RHIZOBIUM MELILOTI*

MASTER OF SCIENCE (1995)
(Biology)

McMaster University
Hamilton, Ontario

TITLE: The β-Ketoadipate Pathway in *Rhizobium meliloti*.

AUTHOR: Gordon MacPerson, B.Sc. (McMaster University)

SUPERVISOR: Dr. T.M. Finan

NUMBER OF PAGES: x, 129

Abstract

Tn*5* mutagenesis was used to generate four independent mutants of *Rhizobium meliloti* that were unable to grow on protocatechuate (Pca^-). Two of the Pca^- mutations were mapped to a region of the second symbiotic megaplasmid (pRmeSU47b) previously shown to be required for growth on protocatechuate. This *pca* locus was shown to consist of the first five structural genes of the protocatechuate branch of the β-ketoadipate pathway, in the order *pcaDCHGB*. This gene order is the same as determined for *Agrobacterium tumefaciens*. An additional reading frame with homology to LysR-type regulators was found to be upstream of, and transcribed divergently to the *pcaDCGHB* operon. This is likely to fulfil the same role as the regulatory gene, *pcaQ*, of *A. tumefaciens*. A cosmid plasmid which carried these *pca* genes failed to complement the Pca^- phenotype of a strain carrying a 300 kb megaplasmid deletion encompassing this *pca* locus. This implies that another *pca* locus, perhaps *pcaIJ*, is present within the deleted region of the megaplasmid. Two Pca^- Tn*5* insertions which did not map to the megaplasmid locus were isolated. One of these insertions appears to be in a catalase gene.

Acknowledgements

I would like to thank my supervisor, Dr. Finan for all of his encouragement and patience. I would like to thank all of the people that I shared the Finan laboratory with, namely Sylvie Bardin, Shan Dan, Brian Driscoll, and Magne Osteras. Thanks also go to Belinda Schoeman, Danielle Segato, and Xiaobang Chen, who rendered considerable technical assistance. I would also like to thank my family; my brother Andrew, my Mom and Dad, and my inlaws Dick and Ann, who are like parents to me. And last but not least, I would thank my wife Marina, for staying beside me through it all, with her love and understanding.

Table I. Contents

Table II: List of Figures

Introduction

Results

III: List of Tables

Introduction

The *Rhizobiaceae*

Rhizobia are gram-negative, rod-shaped bacteria, and include the fast-growing *Rhizobia* and the slow-growing *Bradyrhizobia*. *Rhizobia* are unique in that they induce and colonize root nodules of members of the *Leguminosae* family. Root nodules are plant-derived structures which contain a differentiated form of *Rhizobia*, termed bacteroids. The bacteroids found within the root nodules are capable of reducing atmospheric nitrogen to ammonia, which is then utilized by the plant. Different species of *Rhizobia* are capable of nodulating different host species of legumes. *Rhizobium meliloti* nodulates *Medicago*, *Melilotus*, and *Trigonella* species.

The symbiosis between the legume and *Rhizobium* is the result of a complex process involving both the plant and bacteria. The first step in nodule formation is the colonization of the plant roots by the bacteria (Bauer, 1981). Root hair curling then occurs. This is induced by an oligosaccharide signal molecule produced by the bacteria (LeRouge et al., 1990). This is synthesized by the bacteria

in response to plant-produced polyphenolic signal molecules called flavonoids. In addition to root hair curling, meristematic activity is induced in the root cortex. Bacteria enter the plant via infection threads, which arise from invagination of the cell wall of root hairs. The infection threads develop towards the inner cortical cells of the roots, where the root nodule develops. As the bacteria are released from the infection threads into the plant cell cytoplasm they are surrounded by plant-derived peribacteroid membrane.

As soil bacteria, it is to be expected that aromatic compounds play a role as an important source of reduced carbon for *R. meliloti*. Plant-derived phenolics, from leachates, residues, and animal excreta are common in soil (reviewed in Siqueira et al. 1991). The aromatics may be present as low molecular weight phenolic compounds (see Table 1-1) or as high-molecular weight polyphenolic compounds, such as lignin. Many rhizobial strains are able to grow on laboratory media when supplied with aromatic compounds as sole carbon source (Parke and Ornston, 1984). Thus, the study of bacterial aromatic catabolism is important to understanding of the ecology of bacteria in their native habitat, the soil.

Table 1-1: Phenolic Compounds Found Commonly in soil

p-hydroxybenzoic acid	Cineole acid
Vanillic acid	Salicylic acid
p-Coumaric acid	Hydroxyphenylacetic acid
Protocatechuic acid	Phenylacetic acid
Ferulic acid	4-Phenylbutyric acid
Syringic acid	1,4,5-Trihydroxynaphthalene
Gallic acid	Chlorogenic acid
Gentistic acid	Juglone
Benzoic acid	flavoniods
Caffeic acid	Lignin
Cinnamic acid	Tannin
Hydrocinnamic acid	Chalcones
O-Coumaric acid	Coumarins

A list of phenolic compounds commonly found in soil, taken from Siqueira et al., 1991.

The β-ketoadipate Pathway

In the catabolism of aromatic compounds by aerobic bacteria, a large variety of aromatic compounds are

typically converted to a few common intermediates, which are then metabolized by only a few different pathways. The diphenols protocatechuate and catechol are examples of such intermediates. These diphenols may be cleaved according to two methods. One of these is *ortho* cleavage, in which ring cleavage occurs between the vicinal hydroxyl groups. The other is *meta* cleavage, in which ring cleavage occurs adjacent to one of the vicinal hydroxyl groups.

In the β-ketoadipate pathway (diagrammed in Figure 1-1), the first step is *ortho* ring cleavage. This step is mediated by an oxygenase. Thus, the β-ketoadipate pathway is strictly aerobic. The product of this first step is a *cis,cis*-muconic acid. The muconic acid is then converted to a muconolactone by *cis,cis*-muconate-lactonizing enzyme. At this stage the intermediates in the two branches of the pathway only differ in that the protocatechuate branch intermediates have an additional carboxyl group. The *cis,cis*-muconolactones are converted to the common intermediate β-ketoadipate by muconolactone isomerase and γ-carboxymuconolactone decarboxylase. The remaining steps in the pathway are common to both the protocatechuate and catechol branches. β-ketoadipate is converted to the CoA ester by a CoA transfer from succinyl-CoA. This reaction is catalyzed by β-ketoadipyl succinyl-CoA transferase. β-ketoadipyl-CoA thiolase then catalyzes the removal of

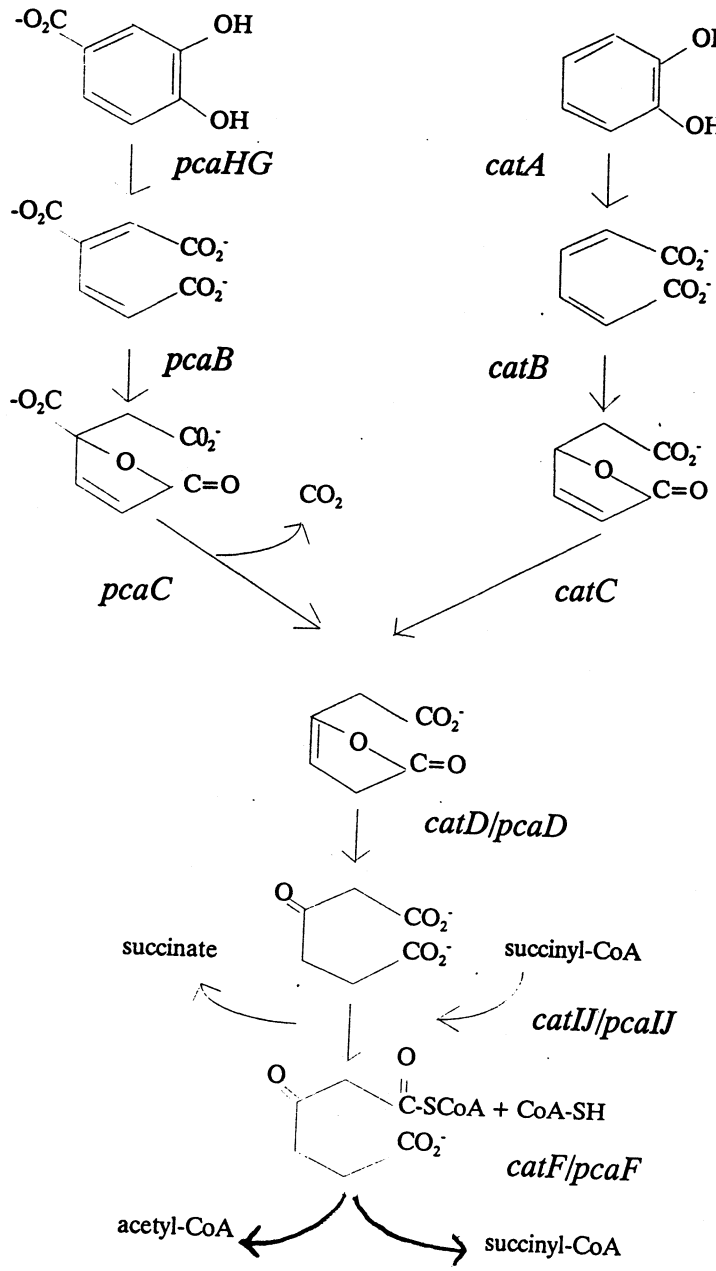


Figure 1-1: The two branches of the β-ketoadipate pathway. From Stanier and Ornston, 1973. The enzymes encoded by the genes are as follows: *pcaHG*, protocatechuate 3,4-dioxygenase; *pcaB*, β-carboxy-*cis,cis*-muconate lactonizing enzyme; *pcaC*, γ-carboxymuconolactone decarboxylase; *catA*, catechol 1,2-dioxygenase; *catB*, *cis,cis*-muconate lactonizing enzyme; *catC*, muconolactone isomerase; *pcaD/catD*, β-ketoadipate enol-lactone hydrolase; *pcaIJ/catIJ*, β-ketoadipate succinyl-CoA transferase; *pcaF/catF*, β-ketoadipyl-CoA thiolase.

acetyl-CoA with the regeneration of succinyl-CoA.

The intermediates of the two branches of the pathway are analogous to each other. The enzymes in the two branches prior to convergence have different proteins catalyzing the analogous reactions, and the enzymes in one branch will not act on analogous substrates from the opposite branch (Stanier and Ornston, 1973).

Genetics of the β-ketoadipate Pathway

Most work on the genetics of the β-ketoadipate pathway have been done with *Acinetobacter calcoaceticus* and *Pseudomonas putida*. The genes of the β-ketoadipate pathway are listed in Figure 1-1.

In *Acinetobacter calcoaceticus*, the *pca* genes are clustered in a single transcriptional unit, and have been sequenced (Doten et al., 1987; Hartnett et al., 1990; Kowalchuk et al., 1994). The gene order is *pcaIJFBDKCHG*. The *cat* genes of *A. calcoaceticus* have also been sequenced (Neidle et al., 1988, 1989; Shanley et al., 1994). While tightly linked, the *cat* genes are organized into three separate transcriptional units. The organization of the *pca* and *cat* genes of *A. calcoaceticus* are compared in Figure 1-2.

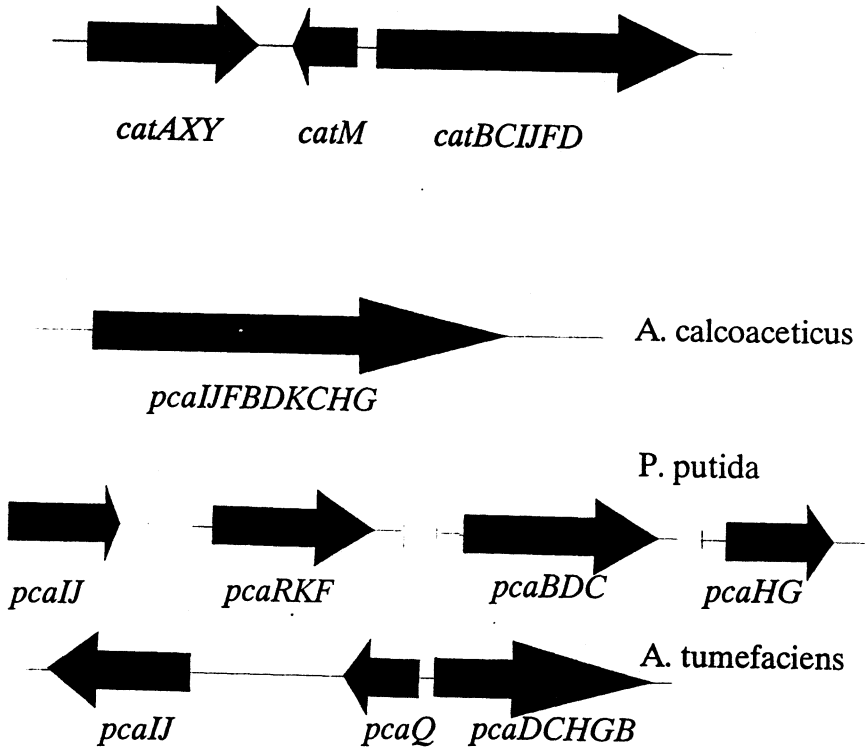


Figure 1-2. The organization of the *pca* and *cat* genes of *Acinetobacter calcoaceticus* (Ornston and Neidle, 1991, Kowalchuk et al., 1994), and the *pca* genes of *P. putida* (Frazee et al., 1993; Harwood et al., 1994; Parales and Harwood, 1992), and *Agrobacterium tumefaciens* (Parke, 1995).

Genes involved in the conversion of substrates shikimate, quinate, and *p*-hydroxybenzoate are linked to the *pca* genes in *A. calcoaceticus*. Genes involved in the conversion of *p*-hydroxybenzoate lie within 10 kb of the *pcaG* gene. These genes include *pobA*, the structural gene for *p*-hydroxybenzoate hydroxylase, which converts *p*-hydroxybenzoate to protocatechuate (DiMarco et al., 1993a); *pobR*, a transcriptional activator of *pobA* (DiMarco et al., 1993b) and *pobS*, which may act as a repressor of *pobA*.

The quinate genes are located between *pcaG* and *pobS*. This was shown by constructing deletion mutants, and demonstrating that these mutants could not grow with either shikimate or quinate as sole carbon source (Elsemore and Ornston, 1994). The presence of all three quinate genes in this region was demonstrated. The gene for the quinate dehydrogenase, *quiA*, was sequenced. The other two quinate genes present were *quiB*, the dehydroquinate dehydratase, and *quiC*, the dehydroshikimate dehydratase. Thus, superoperonic clustering of genes involved in the catabolism of aromatics are present in *Acinetobacter calcoaceticus*.

Superoperonic clustering of genes involved in benzoate catabolism is also observed in *A. calcoaceticus* (Neidle et al., 1987). Four genes involved in the conversion of benzoate to catechol are located adjacent to

the cat genes. These genes are the *benABC* genes, encoding the benzoate-1,2-dioxygenase, and *benD* gene, encoding 2-hydro-1,2-dihydroxybenzoate dehydrogenase. All of the *ben* and *cat* genes lie within a 16 kb region.

In *Pseudomonas putida*, the arrangement of *pca* genes differs from that of *Acinetobacter calcoaceticus* (Figure 1-2), being arranged in four separate operons. Following Tn*5* mutagenesis of *P. putida*, four of the structural *pca* genes were isolated, by cloning the Tn*5* insertions and flanking DNA (Hughes et al., 1988). Three of these genes were shown to be clustered in a single operon, in the order *pcaBDC*. This differs from the order in *A. calcoaceticus*, which is *pcaDBC*. It was also shown that in *P. putida*, the *pcaIJ* genes lie greater than 15 kb away from the *pcaBDC* gene cluster.

Another cluster of *pca* genes lies about 15 kb upstream of the *pcaBDC* gene cluster in *P. putida* (see Figure 1-3). This is the *pcaRKF* gene cluster (Hughes et al., 1988; Harwood et al., 1994). The *pcaR* gene is a regulatory protein required for expression of the *pcaBDC* and *pcaKF* genes (Hughes et al., 1988; Romero-Steiner et al., 1994).

The *P. putida* pcaK gene product, like that of *A. calcoaceticus*, encodes a predicted hydrophobic protein with homology to many membrane-bound transport proteins (Harwood et al., 1994). This gene was shown to catalyze *p*-

hydroxybenzoate uptake in *E. coli* when expressed in that host. The *pcaK* gene was also shown to be involved in chemotaxis to *p*-hydroxybenzoate in *P. putida* (Harwood et al, 1994).

The remaining structural genes for protocatechuate catabolism in *P. putida*, *pcaHG* and *pcaIJ* have also been cloned from *P. putida* and sequenced (Frazee et al., 1993; Parales and Harwood, 1992 respectively). The genes encoding the protocatechuate 3,4-dioxygenase are transcribed in the order *pcaHG*, as in *A. calcoaceticus*. The genes encoding the β-ketoadipate: succinyl-CoA transferase in *P. putida* are also transcribed in the same order, *pcaIJ*, as in *A. calcoaceticus*.

Unlike *A. calcoaceticus*, *P. putida* only has the first three structural *cat* genes (Brown et al., 1992). The three steps common to the catabolism of protocatechuate and catechol are fulfilled by the corresponding *pca* genes.

Some data is also available for *Agrobacterium tumefaciens*. The *pcaDCHGB* genes are located in a single operon (Figure 1-2), and are transcribed divergently from their regulator, *pcaQ* (Parke, 1995). The *pcaIJ* genes are located about 2 kb downstream of the *pcaQ* gene product.The genes encoding the protocatechuate 3,4-dioxygenase have also been cloned and sequenced from *Pseudomonas cepacia* (Zylstra et al., 1989b) and *Bradyrhizobium japonicum USDA110* (Podila

et al., 1993), and are transcribed in the order *pcaHG* in both.

Regulation of the β-ketoadipate Pathway

In the bacteria examined to date, except for one, the enzymes of the β-ketoadipate pathway are inducible. The exception to this is *Bradyrhizobium japonicum* (Parke and Ornston, 1986), in which expression of the *pcaIJ*, *pcaB*, *pcaC*, and *pcaD* genes appears to be constitutive.

The single *pca* operon in *Acinetobacter calcoaceticus* is induced by protocatechuate (Canovas and Stanier, 1967). The regulatory gene(s) involved have not yet been identified. The genes involved in the conversion of the hydroaromatics quinate and shikimate, the *quiA*, *quiB* and *quiC* genes, are also induced by protocatechuate (Canovas et al., 1968; Ingledew et al., 1971). Expression of the *pobA* gene encoding the *p*-hydroxybenzoate monoxygenase, is controlled by the *pobR* and *pobS* gene products. The *pobR* gene product is a transcriptional activator of *pobA* (DiMarco et al., 1993), while the *pobS* gene may encode a repressor of *pobA* (Elsemore and Ornston, 1994). The *pobR* gene, along with the *pobS* gene lie upstream of the *pobA* gene in *A. calcoaceticus*, and both are transcribed divergently from *pobA*.

Expression of *cat* genes in *A. calcoaceticus* occurs in the presence of *cis,cis*-muconate. The *catM* gene flanks *catB* and is transcribed divergently from the *catBCIJFD* operon, and encodes an activator, which, in the presence of *cis,cis*-muconate, activates expression of the structural *cat* genes. There also appears to be an additional regulatory locus involved (Romero-Arroyo et al., 1995). The *ben* genes are induced by both *cis,cis*-muconate and benzoate. Induction of *ben* genes by *cis,cis*-muconate appeared to be independent of *catM* (Neidle et al., 1989).

Regulation of the β-ketoadipate pathway has also been studied in detail in *Pseudomonas putida*. The genes encoding the protocatechuate 3,4-dioxygenase are induced by protocatechuate (Hosokawa, 1970). The gene(s) mediating the regulation of *pcaHG* have not been identified. The genes encoding the remaining enzymes required for protocatechuate catabolism are induced by β-ketoadipate. This induction is mediated by the *pcaR* gene product (Hughes et al., 1988; Romero-Steiner et al., 1994). The *pcaR* gene locus is also involved in the regulation of chemotaxis to aromatic compounds (Harwood et al., 1994).

The *catBC* genes in *P. putida* are induced by *cis,cis*-muconate. This is mediated by the *catR* gene product, which acts as an inducer of *catBC* (Rothmel et al., 1991; Parsek et al., 1992). *Cis,cis*-muconate also acts as an inducer of

catA, encoding the dioxygenase (Wu et al, 1972). The remaining enzymes required to convert catechol to acetyl-CoA and succinate are common to both the protocatechuate and catechol branches of the β-ketoadipate pathways, and are induced by β-ketoadipate.

In *Pseudomonas cepacia*, the regulation of the β-ketoadipate pathway has not been as well studied as for *P. putida*. Protocatechuate 3,4-dioxygenase was shown to be induced by either protocatechuate or β-carboxy-*cis,cis*-muconate. Induction of the carboxymuconate lactonizing enzyme was shown to be induced by either β-ketoadipyl-CoA or β-ketoadipate (Zylstra et al., 1989).

Regulation of the β-ketoadipate pathway has not been studied to the same detail in members of the family *Rhizobiacea* as in *P. putida* and *A. calcoaceticus*. The first five enzymes of the protocatechuate branch of the pathway were shown to be inducible in *Agrobacterium rhizogenes*, *Rhizobium fredii*, *Rhizobium meliloti*, *Rhizobium leguminosarum*, and *Rhizobium trifolii* (Parke and Ornston, 1986). In contrast, of the five enzymes studied, only protocatechuate 3,4-dioxygenase was inducible in *Bradyrhizobim japonicum*, and cowpea *Bradyrhizobium spp.*; the rest of the enzymes were expressed constitutively.

Regulation of the β-ketoadipate pathway has been further studied in *Rhizobium leguminosarum* and *Agrobacterium*

tumefaciens. In *Rhizobium leguminosarum*, a novel pattern of induction was observed (Parke et al., 1991). Protocatechuate was found to be most likely the inducer of *pcaHG*. β-carboxy-*cis,cis*-muconate was found to be the most likely inducer of *pcaC*, *pcaD* and *pcaB*. Induction of *pcaIJ* was found to be due to β-ketoadipate.

In *Agrobacterium tumefaciens*, it was shown that β-carboxy-*cis,cis*-muconate acts as the inducer for the *pcaD* gene (Parke, 1993). This was shown by introducing a clone carrying a regulator of *pcaD*, the *pcaQ* gene, and *pcaD* into *E. coli*. A plasmid carrying the *pcaHG* genes of *A. calcoaceticus* was also introduced into *E. coli*. Growth of this strain in the presence of protocatechuate resulted in the production of β-carboxy- *cis,cis*-muconate intracellularly. Under these conditions, PcaD activity was detected. Thus it is possible that *Agrobacterium tumefaciens* and *Rhizobium leguminosarum* have similar regulatory systems.

Evolution of the β-ketoadipate Pathway

The two branches of the β-ketoadipate pathway are similar in that analogous steps are involved in the degradation of the similar compounds catechol and protocatechuate. These similarities in the analogous

reactions suggest that the analogous enzymes share common ancestry. The enzymes protocatechuate 3,4-dioxygenase and catechol 1,2-dioxygenase share amino acid homology. Their oligomeric structure differs, however (Neidle et al., 1988).

The muconate lactonizing enzyme (MLE) and carboxy muconate lactonizing enzyme (CMLE) do not appear to be related. These two enzymes appear to differ with respect to the reaction methods catalyzed. The reaction catalyzed by MLE proceeds by a syn cycloisomerization, while the reaction catalyzed by CMLE proceeds by an anti cycloisomerization (Chari et al., 1987a). Also, there is little homology between PcaB and CatB at the sequence level (Aldrich et al., 1987; Kowalchuk et al., 1994; Williams et al., 1992).

The same genes in *P. putida* and *A. calcoaceticus* share common ancestry (Ornston and Neidle, 1991), with amino acid identity usually greater than 40%, even though the GC content of the genes in the two organisms differ by 20%. The *catM* gene of *A. calcoaceticus* and the *catR* gene of *P. putida* also share common ancestry, as indicated by 40% amino acid identity (Neidle et al., 1989; Rothmel et al., 1990; Ornston and Neidle, 1991). This is despite the fact that *catM* appears to be a repressor and *catR* is an activator. Both CatM and CatR respond to the same metabolite, *cis,cis*-muconate. Both are members of the LysR family of regulators, and resemble each other more than they resemble

the LysR consensus sequence (Ornston and Neidle, 1991).

Gene rearrangement occurred during the divergence of *cat* and *pca* genes in *A. calcoaceticus* and *P. putida*. Tight linkage of both *cat* and *pca* chromosomal genes has been maintained in *A. calcoaceticus*. In contrast, *pca* genes in *P. putida* are organized in at least three operons on the chromosome, with *pcaBDC* representing the most extensive linkage. It is of interest to note that the order of *pcaD* and *pcaB* is reversed in *P. putida* when compared to the order in *A. calcoaceticus*. Clustering of genes, especially on a plasmid, may be the result of a requirement for co-transfer (Wheelis, 1975). It is of interest to note that in *A. calcoaceticus* strain RJE74, the *ben* and *cat* genes were located on a plasmid (Winstanley et al., 1987). Thus the *cat* and *pca* genes of *A. calcoaceticus* ADP1 may have been recently acquired from a plasmid (Ornston and Neidle, 1991).

It is also of interest that there exists homology at the DNA level between *cat* and *pca* genes in *A. calcoaceticus*. Also, the high-frequency repair of a *pcaJ* mutation appeared to be dependant of *catJ*, implying that gene conversion was taking place. Comparison of the two sequences revealed that *catIJF* and *pcaIJF* are 99% identical (Kowalchuk et al., 1994). In contrast, the regions directly flanking *pcaIJF* and *catIJF* are not similar. Comparison of *pcaD* and *catD* reveals that the two have diverged extensively at the amino

acid level, while remaining isofunctional enzymes, with amino acid identity at 44%.

It cannot be functional constraints on the proteins at the amino acid level which are responsible for the maintenance of the extreme nucleic acid similarity between *catIJF* and *pcaIJF* genes. The *P. putida pcaIJ* gene products only share 67% identity with the *pcaIJ* genes of *A. calcoaceticus* (Parales and Harwood, 1992). The *catIJF* and *pcaIJF* genes may have been recently acquired by *A. calcoaceticus*, and have not had enough time to diverge (Kowalchuk et al., 1994). It is known that *pcaIJF* and *catIJF* can exchange sequence information (Doten et al., 1987), and it has been also suggested that gene conversion may have prevented the *catIJF* and *pcaIJF* regions from diverging (Shanley et al., 1994)

Gene order has been suggested as a factor, which has allowed *pcaD* and *catD* to diverge, while maintaining the similarity between *catIJF* and *pcaIJF*. In *A. calcoaceticus*, *pcaB* lies between *pcaIJF* and *pcaD*, while *catD* is directly downstream of *catIJF*.

Chapter 2

Materials and Methods

Strains, Plasmids and Transposons

Strains were stored in glass vials as frozen permanents at -70°C in LB 7% dimethylsulfoxide (DMS). To prepare frozen permanents, strains were grown in LB (with antibiotic if appropriate). 0.5 ml of this culture was diluted with LB 14% DMS in a glass vial. Stocks were chilled to -20°C prior to placement at -70°C.

Rhizobium meliloti Strains

Strain	Genotype	Reference
RCR2011	SU47 wildtype	Rothamsted Experimental Station
Rm1021	SU47 *str-27*	Meade *et al.* 1982
Rm5000	SU47 *rif-5*	Finan *et al.* 1984
RmF153	Rm1021 *dctB*12::Tn*5-132*	Yarosh *et al.* 1989
RmF460	Rm1021 Ω5071::Tn*5-132* *thi-502*::Tn*5-233*	Yarosh, 1989
RmF726	Rm1021 ΔΩ5149-5079::Tn*5-233*	Charles 1990
RmG212	Rm1021 *lac*⁻ (Rm8501)	Jane Glazebrook
RmG867	Rm1021 *pca-867*::Tn*5*	This work
RmG877	Rm1021 *pca-877*::Tn*5*	This work
RmG878	Rm1021 *katE*::Tn*5*	This work
RmG879	Rm1021 *pcaG879*::Tn*5*	This work
RmH333	Rm5000 *pca-867*::Tn*5-233*	T.M. Finan
RmH334	Rm5000 *pca-877*::Tn*5-233*	T.M. Finan
RmH335	Rm5000 *katE878*::Tn*5-233*	T.M. Finan
RmH336	Rm5000 *pcaG879*::Tn*5-233*	T.M. Finan
RmH337	Rm5000 *pca-867*::Tn*V*	T.M. Finan

RmH338	Rm5000 *pca-877*::Tn*V*	T.M. Finan
RmH339	Rm5000 *katE878*::Tn*V*	T.M. Finan
RmH340	Rm5000 *pcaG879*::Tn*V*	T.M. Finan

E. coli Strains

Strain	Genotype	Reference
MM294A	*pro-82 thi-1 hsdR17 supE44 endA1*	
MT607	MM294A *recA56*	Finan *et al.* 1986
MT609	*thyA36 polA1* Sp^R	T. Finan
MT616	MT607 (pRK600)	T. Finan
DH5α	*endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1* Δ(argF-lacZYA)U169 Φ80 d*lacZ*ΔM15	B.R.L. B.R.L. Inc.
G312	MT607Ω5::Tn5-B20	Driscoll, 1995

Plasmids

Plasmid		Reference
pLAFR1	IncP cosmid cloning vector, Tc^R	Freedman et al. 1982
pRK7813	RK2 derviative (IncP) carrying the pUC9 polylinker	Jones and Gutterson, 1987.
pRK2013	ColE1 replicon with RK2 transfer region	Figurski and Helsinki, 1979.
pRK600	pRK 2013 *npt*::Tn*9*, Cm^R, Nm-Km^R	
pRK415-1	Tc^R, broad host range	Keen et al., 1988
pPH1JI	IncP, Gm^R Sp^R Cm^R	Beringer et al., 1978
pGS220	Tn5 in deletion derivative of pBR322 Ap^R, Nm-Km^R	DeVos et al., 1986
pTH178	pLAFR1 derivative complementing Pca^- phenotype of RmG867, RmG879	This work.
PTH179	pLAFR1 derivative complementing Pca^- phenotype of RmG867, RmG879	This work.
pTH180	pLAFR1 derivative complementing Pca^- phenotype of RmG877	This work.

pTH181	pLAFR1 derivative complementing Pca$^-$ phenotype of RmG877	This work.
pTH182	pLAFR1 derivative complementing Pca$^-$ phenotype of RmG877	This work.
pTH185	5kb *Eco*RI fragment #1 from pTH178 in pUC119	This work.
pTH186	5kb *Eco*R1 fragment #2 from pTH178	This work.
pTH187	5kb *Eco*R1 fragment #1 from pTH178 in pRK7813, orientation 1	This work.
pTH188	5kb *Eco*R1 fragment #1 from pTH178 in pRK7813, orientation 2	This work.
pTH193	5kb *Pst*I fragment from pTH178 in pRK7813, orientation 1	This work.
pTH194	5kb *Pst*I fragment from pTH178 in pRK7813, orientation 2	This work.
pTH215	5kb *Pst*I fragment from pTH178 in pUC119	This work.
pARO2	*pcaCHG* genes from *A. calcoaceticus* strain ADP1 in pRK415-1	Parke et al., 1991
pARO4	*pcaFBFDC* genes from *A. calcoaceticus* strain ADP1 in PRK415-1	Parke et al., 1991

Transposons

Transposon		**Reference**
Tn*5*	Nm^R	Berg and Berg, 1987
Tn*5-132*	Ot^R	Berg and Berg, 1987
Tn*5-233*	Gm^R, Sp^R	DeVos et al., 1986
Tn*V*	Tn*5* containing pSC101 *oriV*, Nm^R	Furichi et al., 1985
Tn*5-B20*	Tn*5* with *lacZ*, Nm^R	Simon et al., 1989

Bacterial Growth

Cultures of *Rhizobium meliloti* were grown at a temperature of 30°C, and *E. coli* cultures were grown at either 30 or 37°C. Components of the growth media were sterilized by filter sterilization through 0.45 μm filters,

for heat labile solutions, or by autoclaving at 18 p.s.i. for 20 minutes. Cultures were started from single colonies. Newly constructed strains were purified to single colonies three times prior to use in experiments.

Complex media used for bacterial growth was Luria-Bertani (LB); the defined media used was M9. For solid media, 1.5% agar was added. Difco Bacto agar was used when growing *R. meliloti* on solid media; Difco Bacto agar or Merck agar was used when growing *E. coli*.

LB medium was composed of, per litre, 10 g of Difco tryptone, 5 g of Difco yeast extract, and 5 g of NaCl. 4 ml of 1M NaOH was added to a litre of liquid LB broth, and 1 ml 1M NaOH was added to a litre of solid media. LB medium as described above was used for liquid cultures of *E. coli*; 2.5 mM $MgSO_4$ and 2.5 mM $CaCl_2$ (LB mc)was used for growth of *R. meliloti*.

M9 medium was composed of, per litre, 5.8 g of Na_2HPO_4, 3 g of KH_2PO_4, 0.5 g of NaCl, and 1 g of NH_4Cl. For solid M9 media, a 2 times concentrate solution was autoclaved prior to the addition of 1.5% agar in distilled water. 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 1 μg/ml D-biotin and the carbon source were added to M9. Carbon sources used were D-glucose (15mM), disodium succinate (15 mM), L-arabinose (15mM), sodium quinate (5 mM), sodium protocatechuate (5 mM), and sodium *p*-hydroxybenzoate (5 mM). Solutions of

aromatic compounds were prepared by dissolving crystals in 50% ethanol:50% water mixture, and adjusting the pH to 7 with sodium hydroxide. Thiamine auxotrophs were grown on media supplemented with 0.01 mM thiamine. Thymidine auxotrophs were grown on media supplemented with 60 μg/ml of thymidine.

LB/M9 media was used for transductions, and consisted of 50% 1X M9 and 50% LB containing 3% agar (LB 2X agar).

Antibiotics

Concentrations of antibiotics used and concentrations of stocks are listed in Table 2-5. For liquid cultures, the lower concentration of antibiotics was used; for growth on solid media, the higher concentration was used. Some antibiotics were used to select for *R. meliloti* (R), some were used to select for *E. coli* (E), and some were used to select for both (R + E). Stocks of antibiotics were filter sterilized with 0.45μm filters. Oxytetracycline, tetracycline, and rifampicin were prepared in ethanol, and chloramphenicol was prepared in 50% ethanol; the remainder were prepared in distilled deionized water.

Table 2-5: Antibiotics

Antibiotic	**Conc. (μg/ml)**	**For use with**	**Stock (mg/ml)**
Neomycin sulphate Nm	100-200	R	20
Spectinomycin dihydorchloride Sp	100-200	R,E	20
Chloramphenicol Cm	10-20	E	2
Rifampicin Rf	25-50	R	5
Kanamycin sulphate Km	20-50	E	6
Gentamycin sulphate Gm	20-70	R,E	6
Tetracycline hydrochloride Tc	2-10	R,E	2
Streptomycin sulphate Sm	100-200	R,E	20
Na-Ampicillin Amp	20-100	E	10
Oxytetracycline hydrochloride Ot	0.3-0.5	R	2

List of antibiotics and concentrations used. Concentrations are given in μg/ml. E indicates used for *E. coli*; R indicates used for *R. meliloti*.

ΦM12 Generalized Transduction

ΦM12 lysates of *R. meliloti* strains were prepared by adding 0.03 ml of wild-type lysate to 5 ml of culture (A_{675} 0.4-0.5). This culture was then incubated for 8 to 16 hours at 30°C. Several drops of chloroform were added to kill any uninfected cells. The culture was then incubated overnight at 4°C. The lysate was transferred to a screw-top vial, and centrifuged at 4000 rpm for 5 minutes in a GPR centrifuge

with a GA-10 rotor. Lysates were then stored at 4°C.

To transfer genetic markers between strains of *R. meliloti*, 1 ml of late log-phase culture was mixed with 1 ml of a 1/40 dilution of lysate in LB mc. The mixture was incubated for 20 minutes at room temperature without shaking. 2.5 ml of 0.85% saline was added, and the cells pelleted (4000 rpm in a GPR centrifuge with a GA-10 rotor). The cells were resuspended in 2.5 ml of 0.85% saline, and pelleted again. The cells were then resuspended in 0.5 ml of 0.85% saline, and 100 μl aliquots were plated on LB/M9 with the appropriate antibiotics. Donor and recipient controls were plated for each case.

The resulting transductants were patched onto M9 or LB/M9 media with appropriate antibiotics for screening of unselected markers, or were streak-purified on LB/M9. The Wu equation was used to convert cotransduction frequencies to physical distances. The equation is $c = (1-d/L)^3$ where c is the cotransduction frequency, d= distance between markers in kb, and L is the size in kb of the genome of the transducing phage (160 kb for ΦM12).

Bacterial Matings

In matings, MT616 (MT607(pRK600)) was used as a mobilizing strain. For spot matings, equal volumes of

donor, recipient, and mobilizer cultures were used. Cultures were washed with 0.85% saline to remove any antibiotics. Equal volumes of cultures were mixed, and 0.1 ml aliquots of these mixtures were spotted on an LB plate and incubated overnight. The spots were then resuspended in 0.85% saline and plated on selective media. For patch matings, equal amounts of donor, recipient, and mobilizer from fresh plates were mixed on an LB plate and incubated overnight. The patches were then struck out on selective media.

Isolation of Complementing Cosmid Clones

In order to isolate clones complementing *pca* insertional mutants, spot matings of mutants with a cosmid clone bank were performed. The clone bank used was the pLAFR1 clone bank of *Eco*R1 partially restricted Rm1021 DNA (Meade et al.,1982). One ml of frozen clone bank was added to 4 ml of LB with 0.5 μg/ml of Tc. This culture was grown to an OD of 0.5 before mixing with the mutant culture and *E. coli* MT616. Spot matings were performed as described above, and strains carrying complementing clones were selected on M9-protocatechuate.

Transposon Mutagenesis of the Rm1021 Genome

Donor and recipient strains were grown to late log phase. The cultures were washed with 0.85% saline and resuspended in LB mc. 1 ml volumes of donor and recipient were mixed, and 1 ml of this was transferred to a sterile, LB mc washed Millepore type HA 0.45 μm filter, under vacuum. The filters were incubated on LB plates for 4 hours at 30°C. The filters were then resuspended in 5 ml of 0.85% saline, and 10^0, 10^{-1}, and 10^{-2} dilutions were plated on LB plates containing Sm (200μg/ml) and Nm (200 μg/ml). The plasmid containing the Tn*5* (pRK602) cannot replicate in *Rhizobium*, so Nm^R colonies of *Rhizobium* contained Tn*5* insertions on the chromosome. The colonies were screened for Pca^- phenotype by patching on M9 media containing glucose, succinate, and protocatechuate as carbon sources.

Site-Directed Tn*5-B20* Transposon Mutagenesis

The plasmid to be mutagenized was transformed into the Tn*5-B20* donor strain, *E. coli* G312 (MT607 Ω5::Tn*5-B20*). Individual transformant colonies were used to start 0.5 ml cultures in LB, and allowed to incubate overnight. Overnight cultures of MT609 recipient and MT616 mobilizer

were washed in saline and mixed together. A 0.5 ml volume of this mixture was added to each of the cultures of donor containing the plasmid. After mixing, 25 μl from each set of cultures was spotted onto LB Thymidine plates, and incubated overnight. The mating spots were resuspended in 0.5 ml of 0.85% saline, and 100 μl from each was plated on LB thymidine (60 μg/ml) plates containing Sp (100 μg/ml) Km (20 μg/ml) and Tc (5 μg/ml). One transconjugant colony was picked from each plate, and was streak-purified. Transposon insertions into insert DNA contained in the plasmid were identified by gel electrophoresis of EcoRI restrictions of the plasmids.

Homogenotization

The Tn*5-B20* insertions on pTH178 were recombined into the genome as follows. Cultures of Rm5000 carrying the insertions on pTH178 were grown overnight in LB Tc. Overnight cultures of the donor strain *E. coli* J53(pPH1JI) were grown in LB Gm Sp. The J53(pPH1JI) cultures were subcultured the next morning. The cells from both sets of cultures were pelleted, and washed with saline that afternoon. The cultures were resuspended in saline. Equal volumes of the Rm5000 pTH178::Tn*5-B20* were mixed with J53(pPH1JI), spotted onto LB plates, and incubated overnight

at 30°C. The following morning the mating spots were resuspended in 1 ml of saline, and 100μl of the undiluted mixture and of a 10^{-1} dilution were plated on LB containing Gm (70μg/ml), Rf (25 μg/ml) and Nm (100 μg/ml). The plates were then incubated at 30°C for four days. Several colonies from each mating were streak-purified on the same media. They were then screened for loss of the Tc^R of the pTH178 (pLAFR1) derivative. Since pPH1JI and pLAFR1 are of the same incompatibility group, they can not coexist in the same cell. By selecting for the Gm^R of pPH1JI, one can select for cells which have lost the pLAFR1. By also selecting for the Nm^R of the Tn*5-B20*, one can also select for cells which have retained the Tn*5-B20*, either by recombination between *R. meliloti* DNA on the pTH178 derivative or by transposition of the Tn*5-B20*. Of the two, the first is most likely to occur.

Replacement of Transposon Insertions

The Tn5 insertion to be replaced with a Tn*5-233* insertion was transduced into Rm5000. The self-transmissible plasmid pRK607 was mated from its *E. coli* host into the Rm5000 bearing the Tn*5* to be replaced. The *E. coli* host was grown up in LB Gm Sp, and subcultured the next morning. The cells were then pelleted, washed, and

resuspended in saline. The Rm5000 strains were grown overnight in LB mc. Equal amounts of the Rm5000 strain and *E. coli* cells were mixed together and spotted on a plate. The plate was incubated overnight at 30°C. Cells from the mating spot were then resuspended in saline and plated on LB Rif Gm Sp. Cells were then purified on the same media, and then screened for loss of the neomycin resistance encoded by the Tn5.

Plant Growth Conditions and Dry Weight Determination

Plants were grown in Leonard assemblies, which are plastic pots sitting in 250 ml beakers. A piece of cotton placed in a hole in the plastic pot extends into the beaker. The bottom of the beaker is also covered with cotton. The pot is filled with a nitrogen-free sand-vermiculite mixture (1:1 w/w), and with 250 ml of 1X Jensen's media (prepared as a 2X solution: 1 g $CaHPO_4$, 0.2 g K_2HPO_4, 0.2 g $MgSO_4 \cdot 7H_2O$, 0.2 g NaCl, 0.1 g $FeCl_3$, 1 ml of 1000X trace mineral solution (1 g H_3BO_3, 1 g $ZnSO_4 \cdot 7H_2O$, 0.5 g $CuSO_4.5H_2O$, 0.5 g $MnCl_2 \cdot 4H_2O$, 1 g $Na_2MoO_4.2H_2O$, 10 g Na_2EDTA, 2 g NaFeEDTA, 0.4 g biotin per litre) per litre). The assembly was covered on top and at the junction of the beaker and pot with aluminum foil, and autoclaved for 4 hours.

Alfalfa (*Medicago sativa* cultivar Iroquois) seeds were surface sterilized in 95% ethanol for 5 minutes, and then for 20 minutes in 2.5% sodium hypochlorite. They were then rinsed once with sterile LB, and then several times with sterile water over a period of 1 hour. Seeds were then placed on sterile water agar plates (1.5% agar) and left in the dark for 2 days to germinate. Ten germinated seeds were planted in each pot, and the foil cover replaced. The pots were then placed in a Conviron plant-growth chamber, with day set at 18 hours and 21°C, night at 6 hours and 17°C. Two days later 0.2 ml of a *R. meliloti* overnight culture was added to 10 ml of sterile ddH_2O for each plant. Independent cultures were used for each pot. This was used to inoculate the seedlings, and the pots were uncovered. Three pots were inoculated for each strain. Two controls were also set up, each in triplicate: wild-type *R. meliloti* (Rm1021) and uninoculated. The plants were grown for 28 days, and were watered during this time with sterile ddH_2O. Symbiotic effectiveness of the strains was determined by measuring the triplicate dry weights of the shoots after 28 days. Plants were harvested 28 days after inoculation, and were cut at the junction of the shoot and root. Shoots from each pot were placed in paper bags and dried at 70°C for one week. Dry weights were then determined by weighing the shoots from each bag.

Boiling Method Miniprep for Isolation of Plasmid DNA

Cells were grown overnight in 5 ml of LB containing the appropriate antibiotic. 1.5 ml of this culture was transferred to a 1.5 ml Eppendorf tube. The cells were pelleted in a microfuge for 30 seconds at 14,000 rpm. The cells were then resuspended in 0.35 ml STET (8% sucrose, 0.5% Triton X-100, 50 mM EDTA pH 8, 50 mM Tris pH 8). Ten μl of a freshly prepared lysozyme solution (50 mg/ml in 10 mM Tris pH 8) was added to this. The tubes were floated in a boiling water bath for 3 minutes, with the heat being shut off right after the tubes were added. The tubes were then centrifuged for 15 minutes at 14,000g, and the pellet was removed. The plasmid DNA was precipitated with 330 μl ice-cold isopropanol, and placed at -70 °C for 15 minutes. The tube was centrifuged for 15 minutes at 14,000g to pellet the DNA, and the supernatant was discarded. The pellet of DNA was washed with 0.7 ml of 70% and then 95% ethanol. The pellet was then dried by placing the tube in a 37°C incubator for 20 minutes. The pellet was resuspended in 50 to 100 μl $T_{20}E_1$.

Alkaline Lysis Miniprep for Isolation of Plasmid DNA

Cells were grown overnight in 10 ml of LB containing the appropriate antibiotic. The culture was transferred to a 15 ml polypropylene Falcon tube, and pelleted for 5 minutes at 4,000 rpm in a GPR centrifuge with a GA-10 rotor. The cells were resuspended in 1 ml of TEG (50 mM Tris pH 7.8, 20 mM EDTA pH 8, 1% glucose) with 10 μl of a freshly prepared lysozyme solution (50 mg/ml in 10 mM Tris pH 7.8). Two ml of alkaline lysis solution (ALS; 0.2M NaOH, 1% SDS) was added and mixed into the tube, followed by 4.5 ml of ddH_2O and 1.5 ml HSS solution (3M potassium acetate pH 4.8). The tube was cooled for 15 minutes at -70 °C and the cellular debris and chromosomal DNA precipitated at 4,000 rpm for 15 minutes in the GPR centrifuge. The supernatant was poured off, and the pellet resuspended in 500μl of $T_{20}E_1$. This was then transferred to an Eppendorf tube. The DNA was cleaned with 1:1 phenol:chloroform-isoamyl alcohol, followed by cleaning with chloroform-isoamyl alcohol. The plasmid DNA was precipitated by mixing in 5 μl of 5M NaCl and 1 ml absolute ethanol and by placing the tube at -70°C for 15 minutes. The DNA was pellet by centrifugation at 14,000g for 15 minutes in a microfuge. The pellet was washed with 0.75 ml of 70% ethanol followed by 0.75 ml of 95% ethanol.

The pellet was dissolved in 50 μl of $T_{20}E_1$.

Large Scale Alkaline Lysis Preparation of Plasmid DNA

A 5 ml culture of *E. coli* carrying the plasmid of interest was grown overnight in LB containing the appropriate antibiotic. This was used to inoculate a 250 ml volume of the same media. The large culture was grown overnight with shaking at 37°C. The cells were pelleted in a sterile 250 ml centrifuge bottle at 6,000 rpm for 15 minutes at 4°C in a GSA rotor. The cells were resuspended in 20 ml of TEG with 0.2 ml of freshly prepared lysozyme solution (50 mg/ml in 10 mM Tris pH 7.8). Forty ml of ALS was added to the solution, and mixed until it became clear. Ninety ml of ddH_2O and 30 ml of HSS were added to the solution and mixed. The bottle was placed at -70°C for one hour. The debris was pelleted at 4°C at 6,000 rpm for 10 minutes, and the supernatant filtered through several layers of sterile cheesecloth into a new 250 ml centrifuge bottle. The DNA was precipitated by adding 90 ml of isopropanol and mixing. The DNA was pelleted by centrifugation for 10 minutes at 6,000 rpm, at 4°C. The pellet of DNA was washed with 10 ml of 70% ethanol, followed by 95% ethanol. The pellet was dried and resuspended in 5.5 ml of $T_{50}E_{20}$. The solution was transferred to a 15 ml polypropylene Falcon

tube, and the volume was adjusted to 9 ml by addition of $T_{50}E_{20}$. Then 9.9 g of CsCl was dissolved in this solution. The solution was added to a 13 ml ultracentrifugation tube. This was followed by addition of 0.5 ml of ethidium bromide (10 mg/ml). The tubes were topped up and balanced with paraffin oil, and the tubes were heat sealed. The tubes were then placed in a NVT65 rotor, with the rotor caps being tightened with a torque wrench, and the rotor was placed in a pre-cooled Beckman L8-70 centrifuge. The tubes were centrifuged at 55,000 rpm at 10°C for a minimum of 18 hours. The tubes were then removed to a dark room. The plasmid DNA band was visualized by using a shortwave UV lamp. A needle was inserted into the top of the ultracentrifuge tube to allow for removal of the DNA. The band was removed using a 3 ml syringe and a 21 gauge needle. This volume of approximately 2 ml was placed in a 50 ml Sorvall centrifuge tube. The ethidium bromide was removed from the solution by washing the solution with isopropanol:CsCl saturated ddH_2O, and removing the isopropanol phase. The remaining solution was diluted with an equal volume of ddH_20, and the DNA was precipitated by adding 6 volumes of 95% ethanol and placing at - 70°C for 30 minutes. The DNA was pelleted at 8,000 rpm for 15 minutes at 4 °C in an SS34 rotor. The pellet was resuspended in 6 ml of ddH_2O, and pelleted again. The DNA was resuspended in 400 μl of $T_{50}E_{20}$, and transferred to an

Eppendorf tube. The DNA was precipitated again by addition of 10 μl of 5M NaCl and 800 μl ethanol, placing it at -70°C for 30 minutes, and pelleting in a microfuge for 15 minutes. The pellet was dried, resuspended in 400 μl of $T_{20}E_1$, and extracted with 0.5 ml of phenol:chloroform-isoamyl alcohol twice, followed by extraction once with chloroform-isoamyl alcohol. The DNA was then precipitated by adding 10 μl of 5 M NaCl and 800 μl of ethanol, mixing and placing at -70°C for 15 minutes, and then pelleting in a microfuge for 15 minutes. The pellet was washed with 70% ethanol, and then with 95% ethanol. It was then dissolved in 200 μl of $T_{20}E_1$.

Total Genomic DNA Isolation

Cells were grown overnight in 5 ml of LB mc. Three ml of this culture were pelleted in a 15 ml polypropylene Falcon tube. The cells were resuspended in 5 ml of ice-cold 0.85% saline, the pelleted again. The cells were then resuspended in 5 ml of ice-cold TES (10 mM Tris pH 7.8, 25 mM EDTA pH 8, 150 mM NaCl), pelleted and resuspended in 2.5 ml ice-cold $T_{10}E_{25}$. Lysozyme (0.25 ml of 2 mg per ml of $T_{10}E_{25}$) was mixed in, and the solution was placed in a 37°C water bath for 15 minutes. Following this, 0.35 ml of sarkosyl-protease (5 mg of pronase E per ml of $T_{10}E_{25}$, allowed to self-digest for 2 hours, then 19% sarkosyl added)

was added, and incubation at 37°C continued for another 30 minutes. The resulting cell-lysate is then extracted with an equal volume of phenol until very little debris at the interface was observed. This is followed by two extractions with chloroform-isoamyl alcohol. NH_4-acetate was added to the solution to a final concentration of 0.3M. 2.5 volumes of isopropanol was then added to the solution. This layer was mixed into the aqueous layer by gently flicking the tube. The DNA then precipitates out as a clump. The DNA was removed from the solution with a sterile Pasteur pipette. It was then washed by dipping it in 70% ethanol and then in 95% ethanol. It was then air-dried, and dissolved in 0.5 ml of $T_{20}E_1$.

DNA Digestion and Agarose Gel Electrophoresis

Plasmid DNA was digested according to the protocols in Joeseph et al (1989). DNA was analyzed on horizontal gels of 0.8% agarose. Minigels consisted of 50 ml of agarose solution, and were run at 50 volts for about 2 hours. Larger gels consisted of 200 to 300 ml of agarose solution. These gels were electrophoresed at 30 volts overnight, or for shorter periods of time at 100 volts. 10X DNA loading buffer consisted of 0.42% bromophenol blue and 25% ficoll type 400. This was added to digestions so as to

be 1X concentration. The molecular weight standard was λHindIII (25 μg of λHindIII DNA, 4 μl of 5M NaCl, 20 μl 1 M Tris pH 7.8, 0.2 μl of 0.5M EDTA and sterile ddH_2O to a final volume of 900 μl, plus 100 μl 10X loading buffer). 10 μl of this was enough for a gel.

DNA Ligations

Equal amounts of insert and vector were digested with the appropriate restriction enzymes. The restriction enzymes were deactivated by placing all of the digestions at -20°C for 30 minutes. The DNA was diluted to 100 μl with $T_{20}E_1$ and precipitated by the addition of 4 μl of 5M NaCl, 200 μl of ethanol, mixing, and placing at -70°C for 30 minutes. The DNA was pelleted by centrifugation in a microfuge for 15 minutes, washed with 70% and 95% ethanol, and resuspended in 10 μl of $T_{20}E_1$. The ligation mixture contained, in an Eppendorf tube, approximately 0.5 μg of vector, 1 μg of insert DNA, 5 μl of 5X ligation buffer (250 mM Tris pH 7.6, 50 mM $MgCl_2$, 25% polyethylene glycol, 5 mM dithiothreitol), 2.5 μl of 10 mM ATP, and 1 μl of T_4 DNA ligase plus ddH_2O to a final volume of 10 μl. The ligations were floated on an ice-water mixture in a beaker, and the beaker was placed at 4°C overnight. A few microlitres of the ligation mixture was used to transform competent *E. coli*

cells. Clones of insert DNA between the vector *lac* promoter and the alpha fragment of the *lacZ* gene were screened for on plates containing X-Gal. Colonies containing vectors with insert DNA were white on this medium. Colonies containing vectors with no insert DNA appeared blue.

Competent Cell Preparation

A single colony of *E. coli* was used to inoculate a 5 ml culture of LB. The culture was incubated overnight at 30 °C. Four ml was used to inoculate 400 ml of LB media in a sterile 2 L flask the following morning. The culture was incubated at 37°C until the culture reached an OD_{590} of 0.375. The culture was aliquoted into 8 sterile 50 ml pre-chilled Falcon tubes, and left on ice for 5 to 10 minutes. The tubes were then centrifuged at 3,000 rpm at 4°C in a GPR centrifuge with the brake off. The pellets were resuspended in 10 ml of ice-cold $CaCl_2$ solution (60 mM $CaCl_2$, 15% glycerol, 10 mM MOPS, pH 7.0; autoclaved). This was done gently on ice. The cultures were then spun down for 5 minutes at 2,500 rpm a GPR centrifuge at 4°C. The pellets were again resuspended in 10 mL of ice-cold $CaCl_2$ solution. The cells were left on ice for 30 minutes. They were then pelleted at 2,500 rpm, 4°C with the brake off. The pellets were resuspended in 2 ml of ice-cold $CaCl_2$ solution. Cells

were then dispensed in 250 μl aliquots into prechilled sterile Eppendorf tubes, and frozen at -70°C until needed.

Transformation of Competent *E.coli* Cells

10 ng of DNA was aliquoted into 10 to 25 μl of sterile ddH_2O in a sterile Eppendorf tube, and placed on ice. Competent cells were rapidly thawed to 4°C, and 100 μl were aliquoted into the tubes containing the DNA. The tubes were gently mixed, and placed on ice for 10 minutes. The cells were then heat-shocked by placing the tubes into a 37°C water bath for 5 minutes. Then 1 ml of LB was added to each tube, and the tubes were incubated for 1 hour at 37°C. Aliquots were then plated on LB plates containing the appropriate antibiotic.

Preparation of Single-Stranded DNA for Sequencing

A culture of XL-1 containing the phagemids in 3 ml of LB Amp^{50} was incubated at 30°C to O.D. 0.1. Ten μl of MK1307 helper phage (10^{10} phage/μl) was added, and the culture was incubated overnight at 30°C overnight. The culture was transferred to 1.5 ml Eppendorf tubes, and spun down in a microfuge at 14,000 rpm for 10 minutes. The supernatant was again spun for 5 minutes at 14,000 rpm.

Then 250 μl of PEG-NaCl (20% w/v PEG 6000, 2.5M NaCl) was added, and the mixture was vortexed and stored at room temperature for 5 minutes. It was then spun in a microfuge for 10 minutes, and the supernatant was discarded. The tube was spun again for two minutes, and any remaining liquid was removed. The pellet was resuspended in 150 μl of TES (20 mM Tris-HCl pH 7.4, 10 mM NaCl, 0.1 mM EDTA), and 100 μl of phenol-chloroform was added. The mixture was vortexed for 15 to 20 seconds, and left on ice for 3 to 5 minutes. The tube was spun at 14,000 rpm for 5 minutes in a microfuge, and the aqueous layer was transferred to a new tube, and the phenol extraction was repeated. To the aqueous layer was added 20 μl of 5M NH_4Ac, and 500 μl of 100% EtOH. The mixture was placed at -70°C for 40 minutes. It was then spun for 15 minutes at 14,000 rpm in a microfuge. The pellet was washed with 100 μl of 70% EtOH. The pellet was then dried at 37°C. It was resuspended in 8 μl of $T_{20}E_1$. Three μl was checked on a gel.

DNA Sequencing

DNA sequencing was done by MOBIX on an ABI 373A Automated DNA Sequencer. 0.1 to 1 μg of single-stranded DNA (5 μl of above preparation) was used with 4 μl of 0.2 picomole/μl primer.

Manipulation of DNA Sequence

DNA sequence was used to search the Genbank library using the BlastX program (Gish et al., 1993; Altschul et al., 1990). The MSDOS computer program "SEQAID II" version 3.81 was used to translate the DNA sequence to protein sequence. Clustal W, an improved version of Clustal V (Higgins et al., 1991; Higgins and Sharp, 1989) was used to align protein sequences.

Southern Blotting of Restricted DNA

DNA (chromosomal or plasmid) was restricted, and loaded on an agarose gel. The gel was electrophoresed at low voltage overnight, stained with ethidium bromide, and photographed with a metric ruler adjacent to the lane containing the molecular size standards. The gel was trimmed, and the DNA was depurinated by soaking in 0.25 M HCl for 10 minutes with shaking. The acid was rinsed away from the gel, and the DNA was denatured by shaking 2 times with 1.5M NaCl, 0.5M NaOH for 15 minutes each time. The gel was rinsed with ddH_2O, and neutralized by shaking 2 times with 1.5M NaCl, 1M Tris pH 8 for 15 minutes each time.

The DNA in the gel was transferred to a piece of nylon membrane (Biotrans - ICN) slightly large than the gel.

This was done by placing the inverted gel onto an SSC (175.3g NaCl, 88.2g trisodium citrate pH 7 per litre yields 20X SSC) soaked piece of Whatman #1 chromatography paper which had its ends in contact with reservoirs of 10X SSC. Strips of Saran Wrap were placed around the gel and over the SSC reservoirs to prevent contact of the chromatography paper with blotting paper. The nylon membrane was soaked in 2X SSC, and placed on top of the gel. Two pieces of chromatography paper (the same size as the gel) were soaked in 2X SSC and placed on top of the nylon membrane. A stack of paper towels approximately 6 cm high cut to the size of the gel were placed on top of the chromatography paper. A flat piece of plastic was placed on top of the paper towels, and was weighted down with a 500 g weight. The DNA was allowed to transfer from the gel to the membrane overnight. The membrane and the gel were then removed, and the positions of the wells of the gel were marked on the membrane with pencil. The membrane was soaked in 6X SSC for 5 minutes to wash off any remaining pieces of gel, and air dried. The membrane was then baked in an 80°C oven for two hours to fix the DNA to the membrane.

Preparation of Dig-Labelled DNA Probes

Plasmid DNA to be digoxigenin (Dig) labelled was purified with CsCl gradients. Labelling was performed with the Boehringer Mannheim DNA Labelling and Detection Kit - Nonradioactive (cat no. 1093 657). An aliquot of the DNA (5 μl) to be labelled was added to an Eppendorf tube (1 to 3 μg). The DNA was denatured by placing the tube in a boiling water bath for 10 minutes, followed by plunging the tube into ice/ethanol for 3 minutes. 2 μl of the hexanucleotide mixture (random primers for DNA synthesis), 2 μl of dNTP labelling mixture (1 mM dATP, 1 mM dCTP, 1 mM dGTP, 0.65 mM dTTP, 0.35 mM digoxigenin-11-dUTP), 10 μl ddH_2O, and 1 μl (2U) of Klenow DNA polymerase, plus the DNA, to a total volume of 20 μl was mixed and placed into the 37°C water bath overnight. The reaction was stopped by addition of 2 μl of 0.2M EDTA pH 8, and the DNA was precipitated by addition of 2.5 μl of 4M LiCl, 75 μl ethanol, mixing and allowing to precipitate at -70°C for 30 minutes. The DNA was pelleted by centrifuging 15 minutes in a microfuge, and then washed with 70% and 95% ethanol. The labelled DNA was then dissolved in 50 ul of $T_{20}E_1$. Prior to addition to hybridization solution, the probe was denatured by placing it in a boiling water bath for 10 minutes.

Hybridization of Probe with Target DNA

Standard solutions used for prehybridization and hybridization were 20X SSC: 3M NaCl, 300 mM Sodium citrate, pH 7.0; and 5X SSC: 750 mM NaCl, 300 mM Sodium citrate, pH 7.0. The membrane was sealed in a plastic bag with 20 ml of standard prehybridization solution per 100 cm^2 of membrane surface. Prehybridization was carried out at 68°C for 1 hour. The probe was heated in a boiling water bath for 10 minutes to denature the DNA, and was placed on ice immediately afterward. The probe was diluted in hybridization solution, 5-25 ng/ml. Hybridization took place overnight at 68°C. After use, hybridization solution was recovered. The membrane was washed two times at room temperature, with 2X SSC and 0.1% SDS, 5 minutes per wash to remove unbound probe. The membrane was then washed two times in 0.1X SSC with 0.1% SDS at 68°C, fifteen minutes per wash.

Detection of Dig-labelled Probe

The membrane to be probed was washed for 1 minute with shaking in Buffer 1 (100 mM Tris pH 7.8, 150 mM NaCl), washed in Buffer 2 (Buffer 1 containing 0.5% blocking reagent) for 30 minutes with shaking, and then washed again

in Buffer 1. The antibody-conjugate (polyclonal sheep antibody directed against Dig, with alkaline phosphatase coupled, diluted 1/5,000 in 50 ml Buffer 1) was exposed to the membrane with shaking for 90 minutes at room temperature. Unbound antibody was washed away by washing the membrane twice with Buffer 1.

The alkaline phosphatase activity was used to locate the positions of the label on the membrane. The membrane was equilibrated with Buffer 3 (100 mM Tris, 100 mM NaCl, 50 mM $MgCl_2$, pH 9.5) for 2 minutes. The membrane was then incubated (no shaking) with the alkaline phosphatase staining solution (10 ml of Buffer 3, 45 μl nitroblue tetrazolium salt solution (75 mg/ml in 70% dimethyl formamide), 35 μl X-phos solution (5-bromo-4-chloro-3-indoyl phosphate toluidinium salt, 50 mg/ml in dimethylformamide)). Under these conditions, the alkaline phosphatase converts X-phos and the nitroblue tetrazolium salt to a blue precipitate. The membrane was allowed to develop in the dark, with no shaking, until blue bands were visible. The reaction was stopped by soaking the membrane in Buffer 4 (10 mM Tris pH 7.8, 1 mM EDTA pH 8) for 5 minutes.

Rothera Test

The Rothera test is a chemical test for the presence of β-ketoadipate, and is indicative of *ortho*-cleavage of an aromatic such as protocatechuate (Evans, 1947). 5 ml cultures of the strains to be tested were started in LB mc, with the appropriate antibiotic if a plasmid was present. The next day the cultures were spun down, and washed with M9 minimal media, and resuspended in M9. 20 μl of the culture was used to inoculate M9 arabinose (0.1%) protocatechuate (5 mM), with the appropriate antibiotic if a plasmid was present. Thiamine was added (2.5 mM) for *R. meliloti* RmF726. The cultures were incubated overnight with shaking at 30°C. The next day, the cultures were spun down at 4,000 rpm for 5 minutes in a Beckman GPR centrifuge. The cells were resuspended to OD_{675} in 0.02M Tris-Cl pH 8. 2.0 ml of this was placed into a test tube. 0.5 ml of toluene was added, and the cells were incubated with shaking for 1 hour at 30°C. 1 gram of $(NH_4)_2SO_4$ was added, and the cells were vortexed. 1 drop of a fresh aqueous solution of sodium nitroprusside (1%) was added, and the cells were vortexed. Finally, 1 drop of concentrated NH_4 (29%) was added, and the cells were vortexed. A purple colour forming within the tube within 5 minutes indicated a positive test for β-

ketoadipate.

β-Galactosidase Assays

The procedure for β-galactosidase assays was taken from Miller (1972). An aliquot of cells suspended in Z buffer was added to a volume of Z buffer, to a total volume of 4 ml, to a final OD_{675} of about 0.1. One ml of this cell suspension was added to each of 3 tubes, in order to perform the assay in triplicate. To permeablize the cells, 10 μl of $CHCl_3$ was added to each tube, followed by vortexing, and the tubes were placed in a 30°C water bath for a minimum of 10 minutes. An aliquot of Z buffer (1 ml) was used as a blank, and was used to zero the Contron spectrophotometer. A solution of ONPG in Z buffer (4mg/ml) was prepared.

In order to begin the assay, 200 μl of ONPG solution was added to each tube, including the blank. The cell suspensions were vortexed, placed into the water bath, and a timer was started, and when the reactions were stopped the time was recorded. The reactions were stopped by adding 0.5 ml of 1M Na_2CO_3, once a yellow colour was noticeable. The tubes were then spun down for 10 minutes at 4,000 rpm in the GPR centrifuge, and the reaction mixtures were transferred to cuvettes, and OD's were measured at 420 nm using the spectrophotometer. The calculation of Miller Units was:

Miller Units = 1000 x A_{420}/time in minutes x A_{675}.

Z buffer consisted of (per litre): 16.1g of $Na_2HPO_4 \cdot 7H_2O$, 5.5g of $NaH_2PO_4 \cdot H_2O$, 0.75g of KCl, 0.246 g $MgSO_4 \cdot 7H_2O$, and 2.7 ml of β-mercaptoethanol. Z buffer was stored at room temperature without the β-mercaptoethanol, which was added just prior to use.

Preparation of Cell Extracts for Protocatechuate 3,4-Dioxygenase Assays

Five ml overnight cultures in LB mc were used to inoculate 200 ml cultures of the appropriate media, which was M9 Succinate (15 mm), and M9 Arabinose (15 mm) Protocatechuate (5 mM). 2.5 mM thiamine was added for RmF726. Cultures were washed with 1X M9 media prior to inoculation of larger cultures. The large cultures were incubated with shaking overnight at 30°C to late log phase. Some of each culture was struck out on LB media to verify that the cultures were pure. The cells were pelleted by using a Sorvall RC-2 centrifuge (GCA rotor) with 250 ml bottles by spinning for 20 minutes at 7,000 rpm and at 4°C. The cells were washed twice in 20 mM Tris pH 7.8, 1 mM $MgCl_2$. The cells were weighed, and resuspended in buffer, 4 ml of buffer per gram of cells. Aliquots of the cells were frozen at -80°C until used in assays for protocatechuate

3,4-dioxygenase.

For sonication, the aliquots were thawed, and 10 μl of 0.1M dithiothreitol were added to each ml of cells. Cells were disrupted with a Heat Systems Sonicator Ultrasonic Processor XL. Cells and disrupted cells were kept at 4°C during and after sonication. The cooling water bath of the sonicator was turned on prior to sonication, and sonication was started when the temperature had stabilized at 4°C. One ml of cells were placed in 15 ml conical polystyrene Falcon tubes. The tubes were placed vertically in the sonicator bath without the caps in order to prevent over-heating. The settings on the sonicator were as follows: pulse time, 30 seconds on/ 20 seconds off; time, 1 hour; power level, 8. The extracts were then centrifuged in the same tubes at 4,500 rpm for 10 minutes at 4°C. This was followed by a spin at 14,000g for 15 minutes in a microfuge. The extracts were examined under phase-contrast microscopy to determine the degree of disruption. When few intact cells were visible, the extracts were not sonicated further.

Protein Determination

The Bradford method was used to determine protein concentration in the cell extracts. The assay solution was prepared by diluting the concentrated dye with 4 volumes of

ddH_2O. The solution was then passed through a Whatman #1 filter. Some of the extract in question was diluted 1 in 10 with distilled water, and 10 to 30 μl was added to 5 ml of the Bradford assay solution. The solutions were mixed, and incubated for 5 minutes at room temperature. The absorbance at A_{595} nm was determined using a Contron Uvikon 930 double beam spectrophotometer, with glass cuvettes. Protein concentration was determined by converting the A_{595} to μg protein using a protein standard curve, and factoring in the volume and dilution of the aliquot. The protein standard curve was constructed using bovine serum albumin as a standard.

Assay for Protocatechuate 3,4-Dioxygenase Activity

The method used to assay for protocatechuate-3,4-dioxygenase was that of Durham et al. (1980). The assay mixture for measurement of protocatechuate 3,4-dioxygenase activity contained, in a total volume of 3 ml, 150 μmol of Tris-HCl pH 8.5 and 0.5 μmol of protocatechuate. The assay was started by the addition of cell extract. The temperature of the reaction was maintained at 30°C by using a waterbath connected to the Contron Uvikon 930 double-beam spectrophotometer, and the reaction was monitored by following the reduction in absorbance at 293 nm.

Protocatechuate absorbs at 293 nm, while the product of the reaction, β-carboxy-*cis,cis*-muconate does not. The molar extinction coefficient of protocatechuate was taken to be 14,560 at 293 nm (Parke et al., 1991). One unit of enzyme activity is defined as the amount of enzyme which oxidizes 1 nanomole of protocatechuate per minute. Specific activity was expressed as units per milligram of protein.

Chapter 3: Results

I. Pca Mutants of *Rhizobium meliloti*

Tn*5* Mutagenesis of Rm1021

Tn5 mutagenesis of *Rhizobium meliloti* was carried out by mating the plasmid pRK602 from *E. coli* MT607 into Rm1021. Plasmid, pRK602, is incapable of replicating in the *Rhizobium* host. Colonies which were resistant to streptomycin and neomycin were isolated. These were cells in which Tn*5* had inserted into one of the three *R. meliloti* replicons. The mating was allowed to occur over a four hour period prior to plating to minimize the occurrence of sibling mutants. Streptomycin (200 μg/ml) and Neomycin (200 μg/ml) resistant colonies were then patched from LB plates onto M9 minimal media plates containing one of protocatechuate, glucose, and quinate as sole carbon sources, and were also patched on LB Sm Nm plates as a positive control. A total of 2,400 colonies from the Tn*5* mutagenesis were patched onto the above plates. From this experiment, a total of five Tn*5* mutants which were unable to grown on protocatechuate (Pca^-) were isolated. All five of these mutants were able to grow with glucose as a sole carbon and energy source. This indicates that the five

mutants were not deficient in a biosynthetic gene required for growth on minimal media.

Since protocatechuate is metabolized to succinate and acetyl-CoA, it was possible that the mutants were deficient in an enzyme required for growth with succinate as a sole carbon and energy source. In such a case it would be possible that the gene that was knocked out would not be specific to the Pca pathway. Of the five mutants obtained, all but one were able to grow on succinate as a sole carbon and energy source. The remaining four Pca$^-$ Tn*5* mutants were designated RmG867, RmG877, RmG878, and RmG879.

Determination of Linkage of Tn*5* to Pca Phenotype

To ensure that the Pca$^-$ phenotype of the above mutants was in fact due to the insertion of the Tn*5* into a *pca* gene, linkage of the Nm^R marker of the Tn*5* to the Pca phenotype was determined. Using a generalized transducing phage, lysates of the four Tn*5* mutants were used to transduce Nm^R into the wildtype Rm1021. For each transduction experiment, 50 colonies were patched on M9 protocatechuate and LB neomycin media. In each case, all 50 transductants failed to grow on protocatechuate but grew on Lb Nm. Thus the neomycin-resistance marker of the Tn*5* was 100% linked to the Pca$^-$ phenotype.

Mapping of *pca*::Tn5 insertions *pca*-867 and *pcaG*879 onto the Second Megaplasmid

The strain RmF726, which has a large region of the second megaplasmid deleted, has a Pca$^-$ phenotype (Charles and Finan, 1991). This implies that a *pca* gene cluster is located within the deleted region. An experiment was performed to determine if any of the above *pca*::Tn*5* insertions were linked to other insertions within this region. Phage grown on the four *pca*::Tn*5* insertion mutants was used to transduce Nm^R into RmF153 carrying a Tn*5-132* insertion in *dctB* (the position of *dctB* on the pRmeSu47b megaplasmid had previously been deduced). Transductants were screened for loss of the oxytetracycline resistance (OT^R) encoded by Tn*5-132*. The results (Table 3-1) indicate that the Tn*5* insertions in RmG877 and RmG878 are not linked to *dctB*, while the Tn*5* insertions in RmG867 and RmG879 are 54% and 39% linked to *dctB*, respectively.

Table 3-1: Linkage of *pca*::Tn*5* Insertions to *dctB*

Donor	OT sensitive recipients	Linkage to *dct*B12::Tn*5-132*
RmG867	54/100	54%
RmG877	0/75	0%
RmG878	0/75	0%
RmG879	23/59	39%

The donor strains were Rm1021 *pca*::Tn*5* insertion strains, and the recipient strain was RmF153(*dct*B::Tn*5-132*). Nm^R transductants were selected on LB Nm (200μg/ml) and were screened for loss of OT resistance on LB OT (0.3 μg/ml).

The exact position of *pca* insertion *pca-867*::Tn*5* was determined by a three-factor transductional cross with strain RmF460, which carried the megaplasmid insertions Ω5071::Tn*5-132* *thi*502::Tn*5-233* (Yarosh, 1989). The results indicate that the *pca* locus is located counterclockwise to the *dct* locus (Table 3-2; Figure 3-1).

Isolation of Cosmids Complementing *pca*::Tn*5* Mutants

Cosmids from the Rm1021 pLAFR1 clone bank (Freidman et al. 1982) were isolated which complemented the Pca^- growth phenotype of strains carrying the *pca*::Tn*5* insertions *pca-867*::Tn*5*, *pca-877*::Tn*5*, and *pca-879*::Tn*5* (Table 3-3).

Table 3-2: Three Factor Mapping of *pca-867*::Tn*5*

Class of Transductant	Fraction of Trans. in Class	Percentage of Trans. in Class
Nm^R OT^R GS^R	2/175	1%
Nm^R OT^S GS^R	104/175	59%
Nm^S OT^R GS^R	40/175	23%
Nm^S OT^S GS^R	29/175	17%

The recipient strain was RmG867 (*pca*-867::Tn*5*), and the donor was RmF460 (Ω5071::Tn*5-132* *thi*502::Tn*5-233*. GmSp resistant transductants were seleicted on LB Gm (Gm 20 μg/ml Sp 100 μg/ml). Transductants were screened for the loss of the unselected marker OT^R and for loss of Nm^R by patching transductants on LB OT (0.3 μg/ml) and on LB Nm (200 μg/ml).

Subsequently, sequencing revealed that the Tn*5* insertion in strain RmG879 was in the *pcaG* gene. Thus the *pca-879* allele was designated *pcaG879*::Tn*5*. An unsuccessful attempt was made to complement the Pca$^-$ phenotype of RmG878. DNA sequence analysis subsequently showed that the transposon lay within a gene with very high homology to the catalase gene (*katE*) of *Brucella abortus*. This allele was therefore designated *katE878*::Tn*5*.

Two classes of cosmid were isolated which complemented *pca-867*::Tn*5* and *pcaG879*::Tn*5*, as judged on the basis of *Eco*R1 digests (Table 3-4). Both of these cosmids also complemented the Pca$^-$ phenotype of the *pcaG879*::Tn*5*

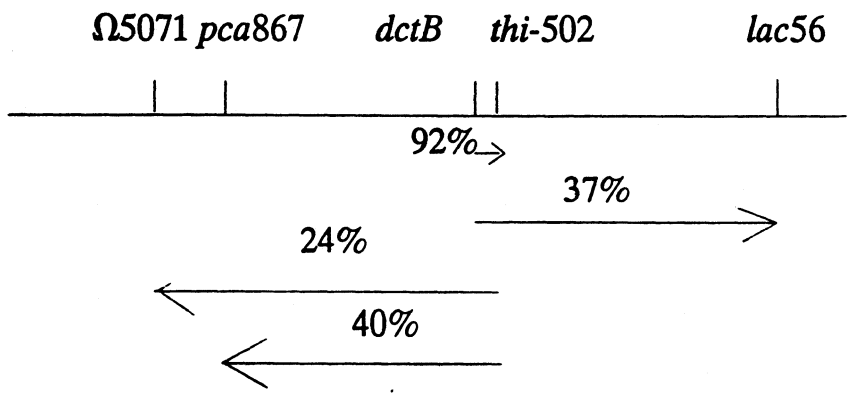


Figure 3-1: Deduced Map position of *pca-867*::Tn*5* on the pRmeSU47b Megaplasmid relative to *dctB*, *thi-502*, and *lac56*. Map position was determined by a three-factor cross. The donor was RmG867 (*pca-867*::Tn*5*) and the recipient was RmF460 (Ω5071::Tn*5-132* *thi502*::Tn*5-233*. Gm Sp^R transductants were selected, and screened for Ot^R and Nm^R. Other markers and locations are from Charles (1990).

mutant, and were designated pTH178 and pTH179. The smaller of two cosmids (pTH178) was used for further work. Three classes of cosmid were isolated which only complemented the Pca$^-$ phenotype of *pca-877*::Tn*5*, and were designated pTH180, pTH181, and pTH182 (Table 3-5).

Table 3-3: Complementation of *pca* Tn*5* Mutants by Cosmids

Cosmid	RmG867	RmG877	RmG878	RmG879	RmF726
pTH178	+	-	-	+	-
pTH179	+	-	-	+	-
pTH180	-	+	-	-	-
pTH181	-	+	-	-	-
pTH182	-	+	-	-	-

Ability of cosmids to complement the Pca$^-$ phenotype of *pca*::Tn*5* insertion mutants was tested. The plasmids were mated into the *R. meliloti* Tn*5* mutants from DH5α cells containing the plasmids using MT616 as a helper. Transconjugants were selected for on LB Streptomycin (200 μg\ml) Tetracycline (2 μg/ml). Transconjugants were screened for growth on M9 minimal media plates with 5 mM protocatechuate as sole carbon source. Plates were scored for growth after 5 days incubation at 30°C. As a positive control, transconjugants were also screened for growth on M9 Glucose (15 mM). All transconjugants grew on M9-Glucose. +, growth like wildtype on M9 protocatechuate; -, no growth on M9 protocatechuate.

Table 3-4: Sizes of EcoR1 Fragments of Plasmids pTH178 and pTH179

pTH178	pTH179
23 kb	23 kb
5 kb	5 kb
5 kb	5 kb
4.1 kb	4.1 kb
3.7 kb	3.7 kb
2.6 kb	2.6 kb
2.3 kb	2.3 kb
2.3 kb	2.3 kb
0.7 kb	1.4 kb
0.6 kb	0.7 kb
0.5 kb	0.6 kb
	0.5 kb

The sizes of the bands were determined by gel electrophoresis. A *Hind*III digest of λ DNA was used as a molecular weight standard. The sizes of the fragments of the plasmids were estimated from a plot of the logarithm of the size of the λ-*Hind*III fragments versus distance migrated.

Table 3-5: Sizes of EcoR1 fragments in plasmids pTH180, pTH181 and pTH182

pTH180	pTH181	pTH182
23 kb	23 kb	23 kb
9.4 kb	9.4 kb	9.4 kb
6.3 kb	6.3 kb	6.3 kb
3.4 kb	2.0 kb	4.8 kb
2.0 kb	2.0 kb	3.4 kb
2.0 kb	1.1 kb	2.0 kb
0.8 kb	0.8 kb	2.0 kb
0.5 kb	0.5 kb	1.1 kb
0.4 kb	0.4 kb	0.8 kb
		0.5 kb
		0.4 kb

The sizes of the fragments were estimated using the same method described in the legend of Table 3-4.

Attempt to Complement *R. meliloti pca*::Tn*5* Insertion Mutants Using *pca* Genes from *Acinetobacter calcoaceticus*

To determine which *pca* genes had been inactivated an attempt was made to complement the *R. meliloti pca*::Tn*5* insertion mutations with clones carrying *Acinetobacter calcoaceticus pca* genes under the control of the *lac* promoter of pRK415-1. Plasmids pARO2 and pARO4 carry the *pcaCHG* and *pcaFBDC* genes of *A. calcoaceticus*, respectively (Parke et al., 1991). These plasmids were previously used to complement *pca* mutants of *Rhizobium leguminosarum*. These plasmids were mated into the *pca*::Tn*5* insertion mutants, and the resulting strains were examined for growth on M9 protocatechuate. Partial complementation was observed for RmG879 (*pcaG879*::Tn*5*) with the plasmid pARO2. This is consistent with the DNA sequence data (see below) which indicates that the Tn*5* insertion in RmG879 is in the *pcaG* gene. The results are summarized in Table 3-6.

Table 3-6: Growth of *pca*::Tn5 Insertion Mutants with Plasmids pARO2 and pARO4

Strain	no plasmid	plasmid pARO2	plasmid pARO4
Rm1021	+	not done	not done
RmG867	-	-	-
RmG877	-	-	-
RmG878	-	-	-
RmG879	-	+/-	-

Plasmids pARO2 and pARO4 were mated into the *R. meliloti* *pca*::Tn5 mutants from *E. coli* DH5α using MT616 as a helper. Transconjugants were isolated on LB Sm (200 μg\ml) Tc (2 μg/ml), and were streaked on M9 minimal media plates with 15 mM protocatechuate as sole carbon source. As a positive control, transconjugants were also streaked on M9 glucose (5 mM). Growth was scored after 5 days incubation at 30°C. +, full growth (as wild-type); +/-, indicates partial growth; - indicates no growth.

Southern Analysis of *pca*::Tn5 Insertions in RmG867 and RmG879

In order to determine the locations of the Tn5 insertions *pca-867* and *pcaG879* in the DNA cloned in pTH178, and to establish that the cloned DNA is contiguous to the *pca* region deleted in RmF726, a Southern analysis of the two insertions was undertaken. Probing digests of total restricted DNA from RmG867 and RmG879 with DIG-labelled pTH178 revealed a number of bands in addition to those corresponding to the DNA in pTH178 (Figure 3-2). Thus, the

position of the insertions *pca-867* and *pcaG879* could not be determined. Another Southern analysis was undertaken, this time using DIG-labelled Tn5 (pGS220) as a probe (Figure 3-3). The results for this experiment are given in Tables 3-7 and 3-8. This revealed that the two insertions were both located within a 5 kb EcoR1 fragment about 2 kb apart. (See Figure 3-4). They also lay within a 5 kb Pst1 fragment.

Table 3-7: Southern Analysis of RmG867 using DIG-labelled Tn*5*

*Pst*I	*Pst*I/*Eco*RI	*Eco*R1	*Hind*III/ *Eco*RI	*Hind*III
5.8 kb	3.5 kb	12 kb	4.0 kb	16 kb
2.2 kb	2.2 kb		3.7 kb	4.2 kb
1.1 kb	1.1 kb		3.7 kb	3.6 kb
1.0 kb	1.0 kb			
0.8 kb	0.8 kb			

Sizes of the bands were determined using a *Hind*III digest of λ, as described earlier (Table 4).

Table 3-8: Southern Analysis of RmG879 using DIG-labelled Tn*5*

*Pst*I	*Pst*I/*Eco*RI	*Eco*RI	*Hind*III/ *Eco*RI	*Hind*III
3.2 kb	3.2 kb	12 kb	5.8 kb	22 kb
2.2 kb	2.2 kb		3.7&kb	3.6 kb
1.1 kb	1.3 kb		1.8 kb	2.1 kb
1.0 kb	1.1 kb			
0.8 kb	1.0 kb			
	0.8 kb			

Sizes of the bands were determined using a *Hind*III digest of λ, as described earlier (Table 3-4).

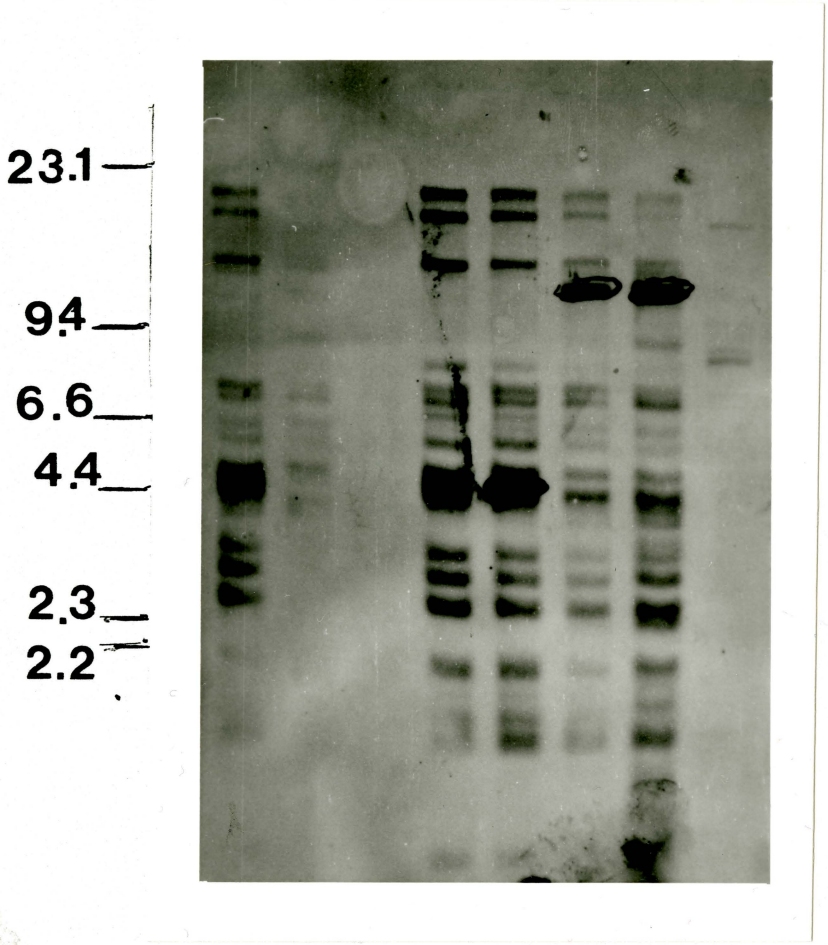


Figure 3-2: Southern analysis of RmG867 (*pca-867*::Tn*5*) and RmG879 (*pcaG879*::Tn*5*) using DIG-labelled pTH178 as a probe. Lanes are as follows: A, Rm1021 (wildtype) *Eco*R1; RmF726 (large megaplasmid deletion) *Eco*R1; C, RmG877 (Rm1021 *pca-877*::Tn*5*) *Eco*R1; D, RmG878 (Rm1021 *katE878*::Tn*5*) *Eco*R1; E, RmG867 (Rm1021 *pca-867*::Tn*5*) *Eco*R1; F, RmG879 (Rm1021 *pcaG879*::Tn*5*) *Eco*R1.

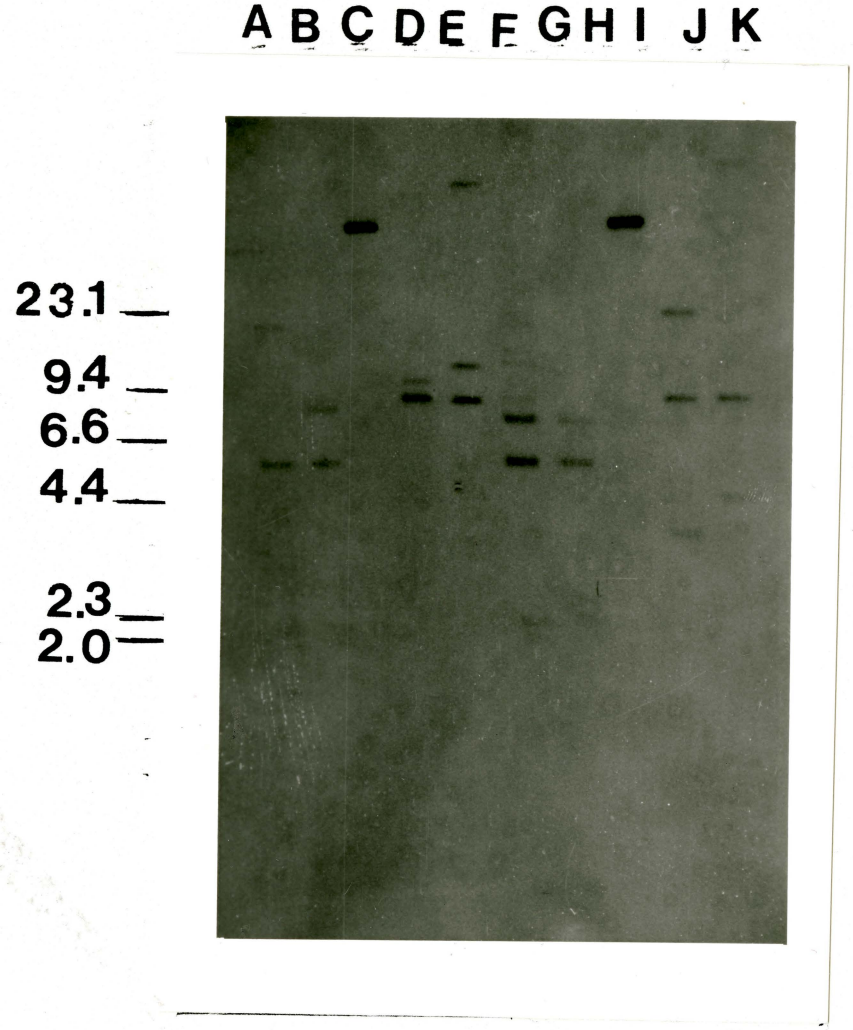


Figure 3-3: Southern Analysis of RmG867 (*pca-867*::Tn*5*) and RmG879 (*pcaG879*::Tn*5*) using DIG-labelled pGS220 (Tn*5*) as a probe. Lanes are as follows: A, Rm1021 *Eco*R1; B, RmG867 *Pst*I; C, RmG867 *Pst*I/*Eco*R1; D, RmG867 *Eco*R1; E, RmG867 *Hind*III/*Eco*R1; F, RmG867 *Hind*III; G, RmG879 *Pst*I; H, RmG879 *Pst*I/*Eco*R1; I, RmG879 *Eco*R1; J, RmG879 *Hind*III/*Eco*R1; K, RmG879, *Hind*III.

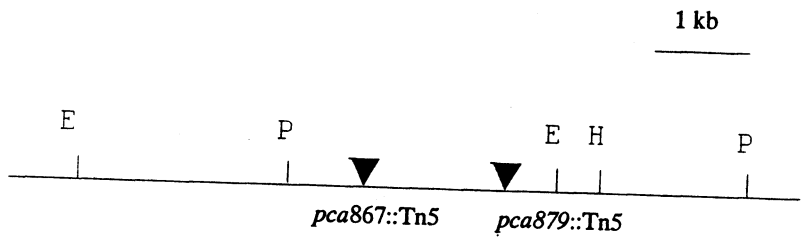


Figure 3-4: Map positions of *pca-867*::Tn5 and *pcaG879*::Tn5 determined by southern analysis.

Plant Phenotypes of the *pca*::Tn*5* Mutants

Plant phenotypes of the mutants were determined by following the inoculation of alfalfa seedlings and growth in a nitrogen-deficient medium for 25 days. The dry weights of the plants were then determined (Table 3-9). The results indicate that they show a wildtype symbiotic phenotype.

Table 3-9: Plant Phenotypes of *pca*::Tn5 Mutants

Strain	Dry Weight (grams) per plant	% Wildtype
Rm1021	$0.225 +/- 4 \times 10^{-3}$	100
RmG867	$0.217 +/- 2 \times 10^{-3}$	99
RmG877	$0.197 +/- 2 \times 10^{-3}$	88
RmG878	$0.228 +/- 2 \times 10^{-3}$	101
RmG879	$0.176 +/- 7 \times 10^{-5}$	78
uninoculated	$0.0563 +/- 6 \times 10^{-5}$	25

The dry weights of the plants were determined after growth of inoculated plants for 25 days, in triplicate. Dry weights were determined in triplicate after cutting the plants at the level where the root began, and drying for 1 week at 70°C.

Rothera Phenotypes of the Pca Mutants

The Rothera test is a chemical test which can be used to determine the cleavage product of an aromatic compound.

A yellow colour is indicative of a meta cleavage, with the meta-cleavage product being a hydroxymuconic semialdehyde; a purple colour is indicative of an ortho cleavage product, namely β-ketoadipate. *R. meliloti* Rm1021 gives a purple colour, indicative of an ortho cleavage (Figure 1-1). Pca$^-$ mutants which are blocked at steps prior to the production of β-ketoadipate would be expected to give Rothera-negative results, whereas those mutations which affect steps after the formation of β-ketoadipate would be expected to give Rothera-positive results. Table 3-10 gives the Rothera results for the *pca*::Tn*5* mutants.

The results indicate that the strains RmG867 and RmG879 are blocked in the pathway prior to the formation of β-ketoadipate, while RmG877 and RmG878 appear to be blocked after the formation of β-ketoadipate. The results also show that pTH178 complements the Rothera$^-$ phenotype of RmG867, RmG879, and RmF726.

pTH178 Restores Protocatechuate 3,4-Dioxygenase Activity to *pca-867* and *pcaG879*

The enzyme protocatechuate 3,4-dioxygenase is the first enzyme in the breakdown of protocatechuate. Pca dioxygenase activity was assayed by measuring the conversion of protocatechuate to carboxy-cis,cis-muconate, indicated

Table 3-10: Rothera Results for Pca Tn5 Mutants

Strain	Rothera Result
RmG867	-
RmG877	+
RmG878	+
RmG879	-
Rm1021	+
RmG867 pTH178	+
RmG879 pTH178	+
RmF726 pTH178	+
RmF726	-

Strains were grown up in LB (with 5 μg/ml tetracycline if a plasmid is present) overnight. Cultures were then spun down, and washed in M9. The cultures were then resuspended in M9. 10 μl were used to inoculate 5 ml of M9 Arabinose (10%) Protocatechuate (5 μM) (5 μg/ml of tetracycline if a plasmid was present). 2.5 mM thiamine was added for strain RmF726 (a *thi* auxotroph). Cultures were incubated overnight at 30°C. They were then spun down, resuspended in 0.02M Tris-HCl pH 8.0 to an OD_{600} of 1.0. 2 ml of the resuspended cells was taken, and 0.5 ml of toluene was added, and the tubes were incubated with shaking for 1 hour. 1 g of $(NH_4)_2SO_4$ was added, and the tubes were vortexed. 1 drop of freshly made 1% sodium nitroprusside was added, followed by

1 drop of concentrated NH_3. A purple colour forming within 5 minutes was scored as a positive result (+), with a - indicating no colour development within 5 minutes.

Table 3-11: PCA Dioxygenase Activity in Mutants With and Without pTH178.

Strain	**Genotype**	**Units of Dioxygenase**
RmF726	Rm1021 ΔΩ5149-5079::Tn*5-233*	0
RmF726 pTH178	Rm1021 ΔΩ5149-5079::Tn*5-233* (pTH178)	68
RmG867	Rm1021 *pca-867*::Tn*5*	0
RmG867 pTH178	Rm1021 *pca-867*::Tn*5* (pTH178)	26
RmG879	Rm1021 *pcaG879*::Tn*5*	0
RmG879 pTH178	Rm1021 *pcaG879*::Tn*5*	18
Rm1021	Rm1021 wildtype	10

Results of assay for protocatechuate 3,4-dioxygenase. Cells were grown in M9 arabinose (0.1%) protocatechuate (5mM), with tetracycline (5 μg/ml) if a plasmid was present. Cells were spun down and washed. Extracts were prepared by sonication. Protocatechuate 3,4-dioxygenase activity was assayed in 3 ml final volumes containing 150 μmol of Tris pH 8.5 and cell extract. Reactions were initiated by adding 0.5 μmol of protocatechuate. The extent of reaction was determined by following the decrease in absorbance at 293 nm. One unit of activity is defined as the amount of enzyme required to convert 1 nanomol of protocatechuate to β-carboxy-*cis,cis*-muconate per minute per mg protein.

by the decrease in absorbance at 293 nm (Table 3-11). The mutants RmF726, RmG867 and RmG879 lack this activity. Activity was restored by the plasmid pTH178.

II. Subcloning and Mutagenesis of the Megaplasmid *pca* Locus

Tn*5*-B20 Mutagenesis of pTH178

Cosmid pTH178 complements the Pca$^-$ phenotype of *pca*::Tn*5* insertions located on the pRmeSU47b megaplasmid (see above). In order to study the expression of the *pca* genes, we wished to obtain *pca-lacZ* gene fusions. The Tn*5* derivative, Tn*5-B20*, generates transcriptional fusions between *lacZ* and the gene in which it inserts. Tn*5-B20* insertions in pTH178 were isolated, and insertions in cloned DNA were selected by examining *Eco*R1 digests of the insertions in pTH178, and selecting those derivatives in which the Tn*5-B20* was located in cloned insert DNA. Since the plasmid pTH178 complements the Rothera phenotype of RmF726, we looked for insertions in pTH178 which did not complement the Rothera phenotype of RmF276. Out of fifty insertions that were screened, four, designated as *pca-6*::Tn*5-B20*, *pca-8*::Tn*5-B20*, *pca-14*::Tn*5-B20*, and *pca-39*::Tn*5-B20* were found to be Rothera negative. The pTH178 plasmids carrying these four insertions were transferred into RmG212 (Rm1021 *lac*$^-$). Following growth in M9 Succinate

and M9 Protocatechuate-Arabinose, transconjugants were assayed for β-galactosidase activity (Table 3-12). Two of these insertions, *pca-6*::Tn*5-B20* and *pca-39*::Tn*5-B20* were found to be induced by protocatechuate. Subsequent mapping of insertions indicated that these two insertions were identical. When these four *pca*::Tn*5B-20* insertions were recombined into the chromosome of Rm1021, the resulting strains were unable to grow on protocatechuate. Introduction of pTH178 into these mutants restored their ability to grow on protocatechuate (data not shown).

Subsequently, sequencing out from the Tn*5-B20* insertions (discussed in III. Sequencing of the *pca* region of *R. meliloti*) revealed which genes the insertions were in. Thus, *pca8*::Tn*5-B20* will be referred to as *pcaC8*::Tn*5-B20*, *pca14*::Tn*5-B20* will be referred to as *pcaD14*::Tn*5-B20*, and *pca39*::Tn*5-B20* will be referred to as *pcaH39*::Tn*5-B20*. The direction of transcription inferred from the β-galactosidase activities of the insertions correlated with that inferred from sequence data.

Complementation of Tn5 Mutants by Tn*5-B20* Insertions on pTH178

Plasmids with the *pca*::Tn*5-B20* insertions were mated in to RmG867 and RmG879 to determine if the plasmids could still complement the Pca$^-$ phenotypes. The pTH178 plasmids

Table 3-12: β-galactosidase Activity of RmG212 carrying pTH178::*lacZ* insertions after growth in M9 Succinate and M9 Protocatechuate-Arabinose

Insertion	**Activity in Succinate**	**Activity in Ara-Pca**	**Fold Induction**
pTH178	8.9 +/- 2.8	55 +/- 42	6.2
pTH178 *pcaH6*	340 +/- 56	3500 +/- 420	10
pTH178 *pcaC8*	170 +/- 10	160 +/- 40	0.94
pTH178 *pcaD14*	32 +/- 1.7	41 +/- 5.3	1.3
pTH178 *pcaH39*	350 +/- 36	5780 +/- 210	16

Cultures grown in LB tetracycline (5 μg/ml) were subcultured into M9 succinate (15mM) tetracycline (5 μg/ml) and into M9 arabinose (0.1%)-protocatechuate (5 mM) tetracycline (5 μg/ml) and following 24 hours growth with shaking at 30°C were assayed for β-galactosidase activity according to Miller (1972). Activities are expressed in Miller Units, which were determined as follows: Miller Units = 1000 x A_{420}/time x A_{675}.

Table 3-13: Growth Phenotypes of *pca*::Tn5 Chromosomal Insertions with Tn*5-B20* insertions on pTH178

Plasmid	**RmG867 *pca-867*::Tn5**	**RmG879 *pcaG879*::Tn5**
pTH178	+	+
pTH178*pcaC8*:: Tn*5-B20*	S + L	S + L
pTH178*pcaD14*:: Tn*5-B20*	-	S + L
pTH178*pcaH39*:: Tn*5-B20*	S + L	+

Plasmids were mated into the Tn5 mutants RmG867 and RmG879 from the *E. coli* strain DH5α using MT 616 as a helper. Transconjugants were selected for on LB streptomycin (200

μg/ml) tetracycline (2 μg/ml). Transconjugants were then struck out on M9 protocatechuate (5 mM) to check for growth after 5 days at 30°C. +, indicates wildtype growth; -, indicates no growth; S+L, indicates that small and large colonies were visible on the plate, suggesting that recombination between the plasmid and the megaplasmid was occurring in these instances.

carrying Tn*5-B20* insertions 6 and 39 still complemented *pca*867::Tn*5* cleanly, indicating that the *pca* gene(s) affected by Tn*5-B20* insertions Ω6 and Ω39 is different than the pca gene(s) affected by *pca*867::Tn*5* (Table 3-13). In other cases, small and large colonies were observed on the M9 protocatechuate plates, suggesting that recombination was occurring between the plasmids and the megaplasmid.

Subcloning of the *pca* Region from pTH178

Southern analysis indicated that the two Tn*5* insertions *pca*867::Tn*5* and *pcaG879*::Tn*5* were located within a 5 kb *Eco*R1 fragment, and also lie within a 5 kb *Pst*1 fragment (Figure 3-4). They also lie within a 3 kb *Pst*1/*Eco*R1 fragment that lies within both the 5 kb *Eco*R1 fragment and the 5 kb *Pst*1 fragment. The plasmid pTH178 contains two 5 kb *Eco*R1 fragments, and one 5 kb *Pst*1 fragment. Of the two 5 kb *Eco*R1 fragments, only one contained a 3 kb *Pst*1/*Eco*R1 fragment. Both 5 kb *Eco*R1 fragments and the 5 kb *Pst*1 fragment were subcloned from

pTH178 into pUC119. The two 5 kb *Eco*R1 fragments in pUC119 were designated pTH185 and pTH186. The 5 kb *Pst*1 fragment in pUC119 was designated pTH215 (See Figure 3-5).

Of the two *Eco*R1 fragments, only one had a *Pst*1/*Eco*R1 fragment that was 3 kb in size. This fragment, in pTH185, was subcloned into pRK7813, a vector that can replicate in *Rhizobium*, in both orientations. These plasmids were designated pTH187 and pTH188. The 5 kb *Pst*1 fragment was also cloned into pRK7813 in both orientations, yielding pTH193 and pTH194.

pTH194 Complements the Rothera Phenotype of pca::Tn*5* Mutants but fails to Complement the Growth Phenotype

The plasmids pTH187, pTH188, pTH193 and pTH194 were checked for the ability to complement the Pca growth phenotype of the mutants *pca*867::Tn*5* and *pcaG879*::Tn*5*, and the deletion mutant RmF726. The results of this experiment are summarized in Table 3-14. Only pTH178 fully complemented the Tn*5* insertion mutants Pca^- phenotype; the other plasmids partially complemented the Pca^- phenotype. None of the plasmids complemented the Pca^- phenotype of RmF726. These plasmids were also examined for the ability to complement the Rothera phenotypes of the pca mutants.

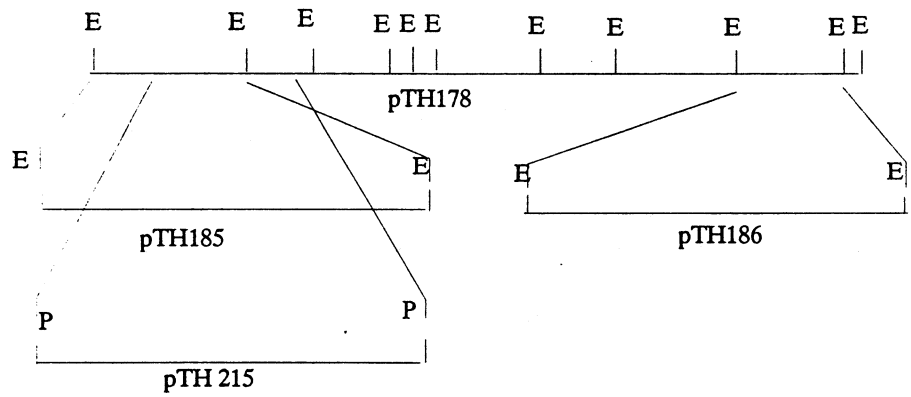


Figure 3-5: Subcloning of 5kb EcoRI fragments and 5kb PstI fragment from pTH178 into pUC119.

Table 3-14: Complementation of Pca⁻ Growth Phenotype of RmG867 and RmG879 by Plasmids Containing 5 kb *Eco*RI and *Pst*I Fragments

Strain	pTH 178	pTH 187	pTH 188	pTH 193	pTH 194
RmF726 ΔΩ5149-5079::Tn*5-233*	-	-	-	-	-
RmG867 *pca-867*::Tn*5*	+	+/-	+/-	+/-	+/-
RmG879 *pcaG879*::Tn*5*	+	+/-	+/-	+/-	+/-

Plasmids were mated into Pca⁻ mutants from DH5α using MT616 as a helper. Transconjugants were selected for on LB streptomycin (200 μg/ml) tetracycline (2 μg/ml). Transconjugants were struck out on M9 protocatechuate (5 mM) to check for growth. +, indicates wildtype growth; +/-, indicates partial growth (less than wildtype); -, indicates no growth. Plates were incubated for 5 days at 30°C.

One of the plasmids, pTH194, fully complemented the Rothera phenotypes of both *pca-867*::Tn*5* and *pcaG879*::Tn*5*, as well as the large deletion mutant RmF726 (Table 3-15). The Rothera result in these cases was much more intense than for wildtype or for that of the mutants with pTH178. This is likely due to increased expression of the *pca* genes on pTH194, driven by p*lac*. Sequence analysis revealed that transcription of *pca* genes on pTH194 was in the same direction as p*lac*.

Table 3-15: Complementation of the Rothera Phenotype of *pca* Insertion and Deletion Mutants by Subclones of pTH178

Strain	Plasmid pTH187	Plasmid pTH188	Plasmid pTH193	Plasmid pTH194
RmG867	-	-	-	++
RmG879	-	-	-	++
RmF726	-	-	-	++

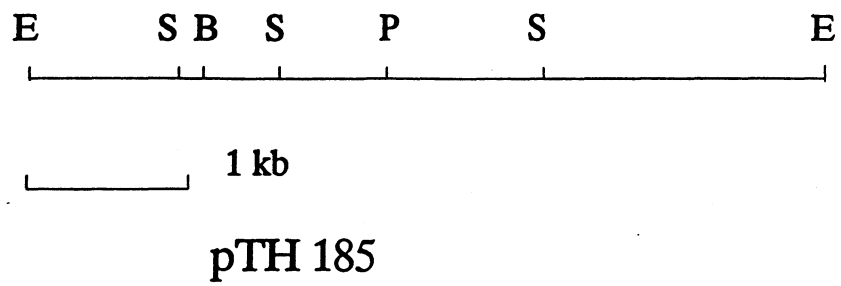
Transconjugants were obtained using the same method as previously (see Table 3-14). Rothera test was performed as described previously (see Table 3-10). -, indicates a negative Rothera test; ++, indicates a positive Rothera result stronger than for wildtype.

Restriction Mapping of the Subcloned *pca* Region

Restriction maps were generated of the regions subcloned in pTH185 and pTH215. The data are summarized in Figure 3-6.

Mapping of Tn5-B20 Insertions in pTH 178

The *pca*::Tn5-B20 insertions were mapped in pTH178 using the restriction digest information from the plasmids pTH185 and pTH215 (see above). All of the *pca*::Tn*5-B20* insertions were located in the region contained within both of pTH185 and pTH215, as *Eco*RI digests showed that the insertions were located within a 5 kb *Eco*RI fragment, and



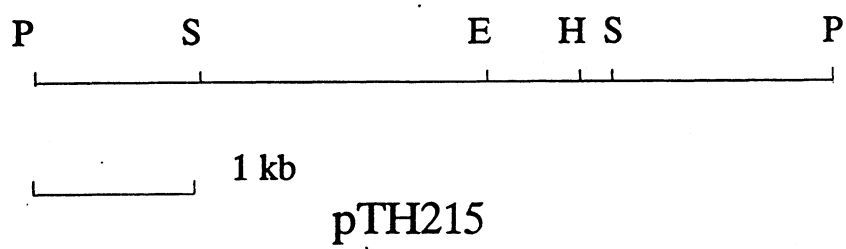


Figure 3-6: Restriction maps of pTH185 (5 kb *Eco*RI fragment #1 from pTH178) and pTH215 (5 kb *Pst*I fragment from pTH178). E indicates an *Eco*RI site, S indicates a *Sal*I site, B indicates a *Bgl*II site, P indicates a *Pst*I site, and H indicates a *Hind*III site.

within a 4.6 kb *Bgl*II-*Hind*III fragment.

Using *Eco*RI, *Eco*RI-*Bam*HI, and *Bgl*II-*Hind*III restrictions it was possible to both locate the *pca*::Tn*5-B20* insertions and to determine the orientation of the *lacZ* gene in each case. The positions and orientations of the insertions are diagrammed in Figure 3-7. Insertions *pcaC8*::Tn*5-B20* and *pcaD14*::Tn*5-B20* are oriented in the same direction; *pcaH39*::Tn*5-B20* is in the opposite orientation. This is in keeping with the results from the β-galactosidase activity; insertions *pcaC8*::Tn*5-B20* and *pcaD14*::Tn*5-B20* were not induced by protocatechuate, while *pcaH39*::Tn*5-B20* was induced by protocatechuate. This implies that the direction of transcription is the same as the orientation of *pcaH39*::Tn*5-B20*. Sequencing showed that this was the case (below).

III. Sequencing of the *pca* Region of *R. meliloti*

The following steps were carried out by Belinda Schoeman, Danielle Segato, and Xiaobang Chen.

Partial DNA sequencing of the *pca* region of *R. meliloti* was undertaken in order to identify the genes located here. Sequence was obtained from pTH185 and pTH215

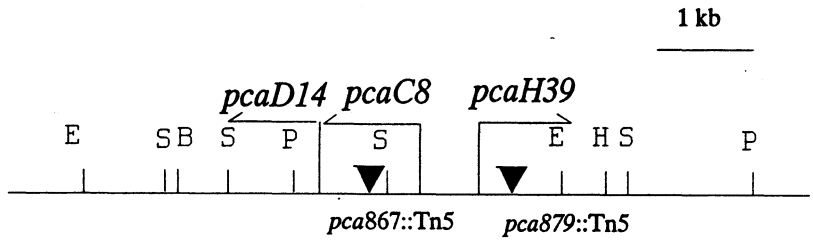


Figure 3-7. Map positions and orientations of *pca*::Tn5-B20 insertions in pTH178. The arrows indicate the direction of transcription of the *lacZ* gene of the *Tn5-B20* insertions.

using the -40 universal primer. As well, sequence was obtained from a *Sal*I deletion and an *Eco*RI deletion of pTH215, also using the -40 universal primer. The 2.6 kb *Sal*I fragment from pTH215 was cloned into pUC119 with the internal *Hind*III site adjacent to the primer annealing site. The -40 universal primer was used to sequence in towards and past the *Hind*III site. The 1.2 kb *Pst*I-*Bgl*II fragment from pTH185 was cloned into both pUC119 and pUC118, and sequence was obtained using the -40 universal primer. Also, the 1.0 kb *Eco*RI-*Bgl*II fragment from pTH185 was cloned into pUC118. A *Sal*I deletion of this fragment in pUC118 was also constructed. The -40 universal primer was used with these plasmids to obtain sequence. In the case of the 1.0 *Eco*RI-*Bgl*II fragment from pTH185 in pUC118, and the *Sal*I deletion of this clone, there was overlap between the sequences. The two sequences were combined to create a large stretch of sequence, identified as G6.

In order to obtain sequence from the site of insertion of the Tn*5-B20*'s in pTH178, the *Eco*RI fragments from the insertions in pTH178 containing the kanamycin resistance gene of the Tn*5-B20* and flanking *R. meliloti* DNA was cloned into pUC118, selecting for kanamycin resistance. This was done for insertions 8, 14, and 39 in pTH178. Sequence was obtained from these plasmids using the *IS*50 (Tn*5*) primer.

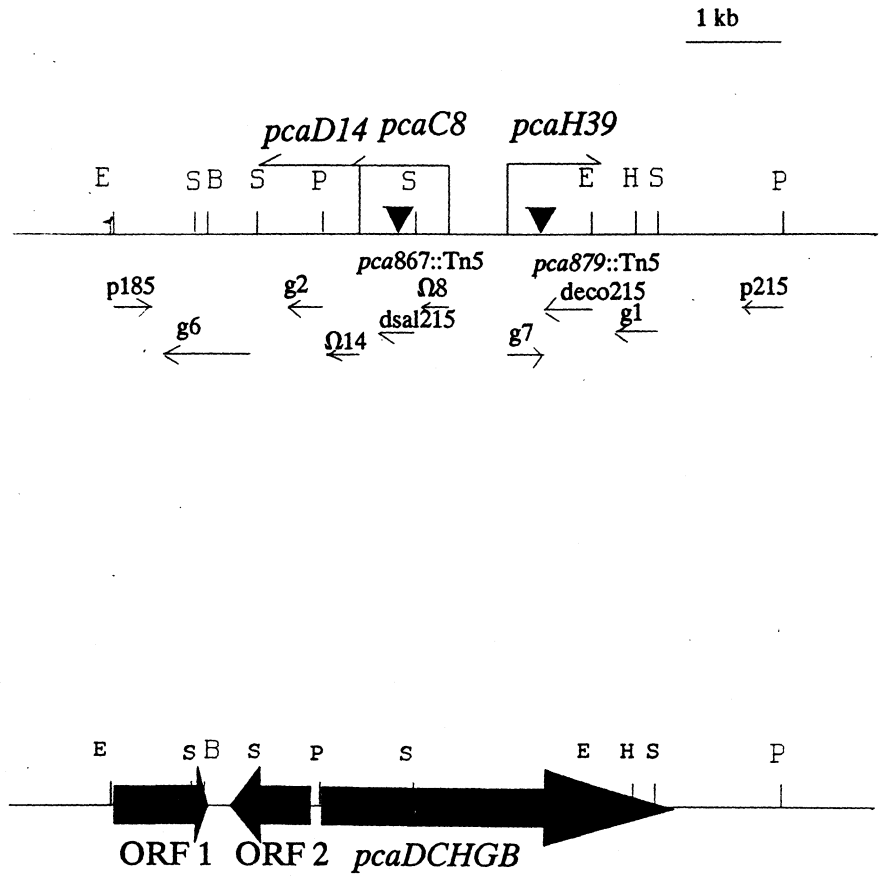


Figure 3-8. DNA sequencing and gene order of the *R. meliloti* pRmeSu47b *pca* gene region.

Sequence was also obtained from a region flanking the *pcaG879*::Tn*5* insertion. The Tn*V* replacement of *pcaG879*::Tn*5* was cloned out of RmH340 (Rm5000 *pcaG879*::Tn*V*) via a ligation of an *Eco*RI digest of RmH340 total DNA, to create pTH320. Strain RmH340 was constructed by T.M. Finan. A 5.5 kb *Hind*III-*Eco*RI fragment was cloned from pTH320 into pUC118 to form pTH329. The *IS*50 (Tn*5*) primer was then used to obtain sequence. This sequence was melded with sequence from *pcaH39*::Tn*5-B20* to yield the sequence G7. The locations of all of the sequence obtained is diagrammed in Figure 3-8. The DNA sequence obtained is listed in Appendix A.

Sequence J59 was obtained from pTHJ59. pTHJ59 was constructed by a ligation of an *Eco*RI restriction of RmH339 (Rm 1021 *katE878*::Tn*5*) total DNA. A 5.5 kb *Eco*RI/*Hind*III fragment from pTH59 containing the left *IS*50 element of the *katE878*::TnV was cloned into pUC118 to generate pTH59-25. Sequence was obtained using the *IS*50 primer to sequence out from the left *IS*50 element of the insertion. Sequence flanking the other *IS*50 of the *katE878*::Tn5 was obtained by cloning a 1.3 kb *Eco*RI/*Hind*III fragment of pTHJ59 into pUC118 to generate pTHJ59-27. Sequence was obtained from pTHJ59-27 using the *IS*50 primer. Sequence J59 was obtained by joining the two sequences at the point of transposon

insertion, and removing the duplicate sequence resulting from Tn*5* insertion.

Sequence obtained was used to search databases of sequence for homology to gene products, at both the nucleotide and protein sequence levels, through use of the blastx program made available via e-mail from the National Centre for Biotechnology Information at the National Institute of Health in the United States. The results of the blastx search, namely the top three homologies, where there were at least three homologies, are summarized for the sequences in Tables 3-16 to 3-23. Results are not given for sequence from the 1.2 kb *Pst*I-*Bgl*II fragment from pTH185 cloned in pUC119, as no significant results were obtained. Results of the blastn search, for sequence which had homology to other nucleotide sequence are summarized in Table 3-24.

Table 3-16: Results of Blastx Search with Sequence from putative *ppdA*

Sequence	**Homology to**	**Smallest Poisson Probability**
P185: Nucleotides 31 to 114 124 to 354	4-hydroxyphenylpyruvate dioxygenase, *Pseudomonas* ssp.	1.0 $x10^{-15}$ In frame +1
P185: Nucleotides 130 to 270 313 to 366	MelA protein, *Shewanella colwelliana*	2.0 x 10^{-9} In frame +1
P185: Nucleotides 124-312	Legiolysin, *Legionella pneumophila*	5.5 x 10^{-6} In frame +1
G6: Nucleotides 15 to 443	Legiolysin, *Legionella pneumophila*	4.8 x 10^{-11} In frame -2
G6: Nucleotides 12 to 260	MelA protein, *Shewanella colwelliana*	1.4 x 10^{-9} In frame -2
G6: Nucleotides 216-287	4-hydroxyphenylpyruvate acid dioxygenase, *Streptomyces avermitilis*	1.6 x 10^{-7} In frame -2

This sequence shows homology to 4-hydroxyphenylpyruvate dioxygenases and the gene is designated *ppdA*.

Table 3-17: Results of Blastx Search with Sequence from putative *pcaQ*

Sequence	**Homology to**	**Smallest Poisson Probability**
G2:Nucleotides 7-366	*gltC* (regulator) gene product from *B. subtilis*	5.9×10^{-20} frame +1
G2:Nucleotides 34-393	hypothetical 31.4 kd protein in CotF-TetB intergenic region in *B. subtilis*	1.6×10^{-18} frame +1
G2:Nucleotides 7-366	Regulatory protein from plasmid pJP4	2.1×10^{-15} frame +1

This sequence shows homology to LysR-type regulators at the amino acid level, and the gene is designated *pcaQ*.

Table 3-18: Results of Blastx Search with Sequence from putative *pcaD*

Sequence	**Homology to**	**Smallest Poisson Probability**
Ω14: Nucleotides 1 to 114	β-ketoadipate-enol -lactone hydrolase (PcaD), *A. calcoaceticus*	3.0×10^{-7} Frame -3

This sequence shows homology to PcaD at the amino acid level, and the gene is designated *pcaD*.

Table 3-19: Results of Blastx Search with Sequence from putative *pcaC*

Sequence	**Homology to**	**Smallest Poisson probability**
D*Sal*215: Nucleotides 1 to 57	γ-carboxymuconolactone decarboxylase (*pcaC*), *A. calcoaceticus*	3.0×10^{-8} In frame -3
Ω8: Nucleotides 49 to 270	γ-carboxymuconolactone decarboxylase (*pcaC*), *A. calcoaceticus*	6.8×10^{-19} In frame -1
Ω8 Nucleotides 58 to 195	γ-carboxymuconolactone decarboxylase homolog, *Paracoccus denitrificans*	1.7×10^{-9} In frame -1

This sequence shows homology to PcaC at the amino acid level, and this gene is designated *pcaC*.

Table 3-20: Results of Blastx Search with Sequence from putative *pcaHG*

Sequence	**Homology to**	**Smallest Poisson Probability**
G7: Nucleotides 155-226 233 to 382	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *A. calcoaceticus*	1.2×10^{-23} In frame +2
G7: Nucleotides 146 to 223 281 to 385	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *P. putida*	5.3×10^{-12} In frame +2
G7: Nucleotides 146 to 223 281 to 385	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *P. aeruginosa*	5.3×10^{-12} In frame +2
G7: Nucleotides 4 to 105	Protocatechuate 3,4-dioxygenase, β subunit (PcaH), *P. putida*	2.2×10^{-7} In frame +1
G7: Nucleotides 4 to 105	Protocatechuate 3,4-dioxygenase, β subunit (PcaH), *P. aeruginosa*	1.2×10^{-6} In frame +1
G7: Nucleotides 4 to 105	Protocatechuate 3,4-dioxygenase, β subunit (PcaH), *A. calcoaceticus*	2.5×10^{-6} In frame +1
DEco215: Nucleotides 158 to 391	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *A. calcoaceticus*	3.2×10^{-17} In frame -2
DEco215: Nucleotides 161 to 382	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *P. aeruginosa*	7.0×10^{-15} In frame -2
DEco215: Nucleotides 161 to 382	Protocatechuate 3,4-dioxygenase, α subunit (PcaG), *P. putida*	7.1×10^{-15} In frame -2

This sequence shows homology to PcaG and PcaH at the amino acid level, in different parts of the sequence, and the two corresponding genes are designated *pcaG* and *pcaH*.

Table 3-21: Results of Blastx Search with Sequence from putative *pcaB*

Sequence	**Homology to**	**Smallest Poisson Probability**
G1: Nucleotides 2 to 301 304 to 390	β-carboxy-cis,cis -muconate lactonizing enzyme (PcaB), *P. putida*	3.8×10^{-10} In frame -2
G1: Nucleotides 2 to 223 304 to 390	β-carboxy-cis,cis -muconate lactonizing enzyme (PcaB), *A. calcoaceticus*	4.6×10^{-10} In frame -2
G1: Nucleotides 236 to 337	aspartate-semialdehyde dehydrogenase, *Streptomyces akiyoshiensis*	1.8×10^{-4} In frame -2

This sequence shows homology to PcaB at the amino acid level, and the gene is designated *pcaB*.

Table 3-22: Results of Blastx Search with Sequence from putative ORF 7

Sequence	**Homology to**	**Smallest Poisson Probability**
P215: Nucleotides 21 to 143 150 to 440	ORF gene product, H -repeat gene, *E. coli*	2.4 x 10^{-57} In frame +3
P215: Nucleotides 117 to 230 272 to 313	Hypothetical protein 2, *Calothrix* sp.	2.9 x 10^{-14} In frame +3
P215: Nucleotides 117 to 230 272 to 313	Gene product for Group II intron-contained ORF, *Calothrix* sp.	2.9 x 10^{-14} In frame +3

This sequence shows homology at the amino acid level to Group II Intron - contained Open Reading Frames.

Table 3-23: Results of Blastx Search with Sequence from *katE878*::Tn5

Sequence	**Homology to**	**Smallest Poisson Probability**
J59:Nucleotides 245-711	Catalase from *Brucella abortus*	1.8 x 10^{-103} frame +2
J59:Nucleotides 245-352 365-670	Catalase from *Bordetella pertussis*	7.8 x 10^{-82} frame +2
J59:Nucleotides 245-670	Catalase from *Campylobacter jejuni*	2.6 x 10^{-80} frame +2

This sequence shows homology to catalase at the amino acid level, and the gene is designated *katE*.

Table 3-24: Results of Blastn Search with Sequence from the *pca* region of *R. meliloti*

Sequence	**Homology to**	**Extent of Homology**	**Smallest Poisson Probability**
G1	*pcaB*, *P. putida*	71 % identity over 59 nucleotides	.0018
d*Sal*215	*pcaGH*, *P. cepacia*	63 % identity over 106 nucleotides	3.9×10^{-5}
G7	*pcaGH*, *P. putida*	71 % identity over 107 nucleotides	3.2×10^{-11}
	pcaGH, *A. calcoaceticus*	62 % identity over 106 nucleotides	8.0×10^{-9}

Figure 3-9: Alignment of partial Sequence of *R. meliloti* PcaD with that of *A. calcoaceticus*

```
A. cal          1 VSEIHSIMITNRQGKTLSVQINGPE-NAPAIVFSNS 36
R. mel            WRRRAAVQFTRINDVAIHYRVVGAVTEKPALVFINS
                       ..  *   .  ..  .. *     **.** **

A. cal         37 LGTDHGMWQPQVAALKSQYRVVTYDTRGHG 69
R. mel            LGTDFRIWRDVVLRLAGDFAIVLYDKRGHG
                  ****  .*.  *  *      * ** ****
```

Figure 3-10: Alignment of partial Sequence of *R. meliloti* PcaC sequence with that of *A. calcoaceticus*

```
R. mel            KRERSIVTIALLAALGQDEEVAMHVRATANTGATRE
A. cal         57 RHTRSLVTIAVLLALGREDELRIDLRACFNNGVTKD 92
                     **.****.* ***...*. . .**. *.* *..

R. mel            DICEALLHVAIYAGVPAANHAIKIVKQVFAEMDAGTVAR---
A. cal         93 ELKELILHCSLYAGLPASMHMNAAAEEVFKDLGIAPEKVNKD
                  .. * .** ..***.**. *      .** ..
```

Figure 3-11: Alignment of partial sequence of *R. meliloti* PcaH with that of *A. calcoaceticus* and *P. putida*

```
A. cal        204 ALIALEDKSNFIEADSRCYRFDITLRGRRATYFENDLT 241
R. mel            HLIAALDWANTVPMDARAYRFDIVLRGRRSTFFENRPEG
P. put        204 QLIAKLDMNNANPMDCLAYRFDIVLRGQRKTHFENC-  240
                   ***  *  *    *. .*****.***.* * ***
```

Figure 3-12: Alignment of partial sequence of *R. meliloti* pcaG with that of *A. calcoaceticus* and *P. putida*

```
A. cal    127 KGSTQAPHISLIIFARGINIGLHTRVYFDDEAEANAKDPVLN 168
R. mel        DGRLMPPHISIWIVARGINIGLHTRMYFPDEAAANAEDPLLA
P. put    127 AGVPMAPHINISLFARGINIHLHTRLYFDDEAQANAKCPVLN 168
               *    *** . . ****** ****.** *** ***  *.*

A. cal    169 SIEWATRRQTLVAKREERDGEVVYRFDIRIQGENETVFFDI 209
R. mel        RIEHRHRAETLVAAGQAPN----YVFDIHLQGEKETVFLDI
P. put    169 LIEQPQRRETLIAKRCEVDGKTAYRFDIRIQGEGETVFFDF 209
               **   *..** *..    *   *.***..*** ****.*.
```

Figure 3-13: Alignment of partial Sequence of *R. meliloti* PcaB with that of *A. calcoaceticus* and *P. putida*

```
P. put      1 LGKVIASGVPEAERYVHLGATSQDAMDTGLVLQLRDALDLIEADL 45
A. cal      1 LTAIVKDADEDAARYVHWGATSQDILDTACILQCRDALAIVQNQV 45
R. mel        RARAPIAQVGRRRVGTPFRSHQPGRGRYQPHAEAEGDSRNFSSGR
              .             .

P.put      46 GKLADTLSQQALKHADTPMVGRTWLQHATPVTLGMKLAGVLGALT 90
A. cal     46 QQCYETALSQAQTYRHQVMMGRTWLQQALPITLGHKLARWASAFK 90
R. mel        VGDVVTALEACGRHWGERALTAHTRMQPAIAITVYDRLRWIEPLL
                   *    .          .    ... . ... .....  ..

P. put     91 RHRQRLQELGPPCWCCSSGGASGSLAALGSKAMPVAEALAEQLK 161
A. cal     91 RDLDRINAIKARVLVAQLGGAVGSLASLQDQGSIVVEAYAKQLK 161
R. mel        DHQDRLDTTGTDIFAVQFGGAAGTLEKLKGKADIVRATLAELAL
              .. .*.   .      . *** *.*. *  ..  *  . *....
```

Figure 3-14: Alignment of Partial Sequence of PcaQ with the Consensus Sequence Amino Terminal Conserved Region. h indicates a hydrophobic residue; p indicates a hydrophillic residue, x indicates any residue.

```
Consensus     hpLRpLRxFxxhxpppphSxAApxLphTQPAhRTxQhpp
R. mel              TFVEVARQKSVVKAAELLHVTQPAVTKTIRRL
                     *           **  *  ****

Consensus             LEpxLGxxLFxRxpRxhxxxTx
R. mel                EEVLGVAVFEREGRGIKITRYG
                       *
```

Figure 3-15: Alignment of Partial Sequence of *R. meliloti* catalase from J59 with that of *Brucella abortus*

```
R. mel      ------------LRIRWRQARVASCCR--IISSSR---SWRQ
B. abo    1 MTDRPIMTTSAGAPIPDNQNSLTAGERGPILMQDYQLIEKLS 43
                        *   *  ...  *  *.

R. mel      HQNRERIPERVVHAKGWGAFGSLKITGDISQYTRAKCLQPGAETP
B. abo   44 HQNRERIPERAVHAKGWGAYGTLTITGDISRYTKAKVLQPGAQTP 88
            ********** ********.*.* ******.**.** *****.**

R. mel      MLARFSTVAGEQGAADHERDVRGFALKFYTDEGNWDLVGNNTPVF
B. abo   89 MLARFSTVAGELGAADAERDVRGFALKFYTQEGNWDLVGNNTPVF 134
            *********** **** ************ **************

R. mel       FIRDPYKFPDFIHTQKRHPKTNLRSATAMWDYWSS-PESLHQVT
B. abo  135 FVRDPLKFPDFIHTQKRHPRTHLRSATAMWDFWSLSPESLHQVT 178
            *.*****************.***********.**  ********

R. mel       ILMSDRGLRRRR
B. abo  179 ILMSDRGLPTDV 190
            ********
```

Discussion

Gene Order in *R. meliloti*

The results in this thesis demonstrate that the order of the genes involved in the catabolism of protocatechuate located on pRmeSU47b in *R. meliloti* is *pcaDCHGB*. An apparent LysR-type regulatory gene (tentatively labelled *pcaQ*) which is transcribed divergently from this operon was also found (Figure 3-8). The primary evidence for the identity of the individual *pca* genes was obtained from sequence comparisons employing partial DNA sequences from the pRmeSu47b megaplasmid *pca* locus (Figure 3-8) linked to *dctB* (Figure 3-1).

Supporting the sequence data, two Pca$^-$ Tn*5* mutations which mapped to this region were shown to lack protocatechuate 3,4-dioxygenase activity (Table 3-11). In addition, these two Tn*5* mutants were negative for the Rothera test for *ortho* cleavage of protocatechuate (Table 3-10). This Rothera phenotype was complemented by the cosmid pTH178, as well as a PstI subclone of this cosmid, pTH194 (Tables 3-10 and 3-15). Also, plasmid pTH194 complemented the Rothera negative phenotype of the large megaplasmid deletion strain RmF726, indicating that the first four genes

of the protocatechuate branch of the β-ketoadipate pathway are located on pTH194. Finally, the well-characterized plasmid pARO2, carrying the *pcaC* and *pcaHG* genes of *Acinetobacter calcoaceticus*, partially complemented the Pca⁻ phenotype of the *pcaH879*::Tn5 mutant (Table 3-6).

It is of interest to compare the partial gene order, as so far determined, in *R. meliloti*, *pcaDCHGB*, with that in other bacterial species. The gene order in *A. calcoaceticus*, *Pseudomonas putida*, *A. tumefaciens* and *R. meliloti* is given in Figure 4-1. In *A. calcoaceticus*, all of the *pca* genes are clustered in a single operon, with the order *pcaIJFBDKCHG* (Kowalchuk et al., 1994). The location of the regulator of the *A. calcoaceticus pca* operon is not known. The difference in gene order between *R. meliloti* and *A. calcoaceticus* does not appear to be the result of a single rearrangement of the genes, because the difference in order is too great. The *pcaK* gene is either lacking in *R. meliloti*, or is located elsewhere, and the *pcaB* gene is located at the end of the transcriptional unit in *R. meliloti*.

The arrangement of *pca* genes in *R. meliloti* also differs from that in *P. putida*, in which *pca* genes are located in four operons. The four operons are *pcaHG* (Frazee et al., 1993), *pcaBDC* (Hughes et al., 1988), *pcaIJ* (Parales and Harwood, 1992) and *pcaRKF* (Harwood et al., 1994).

Comparing the order of genes between *A. calcoaceticus* and *P. putida*, it can be seen that the order is more similar.

In *A. tumefaciens*, the gene order, as far as it has been determined, is identical to that in *R. meliloti* as far as it has been determined (the *pcaK* and *pcaF* genes have not been isolated in either *A. tumefaciens* or *R. meliloti*, and the *pcaIJ* genes have not been isolated in *R. meliloti*). This is not too surprising, as *A. tumefaciens* and *R. meliloti* are both members of the family *Rhizobiaceae*, and *A. tumefaciens* is more closely related to *R. meliloti* than either is to *A. calcoaceticus* and *P. putida* (Parke, 1995).

It should be noted that the order of transcription of the genes encoding the protocatechuate 3,4-dioxygenase, *pcaHG*, is the same in all bacteria for which sequence data is available. This includes *Bradyrhizobium japonicum* USDA110 (Podila et al., 1993) and *Pseudomonas cepacia* (Zylstra et al., 1989).

Although only partial sequences of the *pca* genes of *R. meliloti* were determined in this work, it is of interest to compare the degree of homology of the *R. meliloti* genes with those of *A. calcoaceticus* and *P. putida*. Using the

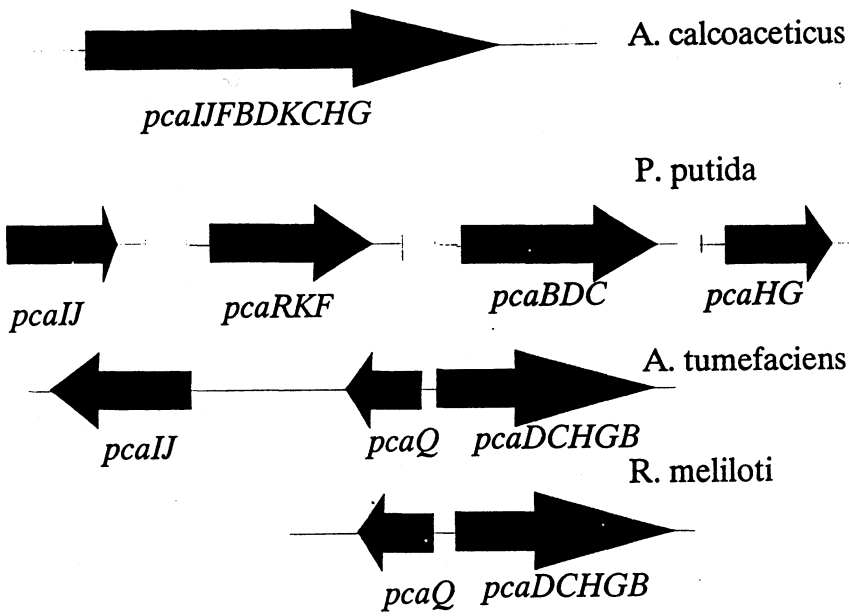


Figure 4-1. *pca* Gene order in *A. calcoaceticus* (Kowalchuk et al., 1994), *P. putida* (Frazee et al., 1993; Harwood et al., 1994; Hughes et al., 1988; Parales and Harwood, 1992), A. tumefaciens (Parke, 1995) and *R. meliloti* (this work).

blastn program, (Altschul et al., 1990) homology was detected at the nucleotide level between the pcaB genes of *R. meliloti* and *P. putida* (Table 3-24). No significant homology at the nucleotide level was detected to the *pcaB* gene of *A. calcoaceticus*. Homology at the nucleotide level was detected between the *pcaHG* genes of *R. meliloti* and the corresponding genes of *P. putida*, *P. cepacia*, and *A. calcoaceticus* (Table 3-24). *R. meliloti* was most similar to *P. putida*. These results suggest that the *pca* genes of *R. meliloti* are more closely related to those of *P. putida* than to those of *A. calcoaceticus*. The evolutionary distance between *R. meliloti* and *A. calcoaeticus* is expected to be about the same as the distance between *R. meliloti* and *P. putida*, as both *A. calcoaceticus* and *P. putida* are both members of Subgroup γ-3 of the γ Group of purple bacteria, while *R. meliloti* is a member of the α Group of purple bacteria (Woese, 1987).

However, comparison of identities at the amino acid level show that the levels of identity of the *R. meliloti* gene products with each of *P. putida* and *A. calcoaceticus* are nearly the same. In any case, evolutionary relationships between *pca* alleles of the different species can not be extrapolated with any accuracy based on the partial sequence data from *R. meliloti*.

Megaplasmid Location of *pca* Genes

In previous work it was shown that *R. meliloti* strains carrying either of two overlapping deletions of the pRmeSU47b megaplasmid failed to grow with C_4-dicarboxylates, lactose, acetoacetate, β-hydroxybutyrate or protocatechuate as sole carbon source (Charles and Finan, 1991). The cosmid clone pTH178 complemented the Pca$^-$ phenotype of the *pca*::Tn*5* mutants RmG867 and RmG879, however it did not complement the Pca$^-$ phenotype of the megaplasmid deletion strain RmF726 (this strain has a 300 kb deletion of the pRmeSU47b megaplasmid, with the *pca-dct* region deleted). This implies that there is another *pca* locus deleted in strain RmF726 that is not carried on pTH178. The *pca* genes not accounted for are *pcaK*, *pcaIJ*, and *pcaF* (see Figure 1-1). One or more of these may be encoded for by the gene knocked out in *pca-877*::Tn*5*. The possibility that protocatechuate inhibits growth in an RmF726 (pTH178) and that there is only one megaplasmid *pca* locus should be examined.

In most of the bacteria studied to date, genes of the *ortho* cleavage pathway are located on the chromosome, with genes for the *meta* cleavage pathways typically being found on plasmids (Assinder and Williams, 1990). An exception to this is a strain of *A. calcoaceticus* (Winstanley et al.,

1987) for which *cat* (*ortho*) genes were located on a plasmid. However, genes for catechol utilization have been found on a plasmid in a strain of *R. leguminosarum* bv. trifolii (Baldani et al., 1992). In the case of *R. meliloti*, it is perhaps not surprising that *pca* genes are located on the pExo megaplasmid, given the variety of other catabolic genes found here (Charles and Finan, 1991). The megaplasmid might also be considered to be a chromosome, given its large size.

It has been shown recently that members of the family *Rhizobiaceae* break down polyaromatic flavonoids, synthesized by plants, that act as inducers of the *nod* genes. The flavonoid breakdown products were shown to include protocatechuate, phloroglucinol, *p*-coumaric acid, *p*-hydroxybenzoic acid, phenylacetic acid, and caffeic acid (Rao et al., 1991; Rao and Cooper, 1994). It is possible that genes catalyzing the breakdown of flavonoids may also be located on the megaplasmid, linked to *pca* genes. The ability to catabolize flavonoids may prove beneficial to *Rhizobium* in the rhizosphere, and flavonoids may serve as a source of reduced carbon available to *Rhizobium* in the soil.

Regulation of *pca* Genes in *R. meliloti*

Of the three independent Tn*5-B20* *pca* insertions in cosmid pTH178, only *pcaH39*::Tn*5-B20* showed induction of β-

galactosidase in the presence of protocatechuate (Table 3-12). The orientation of the *lacZ* gene in the *pcaH39* insertion was the same as the gene order implied by the sequence data, while the other two insertions were oriented in the opposite direction.

The amino acid sequence of the putative *pca* regulator (*pcaQ*) in *R. meliloti* showed significant homology to members of the LysR-type regulators (Table 3-17; Figure 3-14). LysR-type regulators are similarly-sized autoregulatory transcriptional regulators, which activate the transcription of divergently-linked genes. These regulators typically have a greater than 20 percent amino acid identity with other LysR family members, or homology to a consensus sequence for the highly conserved amino acid terminus (Schell, 1993). The putative *pcaQ* of *R. meliloti* shows homology to the LysR N-terminal consensus sequence (Figure 3-14).

It is suggested that the *pca* regulatory gene identified here is *pcaQ*, the same as that in *A. tumefaciens*. Unfortunately the *pcaQ* sequence of *R. meliloti* cannot be compared to that of *A. tumefaciens*, as the latter has not yet been sequenced. The structure of the *pca* operon here is the same as the *pcaDCHGB* operon in *A. tumefaciens*, down to the divergently transcribed regulatory genes. The *pcaQ* gene of *A. tumefaciens* has been shown to induce the expression of

pcaD in the presence of β-carboxy-*cis,cis*-muconate (Parke, 1993). Also, in *R. leguminosarum*, β-carboxy-*cis,cis*-muconate was shown to induce expression of the *pca* genes in the upper portion of the pathway (Figure 1-1). Induction of *R. meliloti pca* genes by β-carboxy-*cis,cis*-muconate has not been tested.

It should be noted that some of the other characterized *pca* and *cat* regulatory genes are members of the *lysR*-type family of regulators. The *catR* gene of *P. putida*, which in conjunction with *cis,cis*-muconate controls expression of the *catBC* operon (Parsek et al., 1992), is a *lysR*-type regulator. The *catM* gene of *A. calcoaceticus*, which, along with the inducer *cis,cis*-muconate, controls the expression of *cat* genes, is also a *lysR*-type regulator. It may be that the *pca* regulatory genes in *R. meliloti*, *R. leguminosarum*, and *A. tumefaciens* are *lysR*-type regulators that respond to the inducer β-carboxy-*cis,cis*-muconate, and share ancestry with *catM* of *A. calcoaceticus* and *catR* of *P. putida*.

Interestingly, the PcaR gene product of *P. putida*, which induces the expression of all of the *pca* genes (except for *pcaHG*) in the presence of β-ketoadipate, is not a member of the family of *lysR*-type regulators. Instead it shows significant homology to PobR from *A. calcoaceticus* (Romero-Steiner et al., 1994), which induces the expression the gene

encoding 4-hydroxybenzoate hydroxylase, *pobA*, in the presence of 4-hydroxybenzoate. It might be expected that all of the regulatory genes of the *pca* branch of the β-ketoadipate pathway would share common ancestry, and that all of the regulatory *cat* genes would share common ancestry, with the two groups not necessarily being related. However, this is not the case.

Evidence for Superoperonic Clustering of Genes Involved in the Catabolism of Aromatics by *Rhizobium meliloti*

Lying adjacent to the *pca* gene region in *R. meliloti* is a gene showing homology to i) *melA* of *Shewanella colwelliana*, and to ii) the gene for 4-hydroxyphenylpyruvate dioxygenase of *Pseudomonas* sp. strain P.J. 874 (Ruetschi et al., 1992) and *Streptomyces avermitilis* (Denoya et al., 1994), and iii) the legiolysin gene of *Legionella pneumophila* (Table 3-16). 4-hydroxyphenylpyruvate dioxygenase catalyzes the formation of homogentisate from 4-hydroxyphenylpyruvate. The *melA* gene product of *S. colwelliana* is also apparently a 4-hydroxyphenylpyruvate dioxygenase, as the product of the reaction catalyzed by *melA* is homogentisic acid (Coon et al., 1994). The *melA* gene product has significant homology to 4-hydroxyphenylpyruvate dioxygenases. This suggests that this

gene encodes a 4-hydroxyphenylpyruvate dioxygenase, or a closely related gene product. There is a precedent for genes involved in aromatic catabolism to be organized in superoperonic clusters. The *pca*, *pobA*, and *qui* (genes involved in the catabolism of quinate) genes have been shown to be in different, linked operons in *A. calcoaceticus* (Averhoff et al., 1992; Elsemore and Ornston, 1994). Also the *ben* (genes involved in the conversion of benzoate to catechol) and *cat* genes in *A. calcoaceticus* are clustered together in separate operons (Shanley et al., 1994), and in fluorescent *Pseudomonas* species (Rosenberg et al., 1969; Wheelis and Stanier, 1970). In *A. tumefaciens*, two *pca* operons are tightly linked (Parke, 1995).

Reiterated DNA Sequence Present on Cosmid pTH178

A Southern blotting experiment was done in order to map the positions of the Tn*5* insertions *pca-867*::Tn*5* and *pcaH879*::Tn*5*. When CsCl purified pTH178 was used as a probe to Southern blots of DNA from wild type and *pca*::Tn*5* mutants, a large number of bands of different sizes were visualized on the filter. Most of these bands did not correspond to fragments in pTH178 and Tn*5*. This suggests that the DNA cloned in pTH178 contains DNA elements which

are repeated on the *R. meliloti* chromosome. Repetitive DNA sequence elements have been reported in prokaryotic genomes. An element termed a repetitive extragenic palindrome (REP) has been identified in *Escherichia coli* and *Salmonella typhimurium* (Gilson et al., 1984; Lupski and Weinstock, 1992; Stern et al., 1984). Another element which has been identified is the enterobacterial repetitive intergenic consensus (ERIC), found in *E. coli* and *S. typhimurium* (Hulton et al., 1991; Sharples and Lloyd, 1990). Recently, repeated elements, termed rhizobium-specific intergenic mosaic elements (RIMEs) were identified in *Rhizobium* species and in *Agrobacterium rhizogenes* (Osteras et al., 1995). These elements were first identified in *R. meliloti* and *Rhizobium* sp. strain NGR234 in the region upstream of the phosphoenolpyruvate carboxykinase gene (*pckA*). It may be that the repetitive elements cloned in pTH178 are RIMEs. Blastn searches for nucleic acid homology using sequence from the *pca* region of *R. meliloti* turned up no homology to any repetitive element described.

Failure of pTH194 to Complement *pca-867* and *pcaG879* Tn5 mutations

Although plasmids pTH193 and pTH194, along with pTH178, carry the *pcaDCHGB* genes of *R. meliloti*, only pTH178 complemented the growth phenotype of the two Tn*5* mutants.

The Rothera results for RmF726 carrying these plasmids indicate that the *pca* genes are expressed on pTH178 (Table 3-10), to a much higher extent on pTH194, and at a very low level or not at all on pTH193 (see Table 3-15). The reason for the difference in expression between pTH194 and pTH193 is likely due to the orientation of the *pca* genes with respect to p*lac* of the vector pRK7813. In pTH194, the direction of transcription from p*lac* is the same as the direction of transcription of the *pca* genes, while in pTH193 the direction of transcription from p*lac* is opposite the orientation of the *pca* genes. Thus it is likely that the p*lac* in pTH194 is causing the *pca* genes to be overexpressed. The Rothera phenotypes of the Tn*5* mutants and the large deletion mutant RmF726 are complemented by pTH194, indicating that the *pcaDCHGB* genes are expressed in these backgrounds. It would be expected that pTH194, in addition to pTH178, should complement the growth phenotypes of the *pca*::Tn5 mutants. The Rothera results indicate that expression of *pca* genes on pTH194 is much higher than on pTH178, likely being driven by the p*lac* of the pRK7813 vector. It may be that higher expression of the *pca* genes on pTH194 affected complementation. One of the intermediates of protocatechuate degradation may accumulate when a subset of the *pca* genes are overexpressed, and have an inhibitory effect on growth (see below).

A Tn5 Catalase Mutant is Unable to Grow on Protocatechuate

Strain RmG878 (*katE*::Tn5) has a Tn*5* insertion in a gene with strong homology (80% at the amino acid level) to catalase (Table 3-23), and is unable to grow on protocatechuate. Catalase catalyzes the formation of water and oxygen from hydrogen peroxide. Presumably the failure of RmG878 to grow on Pca does not result from an inability to catabolize this compound. Rather we suspect that Pca or a Pca degredation product is toxic to a catalase mutant. In fact, cells of RmG878 incubated in the presence of protocatechuate are Rothera positive, indicating that the first four genes of the protocatechuate branch of the β-ketoadipate pathway (*pcaDCHGB*) are expressed (Table 10).

Aromatic compounds often have a inhibitory effect on microbial growth. Benzoic acid, sodium benzoate, methylparaben and ethylparaben are aromatic compounds used as preservatives in food (Frazier and Westhoff, 1988). The mechanisms of action of these preservatives are not known, although it is known that phenolic substances affect the integrity and function of cell membranes (Heipieper et al., 1991).

It has been noted that *E. coli* cells growing on 2-phenylethylamine have increased levels of catalase activity (Parrot et al., 1987). 2-phenylethylamine is metabolized to

phenylacetaldehyde, which is then converted to phenylacetic acid. The increased catalase activity is believed to be due to the increased level of hydrogen peroxide expected to be a product of the amine oxidation catalyzed by the amine oxidase. Whether or not a catalase negative mutant was able to grow on 2-phenylethylamine was not tested. Other enzymatic sources of active oxygen species include autoxidation of various dehydrogenases (Imlay and Fridovich, 1991). Formation of hydrogen peroxide during protocatechuate metabolism does not take place, as ring cleavage occurs via a dioxygenase, and not a dehydrogenase. It should be noted, however, that intermediates in the β-ketoadipate pathway are unstable, and may react with other compounds in the cell. Nonenzymatic sources of reactive oxygen species include the autoxidation of aromatic compounds, such as ubiquinols, catechols, and flavins (Farr and Kogoma, 1991), and it may be that it is protocatechuate that is producing the reactive oxygen species via autoxidation.

Organic-solvent resistance genes encoding resistance to 1,2,3,4-tetrahydronapthalene (tetralin), propylbenzene, and 1,2-dihydronapthalene, were recently isolated from *E. coli* (Ferrante et al., 1995). The genes were identified as coding for alkylhydroperoxide reductase (*ahpCF*). The mutant allele was found to have an activity three times that of the

wildtype in reducing tetralin hydroperoxide to 1,2,3,4-tetrahydro-1-napthol. The authors suggest that the toxicity of solvents like tetralin is due to the formation of toxic hydroperoxides within the cell. Interestingly, the *ahp* mutant was still sensitive to hydrophillic solvents like xylene and toluene, implying that more than one mechanism for solvent toxicity exists.

Deletion mutants that render the metabolism of protocatechuate toxic have been reported in *A. calcoaceticus* (Hartnett et al., 1990), and have been used to select strains in which secondary mutations blocked the degradation of protocatechuate (DiMarco et al., 1993a and 1993b). Deletions of *pcaD* and *pcaK* (shown to transport 4-hydroxybenzoate in *P. putida*) fail to grow in the presence of protocatechuate (Hartnett et al., 1990). This could be due to the accumulation of β-ketoadipate-enol-lactone (product of the reaction catalyzed by the *pcaC* gene product), or it could be that the *pcaK* gene product is required for growth in the presence of protocatechuate. This is apparently not dependent on catalase, although whether it could be more marked in the absence of catalase is unknown. We have not determined if the sensitivity of the *R. meliloti* catalase mutant is due to the presence of protocatechuate, or is due to the metabolism of protocatechuate. Construction of a *pcaG katE* double mutant

would resolve this issue.

Region with Homology to Group II Intron-Contained Open Reading Frames

Introns are sequences of RNA which are spliced out from a primary RNA transcript, resulting in the joining togeather of exons to form an RNA transcript coding for a protein. Group I and Group II introns differ in the splicing mechanism used (reviewed in Saldanha et al., 1993).

Introns were once considered to be a characteristic of eukaryotic genomes, not to be found in prokaryotes. However, both group I and group II introns have since been discovered in prokaryotes (reviewed in Belfort et al., 1995). A Group I intron was first discovered in the thymidylate synthase gene of bacteriophage T4 (Chu et al., 1984). Group II introns have recently been reported in cyanobacteria (*Calothrix* spp.) and the proteobacteria *Azotobacter vinelandii* (Ferat and Michel, 1993), and *E. coli* (Ferat et al., 1994). Sequence from ORF7 in this work shares homology with open reading frames found in the *E. coli* and *Calothrix* spp. group II introns (Table 3-22), and all of these sequences show homology to reverse transcriptases (data not shown). While mobility of group II introns has not yet been observed, the presence of ORFs homologous to reverse transcriptases suggest that they may

be mobile (Belfort et al., 1995).

Conclusion

This work describes the characterization of a megaplasmid *pca* locus in *R. meliloti*. The first four *pca* genes were found to be organized in a single operon, with the order *pcaDCHGB*. This is identical to the gene order found in the closely related *A. tumefaciens*. A gene sharing homology to LysR-type regulators was found to be divergently transcribed from the *pca* operon. This is also the case for the regulator gene *pcaQ* in *A. tumefaciens*. Unexpectedly, a catalase mutant was isolated on the basis of its inability to grow with protocatechuate as sole carbon source. All of the mutants isolated had a wild type symbiotic plant phenotype.

Appendix A: DNA Sequence

Figure 5-1. DNA Sequence P185

GAATTCACCGCCGATCGGGCGGAGGCGGAGACGCTCGGCTCGCTTCTGGGAACCATGGGT
TTTCACGCCGTCGCCTCTCATCGCGAGAAATCGGTGACGCTATGGCGCCAGGGTCGCGAT
CATGGAGGCGTAAACATTGTCATCAACACCGAGCAGCAGGGCTTCGCTCATTCGAGCTAC
CTGGTGCACGGGGCATCGGCCTATGCGGTCGGCCTGAAAGTACCCGATGCCTCCGCCGCC
ATCGAACGCTCTCGCGCGCTCGGCGCCGAAATATTTCTGCCCGAGGCGAACGAAGGCGAG
GGGGCGATGGCCGCAATCCGCGGCATTGGCGGCGGGTTGATCTATTTCCTGGACGAGGCC
GACGACGTCTGGTCGCGCGAGTTCGTCTGGTCGGG

This sequence was obtained from pTH185 using the -40 universal primer (Figure 2-8).

Figure 5-2. DNA Sequence G6

TGGAAATACTCGCCATCCTCATCGCGGTCATAGAGGATGCTCTCATCCCTGAGCCGGTCG
ACGAAGTCAGGTTCGAGCCCGAACCGCACCTCGATGTCGTCGTAATAGTTGCGCGAGATT
GCCAATGGGCGAAAGCCGCTGCGGAGCATATGGTCGGCGGCTGCAAAGATATCGTCGGTG
CGGAAGGCCACATGCTGGACACCCGAGCCAAAGCTCTCTGCGATGAAGTGGCCTGCCAGC
GTGCGGTGGTTGTCCGCTCCGTTCAACGTCAGCCGAAGCGTCCCGCTCGCATTCTCAACG
ACCTGGCTGCGGATCAGACCGGCGGGATCGACGATATCGACCTCTGCCAATTTGTCGGCA
TCGAGGATAGACGTNTAGAAGAGCAGTCAGCTCGGGAGCTCCTCCGCCTGCATGGATTGC
GCCACATGGTCGATGGCGAAGAGCGGCGATGCGCTTGNGGNGTANGCCCCCGACCAGACG
AATCGCGCGACA

This sequence was obtained from the 1.0 kb *Eco*RI-*Bgl*II fragment of pTH185 in pUC118, combined with sequence from a *Sal*I deletion of the 1.0 kb *Eco*RI-*Bgl*II fragment of pTH185 in pUC118 (Figure 2-8) using the -40 universal primer.

Figure 5-3: DNA Sequence G2

GACATTTGTCGAGGTCGCACGCCAGAAGAGCGTGGTGAAGGCGGCCGAACTCCTGCATGT
CACGCAACCGGCGGTCACCAAGACGATCCGGGAACTCGAAGAGGTTCTGGGCGTCGCGGT
TTTCGAGCGCGAGGGGCGTGGCATCAAGATCACGCGCTACGGTGAGGTCTTTCTGCGCCA
TGCTGGCGCCGCGCTGACGGCGCTGNGCCAAGGCCTCGATTCCGTCTCGCAGGAGCGGTC
CGGCGAAGGCCCGCCGATCCGCGTCGGTGCGCTTCCGACCGTTTCGACCCGCATCATGCC
GCGCGCCATCGCGCTCTTCCTGAAAGAGAAGACCGGCGCACGCATCAAGATCGTCACCGG
TGAAAACGCAGTGCTTCTCGAACAATTGGGCATCGGCGATCTTCGACCTCGTCGGTCGG

This sequence was obtained from the 1.2 kb *Pst*I-*Bgl*II fragment from pTH 185 cloned into pUC118 using the -40 universal primer (Figure 2-8).

Figure 5-4: DNA Sequence Ω14

GCCCATGGCCACGCTTGTCATAGAGCACGATGGCAAAATCGCCGGCCAGCCGGAGGACGA
CATCGCGCCAGATACGGAAATCCGTGCCGAGCGAGTTGATGAAGACGAGCGCGGGCTTTT
CCGTGACCGCGCCAACCACGCGATAGTGAATCGCGACGTCGTTGATGCGGGTGAATTGCA
CGGCAGCNCTCCTCCGCCGCCAGATTTGGATGTNTCATCCGGTTAGGTAAAATGATATTT
CGNGGCTTCTCCCTNAACCAGGAGGTTATACGATCCGNGCATCGACGCTCGCGTNAAGTT
TCGCCATCTGCAGACATTTGTCGAGGTCGNACCGCCAGAAGAGCGTTGNTGAAGGGCGGC
C

This sequence was obtained by cloning the *Eco*RI fragment from the Tn*5-B20* insertion *pcaD14* containing the Km^R gene of the Tn*5-B20* and the flanking *R. meliloti* DNA, from pTH178, into pUC118, using the *IS*50 (Tn*5*) primer (Figure 2-8).

Figure 5-5: DNA Sequence D*Sal*215

GTCGACGTGATTGTCGCCTAGAACGGCGCGGCGCGTCGCCATTCCCTGTCGGTAGCGCCC
GGAAGGCGCGAGGGCATCGCTCATGGTTTGCCTCCAGACAACACGGATCGGAAGAAGCGT
CTCAGCACCGACCACAAGGCTTCCGGCTGNTCGACGCAGGGTATATGTCCCGCCCCGCGG
ATCACTTCGAAGCGCGCGCCCGGTATGAGATCGGCCGTCGAGCGGACCAGATCGGGCGGN
GTCGAACCGTCCTGGTCGCCGACGACGCAGAGCACCGGAACCCGCGATCCTGCCTGNGGN
TCCCGTGAAATCG

This sequence was obtained from a *Sal*I deletion of pTH215, using the -40 universal primer (Figure 2-8).

Figure 5-6: DNA Sequence Ω8

GTCGGGCATCGTCCTTCTCCCTCGCCTATCGAGCAACCGTGCCGGCATCCATTTCCGCGA
ACACCTGCTTGACGATCTTGATCGCGTGATTGGCCGCCGGAACGCCTGCATAGATGGCGA
CATGCAGCAGTGCCTCGCAAATATCCTCGCGCGTCGCCCCGGTATTGGCCGTGGCGCGGA
CATGCATGGCCACCTCCTCGTCCTGTCCGAGGGCGGCAAGCAGGGCGATGGTGACGATCG
ACCGCTCCCGCTTCGTCCAGCCGGCACGCGANCAGACATGCCCCCATGCCGCCTCGGTGN
ATCANGATCCTGGAAAGGACGGNCGNAAATCG

This sequence was obtained by cloning the *Eco*R1 fragment, from the insertion *pca8*::Tn*5-B20* in pTH178, containing the Km^R resistance gene of Tn*5-B20* and the flanking *R. meliloti* DNA into pUC118, and using the *IS*50 (Tn*5*) primer (Figure 2-8).

Figure 5-7: DNA Sequence G7

CATCTGATCGCGGCTCTCGACTGGGCCAACACCGTGCCGATGGATGCGCGCGCCTACAGG
TTCGACATCGTGCTGCGTGGCCGCCGCTCGACCTTCTTCGAAAACCGGCCGGAAGGGAAT
TGAGGACGCACCGATGGTTCAGGATCTCAGCACCCTCAAGGAAACCGCGTCGNAGACGGC
CGGACCCTATGTC
CATATCGGCTTGANNCCGAGTTTCTGCGGCATCGGCGGCGTATATGAGGGCGACCTCGGC
GCTTCGATGGTCAACGACAAGACGCTCGGTCAGCGCATCACCGTGACCGGTCGCGTCATC
GACGGCGCCGGAATGCCCCTCAGGGACGCACTTCTCGAAATCTGGCAGGCGGACGCAGCC
GGGCTCTACAAATTCCCCNTCGGAGGTGTGTGGTACGGGGGACCC

This sequence was obtained with sequence from a 5.5 kb *Eco*R1 fragment of pTH320 (*Eco*R1 ligation of total DNA of RmH340 (Rm5000 *pcaG879*::Tn*V*) cloned into pUC118, using the *IS*50 (Tn*5*) primer. This was melded with sequence from *pcaH39*::Tn*5-B20*, which was obtained by cloning the *Eco*R1 fragment from the insertion in pTH178 carrying the Km^R gene from the Tn*5-B20* and the flanking *R. meliloti* DNA into pUC118, and using the *IS*50 (Tn*5*) primer (Figure 2-8).

Figure 5-8: DNA Sequence D*Eco*215

TCGTCTCCGAGAAGGCCGGAAAGATAGGGGTGATCGAAGGCCGAATAGGTCATGCCGCAT
CTCCAGCTCTCGCTCCCTTCGGTTCGAAGGGAGCCTCACACGCATGCGGCAACGCGGGGG
CGGGCTGCCGAACGATCAAATATCCAGGAAGACCGTTTCCTTCTCGCCCTGGAGATGAAT
GTCGAAGACATAATTAGGCGCCTGACCCGCCGCGACAAGCGTCTCGGCCCGGTGGCGGTG
CTCGATGCGGGCAAGCAGCGGATCTTCGGCGTTCGCCGCCGCCTCGTCTGGGAAATACAT
NCGCGTGTGGAGGNCGATATTGATGCCGCGCGCGACGATCCAGATCGTGATATGCGGGGG
CATCAGCCGGCCGTCG

This sequence was obtained from an *Eco*RI deletion of pTH215 using the -40 universal primer (Figure 2-8).

Figure 5-9: DNA Sequence G1

CCAGCGCGAGTTCNTCCGCCAGCGTCGCACGGACGATGTCGGCCTTGCCCTTGAGCTTTT
CGAGCGTGCCGGCGGCTCCTCCGAACTGCACCGCGAAAATGTCCGTGCCCGTGGTATCCA
GCCTGTCCTGATGGTCGAGCAGCGGCTCGATCCAGYTCCGCAGGCGGTCGGTCACCGTAA
TCGCGATCGCCGGCTGCATTCGCGTATGGGCGGTCAGCGCACGCTCGCCCCAGTGCCGCC
CGCAAGCTTCGAGCGCCGTCACGACGTCGCCGACCCGGCCGGAGGAAAAGTTCCGCGAAT
CGCCTTCAGCCTCAGCATGAGGCTGGTATCGACCACGTCCTGGCTGGTGGCTCCGAAATG
GANATGTTCCGNCGACTCTTCGACGACCGA
CCTGGGCAATTGGCGCACGAGCTCGG

This sequence was obtained by cloning the 2.6 kb *Hind*III fragment of pTH215 into pUC119 with the internal *Hind*III site adjacent to the -40 universal primer annealing site. The -40 universal primer was used (Figure 2-8).

Figure 5-10: DNA Sequence P215

CTGCAGCGGTGGACCGCTGGCACGGGATGAATGCGGAAAGCATGGGCCGACTCCCGGCCG
GCGAGAAACCGGACGCTCCGGTCGAGCAATTCCTCTATCAGAGCCTGCTCGGCGTCTGGC
CGATCGCGCCCCTTGGTGACGAAGACGACCTGATTTCCCTGCACGAGCGGATGGTCGATT
TCGCCGTGAAGGCGCTGCGGGAGGCAAAGCTCAGGACGAGCTGGGACGATCCGAACGAGC
GCTATGAAGCAGCGATCAAGGCCTTTCTCGGCGATCTCCTGGATCGGCATAACCGGTCTT
TCCTCGGCGATTTCGAGAAAACGGCCGGACCGTTCATCCATGCCGGGCTGATCAACAGCC
TCTCGCAAGCGCTGGTGAAGCTCACCGCGCCCGGNATCCCCGATT

This sequence was obtained from pTH215, using the -40 universal primer (Figure 2-8).

Figure 5-11: DNA Sequence J59

CACGACAGGTTTGCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG
CTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
ATTGTGAGCGGATAACAATTTCACACAGGACACAGCTATGACATGATTACGAATTCGCTG
GCGGCAGGCCCGCGTGGCGNCATCATGCTGCAGGATTATCAGCTCATCGAGAAGCTGGCG
TCAGCACCAGAACCGCGAGCGCATTCCGGAACGCGTCGTCCATGCCAAGGGCTGGGGTGC
TTTCGGCTCGCTGAAGATCACCGGCGACATCTCGCAATACACCAGGGCGAAGTGTCTCCA
GCCCGGCGCGGAAACGCCGATGCTGGCCCGTTTCTCGACCGTTGCCGGCGAACAGGGCGC
CGCCGACCACGAGCGCGACGTCCGCGGTTTCGCGCTGAAATTCTACACCGATGAAGGCAA
CTGGGACCTCGTCGGcAACAATACGCCGGTCTTCTTCATTCGCGATCCCTACAAGTTCCC
CGACTTCATCCACACGCAGAAGCGTCATCCGAAGACCAATCTCCGTTCCGCCACTGCCAT
GTGGGACTACTGGTCCNTGTCGCCGGAAAGCCTGCATCAGGTGACGATCCTGATGTCCGA
CCGCGGCCTCCGCAGACGCCGATGAATNTGAACGGTTACGGGTCGCAAACTNTTCNTCTG
GAAGAGGC

This sequence was obtained from pTHJ59, an *Eco*RI ligation of RmH339 (Rm 5000 *katE878*::Tn*V*) by sequencing out from both sides of the transposon using the *IS*50 primer. This was done by cloning the 5.5 kb *Eco*RI-*Hind*III fragment from pTH329 containing the left *IS*50 element of the Tn*V* into pUC118, and by cloning the 1.3 kb *Eco*RI-*Hind*III fragment of pTHJ59 into pUC118. The two sequences were joined to form J59, minus the bases duplicated by the transposon during insertion.

References

Aldrich, T.L., B. Frantz, B. Gill, J.F. Kilbane, A.M. Chakrabarty. 1987. Cloning and Complete Nucleotide Sequence Determination of the *catB* Gene Encoding *cis,cis*-Muconate Lactonizing Enzyme. Gene **52**:185-195.

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

Assinder, S.J., and P.A. Williams. 1990. The Tol Plasmids: Determinants of the Catabolism of Toluene and the Xylenes. Advances Micro. Physiol. **31**: 2-69.

Baldani, J.I., R.W. Weaver, M.F. Hynes, and B.D. Eardly. 1992. Utilization of Carbon Substrates, Electrophoretic Enzyme Patterns, and Symbiotic Performance of Plasmid-Cured Clover Rhizobia. Appl. Environ. Microbiol. **58**:2308-2314.

Bauer, W.D. 1981. Infection of legumes by *Rhizobia*. Ann. Rev. Plant Physiol. **32**:407-449.

Belefort, M. M.E. Reaban, T. Coetzee and J.Z. Dalgaard. 1995. Prokaryotic Introns and Inteins: a Panoply of Form and Function. J. Bacteriol. **177**:3897-3903.

Berg, C.M. and D.E. Berg. 1987. Uses of transposable elements and maps of known insertions, p. 1071-1107. In F.C. Niedhardt, J.L. Ingraham, K.B. Low, B. Magaranik, M. Schaecter and H.E. Umbarger (eds.), *E. coli* and *S. typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.

Beringer, J.E., S.A. Hoggan, and A.W.B. Johnston. 1978. Linkage mapping in *Rhizobium leguminosarum* by means of R plasmid-mediated recombination. J. Gen. Microbiol. **104**:201-207.

Canovas, J.L. and R.Y. Stanier. 1967. Regulation of the enzymes of the β-ketoadipate pathway in *Moraxella calcoacetica* 4. Constitutive synthesis of β-ketoadipate succinyl-CoA transferases II and III. Eur. J. Biochem. **1**:289-300.

Canovas, J.L., M.L. Wheelis and R.Y. Stanier. 1967. Regulation of the enzymes of the β-ketoadipate pathway in *Moraxella calcoacetica* 2. The role of protocatechuate as inducer. Eur. J. Biochem. **3**:293-300.

Chari, R.V.J., C.P. Whitman, and J.W. Kozarcik. 1987. Absolute Stereochemical Course of Muconolactone Δ-Isomerase and of 4-carboxymuconolactone Decarboxylase. A ^{1}H NMR "Ricochet" Analysis. J. Am. Chem. Soc. **109**:5520-5521.

Charles, T.C. 1990. Construction of a Genetic Linkage Map of the *Rhizobium meliloti* 1600 kilobase Megaplasmid pRmeSU47b, Generation of Defined Megaplasmid Deletions, and Study of Megaplasmid-borne genes. Ph.D. thesis, McMaster University.

Charles, T.C. and T.M. Finan. 1991. Analysis of a 1600-Kilobase *Rhizobium meliloti* Megaplasmid Using Defined Deletions Generated *in Vivo*. Genetics **127**:5-20.

Chu, F.K., G.F. Maley, F. Maley and M. Befort. 1984. Intervening sequence in the thymidylate synthase gene of bacteriophage T4. Proc. Natl. Acad. Sci. USA **81**:3049-3053.

Coon, S.L., S. Kotob, B.B. Jarvis, S. Wang, W.C. Fuqua and R.M. Weiner. 1994. Homogentisic Acid is the Product of MelA, Which Mediates Melanogenesis in the Marine Bacterium *Shewanella colwelliana* D. Appl. Environ. Microbiol. **60**:3006-3010.

Denoya, C.D., D.D. Skinner and M.R. Morgenstern. 1994. A *Streptomyces avermitilis* Gene Encoding a 4-Hydroxyphenylpyruvic Acid Dioxygenase-Like Protein That Directs the Production of Homogentisic Acid and an Ochronotic Pigment in *Escherichia coli*. J. Bacteriol. **176**:5312-5319.

DeVos, G.F., G.C. Walker and E.R. Signer. 1986. Genetic manipulations in *Rhizobium meliloti* utilizing two new transposon Tn5 derivatives. Mol. Gen. Genet. **204**:485-491.

DiMarco, A.A., B.A. Averhoff, E.E. Kim and L.N. Ornston. 1993a. Evolutionary divergence of *pobA*, the structural gene encoding *p*-hydroxybenzoate hydroxylase in an *Acinetobacter calcoaceticus* strain well-suited for genetic analysis. Gene **125**:25-33.

DiMarco, A.A., B. Averhoff, and L.N. Ornston. 1993b. Identification of the Transcriptional Activator *pobR* and

Characterization of Its Role in the Expression of *pobA*, the Structural Gene for *p*-Hydroxybenzoate Hydroxylase in *Acinetobacter calcoaceticus*. J. Bacteriol. **175**:4499-4506.

Doten, R.C., L.A. Gregg and L.N. Ornston. 1987. Influence of the *catBCE* Sequence on the Phenotypic Reversion of a *pcaE* Mutation in *Acinetobacter calcoaceticus*. J. Bacteriol. **169**:3157-3180.

Doten, R.C., K. Ngai, D.J. Mitchell and L.N. Ornston. 1987. Cloning and Genetic Organization of the *pca* Gene Cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169**:3168-3174.

Durham, D.R., L.A. Stirling, L.N. Ornston and J.J Perry. 1980. Intergenic evolutionary homology revealed by the study of protocatechuate 3,4-dioxygenases from *Azotobacter vinelandii*. Biochemistry **19**:149-155.

Elsemore, D.A. and L.N. Ornston. 1994. The *pca-pob* Supraoperonic Cluster of *Acinetobacter calcoaceticus* Contains *quiA*, the Structural Gene for Quinate-Shikimate Dehydrogenase. J. Bacteriol. **176**:7659-7666.

Evans, W.C. 1947. Oxidation of Phenolic and Benzoic Acid by some soil bacteria. Biochem. J. **41**:373-382.

Farr, S.B. and T. Kogoma. 1991. Oxidative Stress Responses in *Escherichia coli* and *Salmonella typhimurium*. Microbiol. Rev. **55**:561-585.

Ferat, J., and F. Michel. 1993. Group II self-splicing introns in bacteria. Nature **364**:358-361.

Ferat, J., M. Le Gouar and F. Michel. 1994. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. C. R. Acad. Sci. Paris, Sciences de la vie **317**:141-148.

Ferrante, A.A., J. Augliera, K. Lewis and A.M. Klibanov. 1995. Cloning of an organic solvent-resistance gene in *Escherichia coli*: The unexpected role of alkylhydroperoxide reductase. Proc. Natl. Acad. Sci. USA **92**:7617-7621.

Figurski, D.H. and D.R. Helsinki. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. USA. **76**:1648-1652.

Finan, T.M., E. Hartwieg, K. Lemieux, K. Bergman, G.C. Walker and E.R. Signer. 1984. General Transduction in *R. meliloti*. J. Bacteriol **159**:120-124.

Finan, T.M., B. Kunkel, G.F. DeVos, and E.R. Signer. 1986. Second Symbiotic Megaplasmid in *Rhizobium meliloti* carrying Exopolysaccharide and Thiamine Synthetic Genes. J. Bacteriol. **167**:66-72.

Frazee, R.W., D.M. Livingston, D.C. LaPorte and J.D. Lipscomb. 1993. Cloning, Sequencing, and Expression of the *Pseudomonas putida* Protocatechuate 3,4-Dioxygenase Genes. J. Bacteriol. **175**:6194-6202.

Frazier, W.C. and D.C. Westhoff. 1988. Food Microbiology, 4th ed. New York: McGraw-Hill Book Company.

Freidman, A.M., S.R. Long, S.E. Brown, W.J. Buikema and F.M. Ausubel. 1982. Construction of a broad host range cloning vector and its use in the Genetic Analysis of *Rhizobium* mutants. Gene **18**:289-296.

Furuichi, T., M. Inouye and S. Inouye. 1985. Novel one-step cloning vector with a transposable element: application to the *Myxococcus xanthus* genome. J. Bacteriol. **164**:270-275.

Gerischer, U. and L.N. Ornston. 1995. Spontaneous Mutations in *pcaH* and *-G* Structural Genes for Protocatechuate 3,4-Dioxygenase in *Acinetobacter calcoaceticus*. J. Bacteriol. **177**:1336-1347.

Gilson, E., J.M. Clement, D. Brutlag, and M. Hofnung. 1984. A family of dispersed repetitive extragenic palindromic DNA sequences in *E. coli*. EMBO J. **3**:1417-1421.

Gish, W. and D.J. States. 1993. Identification of protein coding regions by database similarity search. Nat. Genet. **3**:266-272.

Hartnett, C., E.L. Neidle, K. 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., N.N Nichols, M. Kim, J.L. Ditty and R.E. Parales. 1994. Identification of the *pcaRKF* Gene Cluster from *Pseudomonas putida*: Involvement in Chemotaxis, Biodegredation, and Transport of 4-Hydroxybenzoate. J. Bacteriol. **176**:6479-6488.

Heipieper, J.J., H. Keweloh and H.J. Rehm. 1991. Influence of phenols on growth and membrane permeability of free and immobilized *Escherichia coli*. Appl. Environ. Microbiol. **57**:1213-1217.

Higgins, D.G. and P.M. Sharp. 1989. Fast and sensitive multiple sequence alignment. CABIOS **5**:151-153.

Higgins, D.G., A.J. Bleasby and R. Fuchs. 1991. Clustal V: improved software for multiple sequence alignment. CABIOS **8**:189-191.

Hughes, E.J., M.K. Shapiro, J.E. Houghton and L.N. Ornston. 1988. Cloning and Expression of *pca* Genes from *Pseudomonas putida* in *Escherichia coli*. J. Gen. Micro. **134**:2877-2887.

Hulton, C.S., C.F. Higgins, and P.M. Sharp. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli, Salmonella typhimurium* and other enterobacteria. Mol. Microbiol. **5**:825-834.

Hussien, Y.A., M.S. Tewfik and Y.A. Hamdi. 1974. Degradation of Certain Aromatic Compounds by Rhizobia. Soil. Biol. Biochem. **6**:377-381.

Imlay, J.A and I. Fridovich. 1991. DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. Science **240**:640-642.

Ingledew, W.M., M.E.F. Tresguerres and J.L. Canovas. 1971. Regulation of the Enzymes of the Hydroaromatic Pathway in *Acinetobacter calcoaceticus*. J. Gen. Micro. **68**:273-282.

Joeseph, S., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, New York.

Jones, J.D.G. and N. Gutterson. 1987. An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. Gene **61**:299-306.

Kowalchuk, G.A., G.B. Hartnett, A. Benson, J.E. Houghton, K. Ngai and L.N. Ornston. 1994. Contrasting patterns of evolutionary divergence within the *Acinetobacter calcoaceticus pca* operon. Gene **146**:23-30.

Keen, N.T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene **70**: 191-197.

Lambowitz, A.M., and M. Belfort. 1993. Introns as mobile genetic elements. Annu. Rev. Biochem. **62**:587-622.

LeRouge, P., P. Roche, C. Faucher, F. Maillet, G. Truchet, J.C. Prome, and J. Denarie. 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature **344**:781-784.

Lupski, J.R., and G.M. Weinstock. 1992. Short, interspersed repetitive DNA Sequences in prokaryotic genomes. J. Bacteriol. **174**:4525-4529.

Meade, H.M., S.R. Long, G. Ruvkun, S.E. Brown and F.M. Ausubel. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon mutagenesis. J. Bacteriol. **149**:114-122.

Miller, J.H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor New York: Cold Spring Harbor Laboratory Press.

Neidle, E.L., M. 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. Bonitz and L.N. Ornston. 1988. DNA Sequence of the *Acinetobacter calcoaceticus* Catechol 1,2-Dioxygenase Structural Gene *catA*: Evidence for Evolutionary Divergence off Intradiol Dioxygenases by Acquisition of DNA Repetitions. J. Bacteriol. **170**:4874-4880.

Neidle, E.L., C. Hartnett and L.N. Ornston. 1989. Characterization of *Acinetobacter calcoaceticus catM*, a Repressor Gene Homologous in Sequence to Transcriptional Activator Genes. J. Bacteriol. **171**:5410-5421.

Ornston, L.N and E.L. Neidle. 1991. Evolution of Genes for the β-ketoadipate Pathway in *Acinetobacter calcoaceticus*. In *The Biology of Acinetobacter* (K.J. Towner et al. Ed). New York: Plenum Press.

Osteras, M., J. Stanley and T.M. Finan. 1995. Identification of Rhizobium-Specific Intergenic Mosaic Elements (RIMEs) within an Essential Two-Component Regulatory System of *Rhizobium* Species. J. Bact. **177**: 5485-5494.

Parales, R.E. and C.S. Harwood. 1992a. Characterization of the Genes Encoding β-ketoadipate:Succinyl-Coenzyme A Transferase in *Pseudomonas putida*. J. Bacteriol. **174**:4657-4666.

Parales, R.E and C.S. Harwood. 1992b. Regulation of the *pcaIJ* Genes for Aromatic Acid Degradation in *Pseudomonas putida*. J. Bacteriol. **175**:5829-5838.

Parke, D. and L.N. Ornston. 1984. Nutritional Diversity of *Rhizobiaceae* Revealed by Auxonography. J. Gen. Micro. **130**:1743-1750.

Parke, D. and L.N. Ornston. 1986. Enzymes of the β-ketoadipate Pathway are Inducible in *Rhizobium* and *Agrobacterium* spp. and Constitutive in *Bradyrhizobium* spp. J. Bacteriol. **165**:288-292.

Parke, D., R. Rynne and A. Glenn. 1991. Regulation of Phenolic Catabolism in *Rhizobium leguminosarum* Biovar trifolii. J. Bacteriol. **173**:5546-5550.

Parke, D. 1993. Positive Regulation of Phenolic Catabolism in *Agrobacterium tumefaciens* by the *pcaQ* Gene in Response to β-Carboxy-*cis,cis*-Muconate. J. Bacteriol. **175**:3529-3535.

Parke, D. 1995. Superoperonic Clustering of *pca* Genes for Catabolism of the Phenolic Compound Protocatechuate in *Agrobacterium tumefaciens*. J. Bacteriol. **177**:3808-3817.

Parrott, S., S. Jones and R.A. Cooper. 1987. 2-Phenylethylamine Catabolism by *Escherichia coli* K12. J. Gen. Microbiol. **133**: 347-351.

Parsek, M.R., D.L. Shinabarger, R.K. Rothmel and A.M. Chakrabarty. 1992. Roles of CatR and *cis-cis*-Muconate in Activation of the *catBC* Operon, Which Is Involved in Benzoate Degradation in *Pseudomonas putida*. J. Bacteriol. **174**:7798-7806.

Podila, G.K., S. Kotagiri and S. Shantharam. 1993. Cloning of Protocatechuate 3,4-Dioxygenase Genes from *Bradyrhizobium japonicum* USDA110. App. Env. Microbiol. **59**:2717-2719.

Rao, J.R., N.D. Sharma, J.T.G. Hamilton, D.R. Boyd and J.E. Cooper. 1991. Biotransformations of the Pentahydroxy Flavone Quercetin by *Rhizobium loti* and *Bradyrhizobium* Strains (*Lotus*). Appl. Enviro. Microbiol. **57**:1563-1565.

Rao, J.R. and J.E. Cooper. 1994. Rhizobia Catabolize *nod* Gene-Inducing Flavonoids via C-Ring Fission Mechanisms. J. Bacteriol. **176**: 5409-5413.

Rothmel, R.K., D.L. Shinabarger, M.R. Parsek, T.L. Aldrich and A.M. Chakrabarty. 1991. Functional Analysis of the *Pseudomonas putida* Regulatory Protein CatR: Transcriptional Studies and Determination of the CatR DNA-Binding Site by Hydroxyl-Radical Footprinting. J. Bacteriol. **173**:4717-4724.

Romero-Arroyo, C.E., M.A. Schell, G.L. Gaines III, and E.L. Neidle. 1995. *catM* Encodes a LysR-Type Transcriptional Activator Regulating Catechol Degredation in *Acinetobacter calcoaceticus*. J. Bact **177**: 5891-5898.

Romero-Steiner, S , R.E. Parales, C.S. Harwood, and J.E. Houghton. 1994. Characterization of the *pcaR* Regulatory Gene form *Pseudomonas putida*, Which Is Required for the Complete Degradation of *p*-Hydroxybenzoate. J. Bacteriol. **176**:5771-5779.

Ruetschi, U., B. Odelhog, S. Lindstedt, J. Barros-Soderling, B. Persson and H. Jornvall. 1992. Characterization of 4-hydroxyphenylpyruvate dioxygenase. Eur. J. Biochem. **205**:459-466.

Schell, M.A. 1993. Molecular Biology of the LysR Family of Transcriptional Regulators. Ann. Rev. Microbiol. **47**:597-626.

Saldanha, R., G. Mohr, M. Belfort, and A.M. Lambowitz. 1993. Group I and group II introns. FASEB J. **7**:15-24.

Shanley, M.S., A. Harrison, R.E. Parales, G. Kowalchuk, D.J.

Mitchell and L.N. Ornston. 1994. Unusual G+C content and codon usage in *catIJF*, a segment of the *ben-cat* supra-operonic cluster in the *Acinetobacter calcoaceticus* chromosome. Gene **138**:59-65.

Sharples, G.J, and R.G. Lloyd. 1990. A novel repeated DNA sequence located in the intergenic regions of bacterial chromosomes. Nucleic Acids Res. **18**:6503-6508.

Simon, R. J. Quandt and W. Klipp. 1989. New derivatives of transposon Tn5 suitable for mobilizations of replicons, generation of operon fusions and induction of genes in gram negative bacteria. Gene **80**:161-169.

Siqueira, J.O., M.G. Nair, R. Hammerschmidt and G.R. Safir. 1991. Significance of Phenolic Compounds in Plant-Soil Microbial Systems. Crit. Rev. Plant Sci. **10**:63-121.

Stanier, R.Y., and L.N. Ornston. 1973. The β-ketoadipate Pathway. Advances Microbiol. Physiol. **9**:89-151.

Stern, M.J., F.F. Ames, N.H. Smith, E.C. Robinson, and C.F. Higgins. 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37**:1015-1026.

Thompson, J.D., D.G. Higgins and T.J. Gibson. 1994. ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, postion specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673-4680.

Wheelis, M.L. 1975. The Genetics of Dissimilarity Pathways in *Pseudomonas aeruginosa*. Ann. Rev. Microbiol. **29**:505-524.

Williams, S.E., E.M. Woolridge, S.C. Ransom, J.A. Landro, P.C. Babbitt and J.W. Kozarich. 1992. 3-Carboxy-*cis,cis*-muconate Lactonizing Enzyme from *Pseudomonas putida* Is Homologous to the Class II Fumarase Family: A New Reaction in the Evolution of a Mechanistic Motif. Biochemistry **31**:9768-9776.

Winstanley, C., S.C. Taylor and P.A. Williams. 1987. pWW174: A large plasmid from *Acinetobacter calcoaceticus* encoding benzene catabolism by the β-ketoadipate pathway. Mol. Microbiol. **1**:219-227.

Woese, C.R. 1987. Bacterial Evolution. Microbiol. Reviews **51**:221-271.

Wu, C.H., M.K. Ornston and L.N. Ornston. 1972. Genetic control of enzyme induction in the beta-ketoadipate pathway of *Pseudomonas putida*: Two point crosses with a regulatory mutant strain. J. Bacteriol. **109**:796-802.

Yarosh, O.K. 1989. Analysis of C_4-Dicarboxylic Acid Transport Genes in *Rhizobium meliloti*. M. Sc. Thesis, McMaster University.

Yarosh, O.K., T.C. Charles and T.M. Finan. 1989. Analysis of C_4-dicarboxylate transport genes in *Rhizobium meliloti*. Mol. Micro. **3**:813-823.

Zylstra, G.J., R.H. Olsen and D.P. Ballou. 1989a. Cloning, Expression, and Regulation of the *Pseudomonas cepacia* Protocatechuate 3,4-Dioxygenase Genes. J. Bacteriol. **171**:5907-5914.

Zylstra, G.J., R.H. Olsen and D.P. Ballou. 1989b. Genetic Organization and Sequence of the *Pseudomonas cepacia* Genes for the Alpha and Beta Subunits of Protocatechuate 3,4-Dioxygenase. J. Bacteriol. **171**:5915-5921.