Genetic analysis of indigenous plasmid, pPDL2 of *Flavobacterium* sp. ATCC 27551and its use in engineering benzoate degrading *Acinetobacter* sp. DS002

Thesis submitted for the Degree of
Doctor of Philosophy
In
Animal Sciences

Supervisor
Prof. S. Dayananda

By
P. Emmanuel Vijay Paul

Dept. of Animal Sciences
School of Life Sciences
University of Hyderabad
Hyderabad – 500 046, India

March, 2011

Enrolment No. 05LAPH06

University of Hyderabad

(Central University established in 1974 by an act of Parliament)

HYDERABAD-500 046, INDIA

CERTIFICATE

This is to certify that **Mr. P. Emmanuel Vijay Paul** has carried out the research work embodied in the present thesis under my supervision and guidance for a full period prescribed under the Ph.D. ordinance of this University. We recommend his thesis entitled **"Genetic analysis of indigenous plasmid, pPDL2 of *Flavobacterium* sp. ATCC 27551 and its use in engineering benzoate degrading *Acinetobacter* sp. DS002"** for submission for the degree of Doctor of Philosophy in Animal Sciences of this University.

Prof. S. Dayananda _____
Supervisor

Head _____
Department of Animal Sciences

Dean _____
School of Life Sciences

University of Hyderabad

(Central University established in 1974 by an act of Parliament)

HYDERABAD-500 046, INDIA

Declaration

I hereby declare that the work embodied in this thesis entitled **"Genetic analysis of indigenous plasmid, pPDL2 of *Flavobacterium* sp. ATCC 27551 and its use in engineering benzoate degrading *Acinetobacter* sp. DS002"** has been carried out by me under the supervision of Prof. S. Dayananda and this has not been submitted for any degree or diploma of any other university earlier.

Prof. S. Dayananda
(Research Supervisor)

P. Emmanuel Vijay Paul
(Research Scholar)

Acknowledgements

I would like to acknowledge my sincere thanks and heartfelt gratitude to my supervisor ***Prof. S. Dayananda*** *for his guidance, unwavering support throughout the course of my work.*

I thank the Dean ***Prof . Ramanadham*** *and the former Dean* ***Prof. A. S. Raghavendra*** *for allowing me to use the school facilities.*

My thanks to the Head, Department of Animal Sciences, ***Prof . Manjula Sritaran,*** *and former Head* ***Prof. S. Dayananda*** *and* ***Prof. Aparna Dutta Gupta*** *for allowing me to use the Department facilities.*

I would like to thank my Doctoral Committee members ***Prof . Manjula Sritaran*** *and* ***Prof. Appa Rao Podile*** *for their valuable suggestions.*

I would like to thank ***all the faculty members*** *of School of Life Sciences for their help whenever needed. I thank* ***Dr. Ch. Venkatramana*** *for his valuable suggestions. I also thank* ***Dr. O.H. Shetty*** *and* ***Prof. P. Reddanna*** *for allowing me to use their lab facilities. I thank* ***Gopi*** *for helping me in the oxygraph studies.*

I thank ***Mr. Anil, Mansi, Shailly,*** *and* ***Devilal*** *for their timely help during annotation of the sequences.*

I thank ***CSIR for*** *giving me financial assistance throughout my Doctoral studies. I thank the University for providing the financial assistance to visit TAMU, USA.*

I thank ***DST, DRDO, CSIR, UGC*** *for funding our laboratory .*

I thank ***Prof. James R. Wild****, Dept of Biochemistry and Biophysics, TAMU for inviting me to his lab as a visiting Research Scholar. Many thanks to* ***Dr. Melinda Wales, Dr. Boris, Dr. Janet Grimesly*** *for their useful discussions and valuble suggestions. I thank* ***Manoj, Gaurav, Piush, Nerraj, Subrithi*** *and* ***Dheeraj*** *for making my stay at TAMU enjoyable.*

I thank ***Mr.Lallan, Mr. Ankinedu*** *,* ***Mr.Jagan,*** ***Mr. Gopi*** *and* ***Mrs. Jythothi*** *who helped me in different endeavors during my work here.*

I thank ***Dr. Madhura Rekha, & Mrs. Leena Bhashyam*** *for their technical help whenever needed.*

I thank my current ***lab mates Purushotham, Aleem, Gopi, Prasanna, Toshi, Swetha, Sunil, Santosh*** *and* ***Ashok*** *for their constant support during my work.*

I specially thank my senior colleagues ***Dr. Pakala Suresh Babu , Dr. Syed Khajamohiddin, Dr. Pandey*** *and* ***Dr. Elisha*** *for their constant support during my work.*

Thanks to ***Manikanta (lab Asst) , Krishna (Lab Attendant)*** *and* ***Mallesh*** *for their assistance in administrative work.*

I thank ***Dr. Rajesh, Dr. Vijay Chiatanya, Girish, Ranjith, Vijay*** *all my friends at School of Life Sciences for all the help and support and making my stay at University of Hyderabad memorable.*

Special thanks to ***Naresh, Dinakar, Kiran, Vinay, Hari, Pavan, Madhu, Boney, Kauser, Christopher*** *and all my friends at CC Church for all their support and prayers. I greatly acknowledge the support of my Pastors.* ***Rev.Pascal Prakash*** *and* ***Rev. Promod****.*

I am grateful to my mother, wife, my daughter and son for their love, affection and cooperation. I also thank my brother, sister and her family for their love and support. I thank my grandfather and grandmother for their prayers and blessings.

Special thanks to all my relatives and friends for supporting me in time of need.

I thank my ***LORD God, Jesus*** *for his Grace, Constant presence and for help all through my life.*

P. Emmanuel Vijay Paul

Contents

Introduction

Materials & Methods

Contents

Contents

Results & Discussion

Chapter-1

Contents

Chapter-2

Chapter-3

INTRODUCTION

"Nothing in biology makes sense except in the light of evolution". Theodosius Dobzhansky (1900-1975), a noted evolutionary biologist made this statement about four decades ago (Dobzhansky, 1973). A thorough literature survey only reveals more and more evidences in support of the statement, not even a single valid statement is seen that contradicts it. It is uniformly applicable for all types of living cells. No exceptions are seen based on their morphological complexities, habitats and life span etc.

Bacterial adaptations to various environmental conditions are believed to occur through the process of clonal divergence and periodic selection (Levin, 1981; Nojiri et al, 2004). Genetic variation, the very basis for clonal divergence, is considered as a prerequisite of evolution. They are generated either due to errors in DNA replication process or due to the failures in repair of DNA damages occurred for a variety of reasons. The mixed populations obtained through genetic variations are the substrates for natural selection, the process in which the living conditions of the organism plays a major role (Arber, 1993; Arber, 1995). The evolutionary geneticist, Arber has used the term 'evolution genes' for genes that are involved in generation of genetic variations (Arber, 2000). He has further grouped the genes into generators and modulators based on their ability to generate and modulate genetic variations (Arber, 2000). The generators are the genes that code for enzymes of DNA replication, transposable elements and site-specific recombinases that contribute for genetic variations (Arber, 2000). The genes that code for DNA repair enzymes and restriction-modification systems are considered as modulators due to their ability to modulate the frequency of genetic variations (Arber, 2000). In addition to the genetic factors, a number of non-genetic factors are shown to

be involved in generation of genetic variations (Arber, 2000). Non-genetic factors include intrinsic instability of the nucleotides (Goodman et al, 1993) and extrinsic factors such as structural flexibility of the biological molecules (Arber, 1993), chemical and physical mutagens etc.

1.1 Generation of genetic variations- Natural strategies

Three qualitatively different natural strategies are seen contributing for the overall production of the genetic variants (Arber, 2004). The first natural strategy is the local sequence changes which occur due to mutations. The generation of new biological function due to mutations is extremely rare in absence of natural selection on the primary function displayed by the gene. Alteration of an already existing function into a new function occurs at a greater frequency than evolution of new function *de novo* (Arber, 2004). This strategy also applies to functional domains, entire genes, as well as to the functional systems that are composed of different genes (Arber, 2004).

The second natural strategy is DNA rearrangement or intragenomic reshuffling of DNA which occurs due to recombination between homologous regions. Existence of related sequence motifs and functional domains in different functional genes is seen as an evidence for intragenomic reshuffling of DNA (Lagomarsino et al, 2009). The same is true for the reassortment of expression control signals with different reading frames serving in protein expression (Yamanaka et al, 1998)**.** Deletion of the DNA segments due to rearrangement process leads to loss of non-essential sequences from the genome (Davis et al, 2002). Examples for genetic variation through gene loss are seen in a number of instances. Genetic variations seen among various pathogenic mycobacterial strains can

be cited as a classical example for this phenomenon (Gómez-Valero et al, 2007). Similar situation is also seen in other organisms, like *Buchnera,* an endosymbiont of insects (Van Ham et al, 2003) and in *Bordetella pertusis,* a pathogenic strain (Parkhill et al, 2003). Comparison of genomes of pathogens such as *Bordetella pertusis, B. parapertusis* and *B. bronchiseptica* have shown existence of a very few unique genes and a common gene complement. Reduction in genome size of *B. pertusis* and *B. parapertusis* when compared with *B. bronchiseptica* shows that the two species might have evolved independently from the common ancestor *B. bronchiseptica* through gene loss (Parkhill et al, 2003). The third natural strategy for production of genetic variations is by acquisition of genetic information through the process of Horizontal Gene Transfer (HGT) or Lateral gene transfer (LGT) (Arber, 2006). Unlike the other two processes, genetic variations through HGT greatly facilitates for the rapid genetic adaptation of bacteria to environmental alterations. In bacteria, a number of strategies contribute to horizontal mobility of genes. Some of them just facilitate to lateral transfer of few genes, whereas the others have shown ability to mobilize entire genome. A brief description is given below before presenting the significance of HGT in bacterial adaptive responses to the environmental pollution, the primary aim of investigation of the present study.

1.2 Mechanisms of Horizontal gene transfer

Horizontal gene transfer in bacteria is an important driving force which plays a critical role in generating genetic variations among bacteria and contributes to organismal fitness by shaping their genome (Preston et al 1998; Hacker and Carniel, 2001). In bacteria, HGT occurs in three forms: transformation, transduction and conjugation. Brief descriptions of the three mechanisms are given here.

1.2.1 Transformation

Natural transformation is considered as the stable uptake, integration and functional expression of extracellular DNA that occurs under natural bacterial growth conditions. Transformation is the only mechanism that can potentially explain how bacteria acquire DNA from foreign species beyond the host range of mobile genetic elements or bacteriophages (Jain et al, 2002; Thomas and Nielsen, 2005). Prerequisites of natural transformation include existence of extracellular DNA in the environment, the presence of the bacterial cells and the stabilization ability of translocated DNA either by integration into the bacterial genome or by recircularization into a self-replicating plasmid (Thomas and Nielsen, 2005). Bacterial cells can only participate in transformation if the cells are in the physiological state of competence. Acquiring a physiological state of competence depends on expression of about 20-50 proteins which contribute to transgenes uptake, protection and maintenance (Dubnau, 1999; Alm and Mattick, 1997). With the exception of *Neisseria gonorrhoeae*, most naturally transformable bacteria develop time-limited competence in response to specific environmental conditions such as altered growth conditions, nutrient access, cell density or starvation (Yamane et al, 1999; Hamoen et al, 2001; Hui and Morrison, 1991; Frosch and Meyer, 1992; Karlin et al, 1997). DNA continuously enters into the environment through cellular decomposition, disruption, excretion from living cells and finally through existence of viral particles (Steinmoen et al, 2002). Secretion of DNA by bacteria at a specific stage of growth phase through specialized secretory system such as the Type IV secretion system is known as active excretion or secretion (Draghi and Turner, 2006). Active excretion of DNA has been reported for many genera of bacteria, including *Acinetobacter*, *Alcaligenes*,

Azotobacter, *Bacillus*, *Flavobacterium*, *Micrococcus*, *Pseudomonas* and *Streptococcus* (Paget and Simonet, 1994; Lorenz and Wackernagel, 1994; Moscoso and Claverys, 2004). Upon exposure to competent bacteria, the extracellular DNA binds non-covalently to sites present on the cell surface and double stranded DNA is converted to single stranded DNA during the translocation across the inner membrane. Plasmids when translocated across the membrane reconstitute to a replicative form whereas the linear single stranded DNA must integrate into the bacterial genome to persist for further generations (Norman et al, 2009). For homologous recombination to occur the incoming DNA must have 25 to 200 bp sequence of high similarity with the recipient genome for initiation of DNA pairing and strand exchange. The fate of internalized DNA to integrate in to the bacterial genome is varied and depends on the species (Thomas and Nielsen, 2005). It is estimated that only 0.1% of internalized DNA is integrated in *Acinetobacter baylyi* whereas 25 to 50% of internalized DNA is integrated in *B. subtilis* and *S. pneumonia* (Palmen and Hellingwerf, 1997). In addition to homologous recombination, the internalized DNA may undergo DNA replacement through illegitimate recombination but it is shown to occur at low frequency (de Vries et al, 2001; Nielsen, 2003). Instead of DNA replacement, DNA additive integration is also possible through homologous recombination or homology-facilitated illegitimate recombination (HFIR). Recombination that occurs between two circular molecules or between a circular molecule and a bacterial chromosome leads to additive integration (Dempsey and Dubnau, 1989). In HFIR, recombination is initiated at the region of high similarity and strand exchange extends into regions of little or no sequence similarity, resulting in substitution of DNA sequences or integration of additional DNA sequences (Meier and Wackernagel, 2003; de

Vries et al, 2004). Instances of HFIR are reported from *A.baylyi* and *S. pneumonia* in which a single region of approx 200 bp in length has shown to initiate recombination events that lead to the additive integration of more than 1000 bp long heterologous DNA fragments into the bacterium's genome. Some of the host DNA sequences are shown to be deleted during integration process (de Vries et al, 2004; Prudhomme et al, 2002). HFIR events that initiate at regions of high sequence similarity are shown to occur at very high frequency than that of the illegitimate recombination (de Vries and Wackernagel, 2002; Hulter and Wackernagel, 2008).

Recombination frequencies are higher in certain strains and species that express the mutator phenotypes, the property where the frequency of mutations are shown to be high when compared to the wild type strains (Taddei et al, 1997). Mutator phenotypes can arise transiently in response to altered growth conditions (Feng et al, 1996) or more stably through mutations in the genes that belong to the DNA-repair and maintenance machinery (Matic et al, 1997; Funchain et al, 2001; LeClerc et al, 1996). Methyl-directed mismatch (MMR) repair system is initiated after replication and is involved in removal of base-base mismatches or small insertion-deletion loops generated by mis-incorporation or strand slippage during replication of DNA. MMR distinguishes the parental strand from the newly synthesized strand based on the methylation states. The newly synthesized strand is unmethylated whereas the parent strand is methylated. MMR system also repairs DNA after homologous recombination and usually results in non-incorporation of the exogenous DNA (Hsieh, 2001). Mutator stains with defects in the methyl-directed DNA repair genes has been shown to recombine at elevated frequencies (Vulic et al, 1997; Majewski and Cohan, 1998; Majewski et al, 2000; Young and Ornston, 2001). It has

been proposed that the mutator phenotypes have access to a different gene pool than normal cells (Townsend et al, 2003) as these strains allow incorporation of exogenous DNA into the chromosome which are otherwise prevented from incorporation by the MMR system.

1.2.2 Transduction

Transduction is a mechanism of DNA acquisition by which non-viral DNA can be transferred from an infected host bacterium to a new host via infection. Host DNA is fortuitously packaged into the empty phage head when the phage particle is produced. Defective phage particles which are released from lysed host cells can adsorb to new host cells and deliver the DNA carried in the capsid into the new host. The injected bacterial DNA can be integrated into the recipient genome. Although most bacteriophages infect only a narrow range of hosts, this mechanism of gene transfer has several advantages. Mainly, the transducing phages can persist under environmental conditions, require no cell-cell contact, and the DNA to be transduced is protected by the phage particles (Wommack and Colwell, 2000). Recent studies have shown existence of bacteriophages in soils and marine environments (Ashelford et al, 2003) and there is evidence of the presence of the phage related genes in bacterial genomes (Canchaya et al, 2003), suggesting that transduction is also a key player in HGT. Some of the examples of DNA acquisition by bacteria through transduction are the pathogenicity determinants (toxins) of *Corynebacterium diphtheria*, *Clostridium botulinum*, *Streptococcus pyogenes*, *Staphylococcus aureus* and Shiga toxin producing *E. coli* (Brüssow et al, 2004). Pathogenicity islands (PAI) and their non-pathogenic derivatives are thought to have evolved from lysogenic bacteriophages (Dobrindt et al, 2004; Hacker et al, 2003;

Grozdanov et al, 2004). Uptake of such genes usually contributes to general adaptability, fitness and competitiveness of the host to various environmental conditions.

1.2.3 Conjugation

The third mechanism of horizontal gene transfer is conjugation which is known to be quantitatively more important than transformation and transduction (Halary et al, 2010). The two main functions that are essential for persistence of plasmids are replication and horizontal spread. The horizontal spread of the plasmid occurs mainly by conjugation and the successful propagation of a plasmid by conjugation requires two important processes: the first one of them is the mating pair formation (MPF) and second one is the rolling circle replication (RCR) (Llosa et al, 2002). Conjugation of a plasmid is possible if it carries the two sets or modules of genes: *mob* genes that are involved in the DNA transfer replication and genes that code for the Type 4 secretion system (T4SS). The initial requirement for conjugation is mating pair formation (MPF), a complex where the donor and recipient connect physically through a sex pilus formed by the T4SS. Subsequently the pilus retracts to facilitate cell contact (Christie et al, 2005; Clarke et al, 2008). The second step in the conjugation process (DNA processing and transfer) involves relaxase-mediated nicking of the plasmid at origin of transfer (*oriT*) and formation of the relaxosome. Relaxosome is a nucleoprotein that contains single-stranded plasmid DNA, the relaxase and a number of other proteins (some host encoded)(Lanka and Wilkins, 1995). A cognate coupling protein usually helps the relaxosome to dock with T4SS, where it is subsequently transported into the recipient cell. Finally, RCR ensures that a second strand is synthesized in both the donor and the recipient (Lanka and Wilkins, 1995). Plasmids that code for MPF and relaxosome formation and DNA processing are known as

conjugative or self-transmissible plasmids but plasmids that carry only the genetic information necessary for relaxosome formation and DNA processing are termed as mobilizable plasmids (Lessard et al, 2004). Some plasmids are called non-mobilizable because they are neither conjugative nor mobilizable. Hence, plasmids can be classified into three categories based on their mobility: conjugative, mobilizable and non-mobilizable plasmids (Smillie et al, 2010).

1.3 Entities of Horizontal Gene Transfer (HGT)

Genome evolution by DNA acquisition occurs through various mobile genetic elements such as plasmids, bacteriophages, genomic islands, Integrons, transposons, conjugative transposons and IS elements (Dobrindt et al, 2004). Some of the characteristics of mobile genetic elements are briefly described in the following sections.

1.3.1 Plasmids

Plasmids are present in all branches of the 'bacterial tree of life' and have been found in all bacterial communities studied to date, including those living in soil, marine and clinical environments (Brainbridge and Typas, 1984; Stanisich, 1988; Del Solar et al, 1988). A plasmid is a collection of functional genetic modules that are organized into a stable, self-replicating entity or 'replicon', which usually does not contain genes required for essential cellular functions. Plasmids are defined as covalently closed, circular double stranded DNA molecules that replicate autonomously in a host cell. However, linear double stranded DNA molecules are found in *Borrelia burgdorferi*, *Streptomyces rochei* and *Thiobacillus versutus* etc., (Barbour and Garon , 1987; Harochika et al, 1984; Hinnebusch and Tilly, 2006) and they range in size from a few to several hundred

kilobases (Broda, 1979; Darnell et al, 1986). Plasmids replicate, control their copy number and ensure their inheritance at each cell-division by following various partitioning mechanisms (Austin and Abeles, 1985). Bacterial plasmids are typically composed of different regions or modules, which encode functions involved in plasmid replication, stability, and transfer, which make-up what is often defined as the plasmid backbone. Apart from these household functions most plasmids very often harbour one or more accessory elements that encode useful selectively advantageous traits to the host bacterium whereas some plasmids are small and does not code for any known function. These plasmids are called cryptic plasmids. Some of the advantageous traits are antibiotic resistance (R plasmids), use of novel carbon sources (catabolic plasmids), sugar utilization, colicin activity, nitrogen fixation and conjugation (Werneegreen et al, 1997; Zawadzaki et al, 1996; Riley and Gordon, 1992; Mercer et al, 1984).

1.3.1.1 Resistance plasmids

Dissemination of antibiotic resistance is possible through all the mobile genetic elements and the plasmids are the first ones to be identified as the entities involved in the spread of the antibiotic resistance genes. Plasmids that contain genes that confer antibiotic resistance to the host are known as resistance plasmids (Yang et al, 1976). Plasmid-encoded antibiotic resistance encompasses most, if not all classes of antibiotics currently in clinical use and includes resistance to many that are at the forefront of antibiotic therapy (Bennett, 2008). Notable among these are the commonly used β-lactams, cephalosporins, fluoroquinolones and aminoglycosides (Bennett, 2008).

Retrospective analysis has shown that the antibiotic resistance genes are acquired by the plasmids either by transposons or Integrons (Fig. 1.1) (Carattoli, 2003). A plasmid

borne resistance genes, located within a transposon or an integron, can be transferred to other strains and species, enabling it to penetrate into niches not accessible to its original host strain (O'Brien, 2002). As IS elements can generally be found at many different sites, particularly on different plasmids, the potential for dissemination of the resistance genes is high (Bennett, 2008). The transposition modules, the IS elements, generally retain their ability to transpose as individual elements as well as part of the compound structure

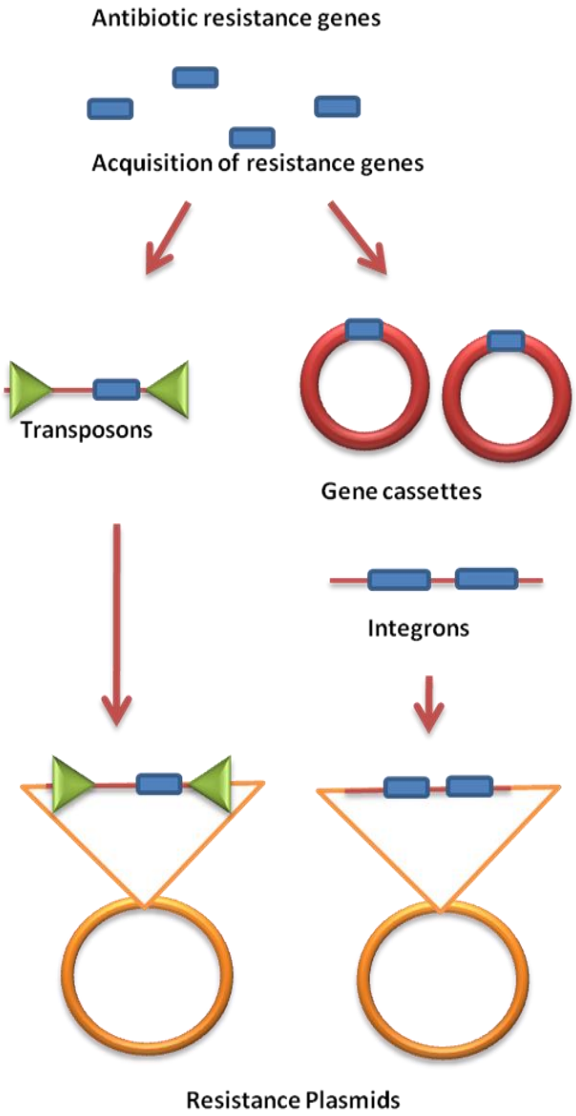


Fig. 1. 1. Schematic representation of plasmid mediated horizontal gene transfer of antibiotic resistance

whereas complex elements are indivisible with respect to transposition as they can only transposs as compound structures. The well-known transposons such as Tn*5*, encoding resistance to aminoglycosides such as kanamycin and neomycin, and Tn*10* encoding resistance to tetracycline, are compound elements found in a number of Gram-negative bacteria. The transposon Tn*3*, encoding resistance to a number of β-lactam antibiotics, including ampicillin, and Tn*21*, encodes resistance to streptomycin, spectinomycin and sulphonamides as well as mercuric ions. Tn*3* and Tn*21* are examples of complex transposons and are also commonly found on plasmids in members of enterobacteriaceae (Bahl et al, 2007; Bahl et al, 2009; Bennett, 2008).

Many resistance plasmids can also have conjugative or mobilizable functions. Accordingly, mobilizable resistance plasmids tend to be relatively small, often less than 10 kb in size, encoding only a handful of genes including the resistance gene(s), whereas conjugative plasmids tend to be somewhat larger, 30 kb or more, reflecting the sizable amount of DNA (20–30 kb) needed to encode the conjugation functions that permit cell-to-cell coupling, particularly between Gram-negative bacteria. One broad host range plasmid is the resistance plasmid RP1 (also known as RP4 and RK4), first identified in a clinical strain of *Pseudomonas aeruginosa*. This plasmid appears to be able to transfer to most, if not all Gram-negative bacteria (Salyers and Shoemaker, 1994). Many other, unrelated resistance plasmids are also known to have broad host ranges (Thomas, 1989).

1.3.1.2 Conjugative Plasmids

The F plasmid of *E.coli* is the well-studied example of the conjugative plasmids. Plasmids usually consist of relatively modest number of compactly arranged genes, often accompanied by *cis*-elements involved in their functions such as replication, segregation

and transfer. The genes usually organized into modules based on the function they perform.

a) Maintenance modules

Maintenance modules are involved in replication and control of the copy number of the plasmid. A large number of different mechanisms have evolved for the vertical transmission of circular plasmids (Giraldo, 2003), but most follow either of two schemes: one is the *theta* mechanism, in which replication initiates by melting the double stranded DNA at the *oriV*, thus allowing for assembly of the replisome. Replication may then proceeds either uni- or bi-directionally leading to the formation of a structure resembling the greek letter θ (Kues and Stahl, 1989). In the other mechanism, rolling circle replication is initiated from a 3'-OH primer, generated by nicking one strand of the plasmid, and then proceeds by strand displacement (del Solar et al, 1998). Plasmid copy number is almost universally controlled at the level of replication initiation by employing plasmid-encoded *trans*-activators or inhibitors. These are in the form of Rep replicases (del Solar et al, 1998).

b) **Plasmid stability modules**

Apart from maintaining stable copy numbers with replication, plasmids must cope with enzymatic degradation, damaging insertions of foreign DNA and homologous recombination that leads to the formation of plasmid multimers. If a plasmid fails to be transmitted to a daughter cell it leads to the formation of a plasmid –free or cured strain. Mechanisms that promote the stability of the plasmid by properly segregating them into the daughter cells are identified (Hiraga, 1992; Nordstrom and Austin, 1989; Williams and Thomas, 1992).

Multimer resolution

Smaller plasmids with high and medium copy number usually encounter the problem of multimer formation. Plasmid multimer formation potentially prevents segregation of the plasmid into the daughter cells (Summners et al, 1993). Many smaller plasmids have therefore acquired site-specific recombinase systems which are called as multimer resolution systems (*mrs*). The Xer-cer system of the ColE1 plasmid is a well-studied example (Summers, 1998). The host-encoded XerCD complex resolves the multimers formed at the recombinase site *cer* (Summners, 1998). Large plasmids of the Inc-P1 group solve the problem of multimer formation by employing resolvases of the *parCBA* system. The transposons that are present on the large plasmids also contain resolvases which are also shown to resolve the multimers formed (Tolmasky et al, 2000).

Active partitioning

Large plasmids relay on specialized active systems for proper segregation of the plasmids. One of the well-studied examples is the Par system comprises of an ATPase coded by *parA* and a DNA binding protein *parB* which act in trans on the centromere-like region sometime called as parS (Tolmasky et al, 2000; Hayes and Brilla, 2006). The ParAB proteins form a nucleo-protein complex with the *parS* region which is known as a segrosome (Hayes and Brilla, 2006). ATPase action of ParA provides the necessary energy for the intracellular movement of segrosome towards the ends of the cells during cell division thus allowing proper distribution of plasmids to the daughter cells (Bignell and Thomas, 2001).

Plasmid addiction

Another system which maintains stability of plasmids is the addiction system which employs the post-segregational killing of the cured strains. Usually a gene product of the plasmid addiction system performs an action that arrests the growth or kills the host (Gerdes et al, 1986). A second gene product counteracts the effects of the first gene. Some the well-studied examples are the toxin-antitoxin system are the Ccd system of F plasmid (Bernard and Couturier, 1992), ParDE system of the IncP plasmid RK2 (Jovanovic et al, 1994).

c) Dissemination modules

Two sets of genes are essential for successful self-transfer of a plasmid–genes responsible for mating pair formation (transferosome) and relaxosome formation. The genes on the *tra* operon are responsible for mating pair formation which forms a T4SS type secretion system. T4SS-like type T2SS and T3SS secretion systems are also employed by different groups of bacteria for the formation of extracellular filaments or pili (Lawley et al, 2003). Genes responsible for mobilising the plasmid is a site-specific relaxase which nicks DNA to give a single-stranded substrate that is suitable for transfer. Basing on the existence of these two sets of genes, plasmids are characterized into mobilizable or self-transmissible plasmids.

1.3.2 IS elements and Transposons

IS elements and transposons are mobile genetic elements that have the ability to move both intra and inter-molecularly from one site to another within a DNA molecule or from one DNA molecule to another and do not require any DNA homology between the element and the sites of insertion (Craig, 1997). IS elements are short mobile elements that code for a transposase and is flanked by two identical inverted repeats (IR). Mobile

elements can be distinguished by their structure, genetic relatedness and mechanism of transposition (Bennett, 2005). A transposon differs from an IS element if it encodes at least one function that changes the phenotype of the cell in a predictable fashion, for example a resistance transposon confers resistance to a particular antibiotic(s). Transposons are either modular systems, referred to as composite transposons, constructed from a pair of IS elements and a central DNA sequence that is not inherently able to transpose, the expression of which alters the cell phenotype or complex systems where transposition and non-transposition functions have not obviously been assembled in a modular fashion. Transposons are classified into three classes based on their structure (Fig. 1. 2).

Class I elements include simple Insertion sequences (IS) which carry only genetic determinants for its own transposition and composite transposons which carry genes for its own transposition and various other genes. The class II elements are ancestrally related elements, which are also known as Tn*3* family of transposons (Grindely, 2002). These elements encode for a large transposase of about 1000 amino acid residues and have similar inverted repeat (IR) sequences of 40 bp.

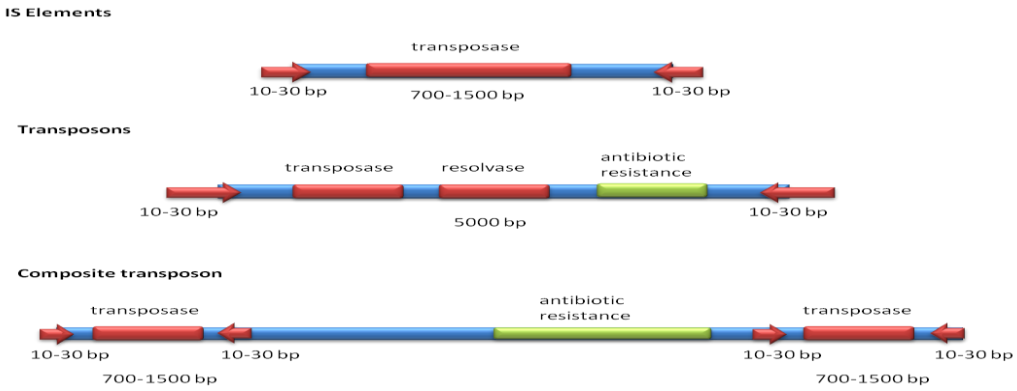


Fig. 1. 2. Schematic representation showing structural features of IS elements and transposons

The class II transposons transpose in a replicative pathway forming an intermediate called co-integrate. The resolution of the co-integrate is done at a *res* site by a site specific recombinase which resolves it into a target with a simple insertion and a regenerated donor replicon (Kitts et al, 1983). The resolution system of class II transposons work even when the *res* sites are far apart, a situation seen in large transposons. This resolution system allows successful resolution of large molecules and therefore is involved in transposition of large transposons (> 50 kb). Class II transposons are advantageous for dispersal of whole catabolic genes/operons due to their large size. The third class of bacterial transposons proposed by Grindely and Reed (1985) consists of bacteriophage Mu and related phages. This virus employs a transposition like mechanism for its replication.

1.3.3 Integrative and conjugative elements (ICEs)

ICEs are self-transmissible mobile genetic elements (MGEs) that encode conjugation machinery as well as for intricate regulatory systems to control their excision from the chromosome and their transfer by conjugation (Salyers et al, 1995; Osborn and Boltner, 2002; Burrus and Waldor, 2004). In conjugative plasmids, the conjugative functions are encoded by the plasmids whereas the conjugative functions of ICEs are coded by the chromosome itself (Woznik and Waldor, 2010). Therefore, they can be called as chromosome-borne self-transmissible elements (Burrus and Waldor, 2004; Woznik and Waldor, 2010). Unlike plasmids they are not self-replicating. ICEs encode site-specific recombinase known as Integrase (Int) which integrates the ICEs at specific sites (*attB*) into the chromosome of the host and they replicate as part of the host chromosome. Certain conditions induce the excision of the ICEs from the chromosome

which is the reverse process of integration and it also requires the enzyme, integrase (Int). The Int-mediated excision is brought about by recombination directionality factors (RDFs) which bias the processes towards excision than integration (Lewis and Hatfull, 2001). Excisionase (Xis) is one of the RDF involved in excision of ICEs such as Tn*916* (Hinerfeld and Churchward, 2001). Many factors prevent the premature excision of the ICEs and allow stable maintenance of the ICEs (Klockgether et al, 2004; Wozniak and Waldor, 2009). The ICE once excised undergoes processing in its extrachromosomal state before conjugal transfer. A relaxase binds to *oriT* and nicks it and the ICE element undergoes rolling circle replication (Lanka and Wilkins, 1995). A single stranded ICE received by the recipient is reconstituted it into a double stranded circular molecule by the host replicative machinery integrating into the chromosome at the *attB* sites (Burrus and Waldor, 2003). ICEs combine the features of all MGEs, such as phages, transposons and plasmids (Woznik and Waldor, 2010).

The genes that are present on the ICEs are varied; SXT the ICE of *Vibrio cholerae* has antibiotic resistance genes (Cm^R, Sm^R), PAPI-1 of *Pseudomonas aeruginosa* has virulence factors and regulation of biofilm formation (Drenkard and Ausubel, 2002; Qiu et al, 2006), The ICE, $MISym^{R7A}$ of *Mesorizobium loti* has the genes responsible for nodulation and nitrogen fixation (Sullivan and Ronson, 1998) and ICE*clc*^{B13} has genes responsible for degradation of 3-chlorobenzoic acid (Ravatn et al, 1998). Basing on studies conducted on plasmids and transposons, it is thought that expression of genes present on ICEs in a new host results in decrease of host fitness (Dahlberg and Chao, 2003; Doyle et al, 2007). However, studies on ICEs such as CTnDOT of *Bacteroides thetaiotamicron* VPI-5482 and ICE*clc*^{B13} of *Pseudomonas aeruginosa* PAO1 has shown

that the effects on host fitness is minimal due to expression of genes on these ICEs (Moon et al, 2004 and Gaillard et al, 2008).

1.3.4 Genomic Islands

A new type of genetic entity, known as pathogenicity islands (PAIs), has been shown to contribute to the pathogenicity of uropathogenic isolates of *Escherichia coli* (UPEC) (Hacker et al, 1990). PAIs found in their study were unstable chromosomal regions with variable virulence associated characteristics and phenotypes (Groisman & Ochman, 1996). Elements similar to PAIs are also found in non-pathogenic microbes which were named as Genomic Islands (GEIs). GEIs vary in size and are present in many bacterial genomes (Dobrindt et al, 2004). Different GEI families have been recognized on the basis of predicted sequence and functional homologies (Burrus et al, 2002; Juhas et al, 2007b; Vernikos & Parkhill, 2008). The coding capacity of GEIs is not limited to pathogenicity functions, but can be very diverse, including traits as symbiosis (Sullivan et al, 2002), sucrose and aromatic compound metabolism (Gaillard et al, 2006), mercury resistance and siderophore synthesis (Larbig et al, 2002b). The fact that GEIs come in a large spectrum of varieties in terms of genetic organization and functionality, it makes more difficult to provide an exact definition of a GEI. Juhas and co-workers have proposed that the term GEI should be used for the overarching family of discrete 'DNA elements', which are part of a cell's chromosome and can drive or have driven strain differentiation (Juhas et al, 2009).

These are the following features that are shared by most GEIs:

(1) GEIs are relatively large segments of DNA, usually between 10 and 200 kb detected by comparisons among closely related strains. Discrete DNA regions detected by comparative genome sequencing with sizes smaller than 10 kb have been named genomic islets (Hacker & Kaper, 2000).

(2) GEIs may be recognized by nucleotide statistics (e.g. GC content, cumulative GC skew, tetranucleotide frequencies or codon usage) that usually differ from the rest of the chromosome (Juhas et al, 2009).

(3) GEIs are often inserted at tRNA genes, in which case they might be ICEs (Woznik and Waldor, 2010).

(4) GEIs are often flanked by 16–20-bp perfect or almost perfect direct repeats (DR). DRs usually arise by the site-specific integration of the GEIs into the target site and can act as recognition sequences for their enzymatic excision (Schmidt & Hensel, 2004).

(5) GEIs often harbour functional or cryptic genes encoding integrases or factors related to plasmid conjugation systems or phages involved in GEI transfer (Juhas et al, 2009).

(6) GEIs often carry insertion elements or transposons, which may have been implicated in mobilizing genetic material onto or deleting DNA from the element (Buchrieser et al, 1998; Gal-Mor & Finlay, 2006).

(7) GEIs often carry genes offering a selective advantage for host bacteria. According to their gene content, GEIs are often described as pathogenicity, symbiosis, metabolic, fitness or resistance islands (Dobrindt et al, 2004; Schmidt & Hensel, 2004) (Fig. 1. 3).

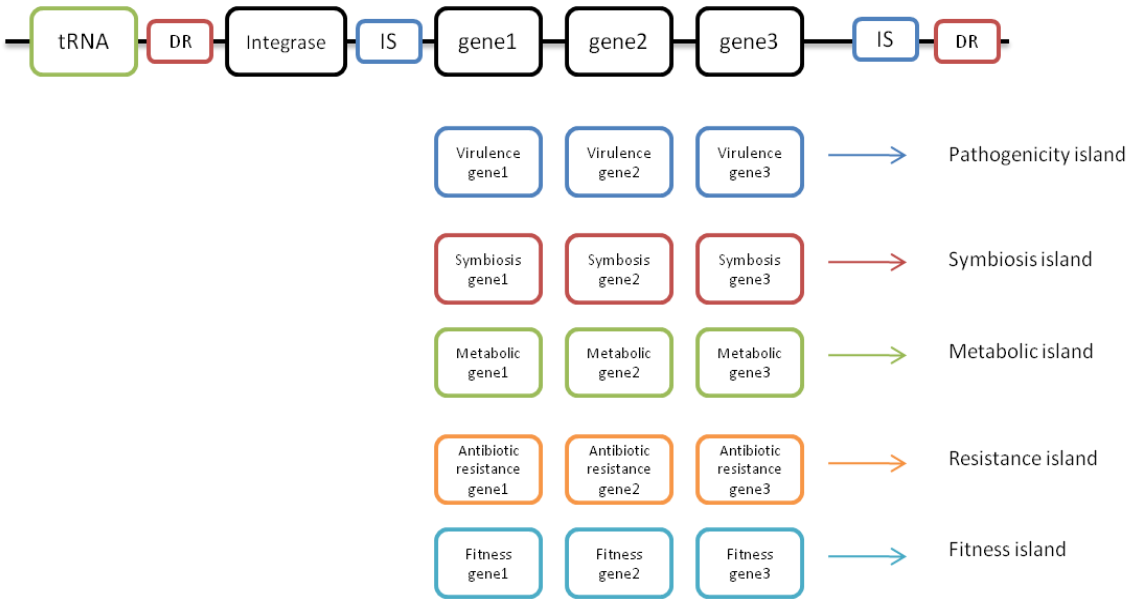


Fig. 1.3. Schematic diagram showing general features of Genomic Islands.

The integration of GEIs is catalyzed by site-specific phage-like recombinases called integrases (Int) and occurs usually at the 3'-end of tRNA genes (Ravatn et al, 1998a; Larbig et al, 2002b; Williams, 2002). Integrases are not strictly conserved among all GEIs, and have shown no preference for tRNA genes to be used as target sites (Williams, 2002). Several GEI-encoded integrases have shown similarities to the integrases coded by lambda, P4 or XerD families (van der Meer et al, 2001; Burrus et al, 2002; Mohd-Zain et al, 2004). Excision of GEIs is also catalyzed by the integrase which may be assisted by an excisionase (Burrus & Waldor, 2003; Lesic et al, 2004; Ramsay et al, 2006). The *int* gene encoding integrase is often situated at one extremity of the island and adjacent to the tRNA gene in the integrated GEI form (Juhas et al, 2009). Following excision, both GEI

ends close up to form a single copy of the recombination site (attP) (Ravatn et al, 1998b; Doublet et al, 2005; Qiu et al, 2006; Ramsay et al, 2006). Excision is mostly non-replicative and a single copy of the recombination site is again formed on the chromosome (attB) (Burrus & Waldor, 2003; Ubeda et al, 2007). GEI can reintegrate back into the attB site of the same host or, after transfer to a new cell, in a suitable attB site (Ravatan et al, 1998a; Williams, 2002). GEIs do not contain replication modules and hence integrate and replicates along with the chromosome. An exception to this is the Staphylococcal pathogenicity islands (SaPIs) which share many features of temperate phages and are known as Phage type GEIs. These Staphylococcal pathogenic islands are proven to be capable of existing in a self-replicating plasmid-like state (Ubeda et al, 2007, Ubeda et al, 2008).

1.3.5 Integrons

Integrons are defined as the genetic elements that encode a site-specific recombinase, known as DNA integrase, and a corresponding recombination site, *attI*, into which gene cassettes may be specifically inserted (Bennett, 1999). The first component of the integron is the site-specific recombinase which is a tyrosine recombinase involved in integration and excision of the gene cassette (Hall and Collis, 1995). Recombination most commonly targets an integron-associated attachment site (*attI*) immediately adjacent (upstream) to the integrase gene (*intI*) and a site, *attC*, found within individual mobile circular gene cassettes (Collis et al, 2002). Plasmid–borne Integrons of class-I usually carry genes that convey resistance to different groups of antimicrobial agents including β-lactams, aminoglycosides, trimethoprim, chloramphenicol, quaternary ammonium compounds, rifampicin, sulphonamides, macrolides, lincosamides and

quinolones (Levings et al, 2006; Fluit and Schmitz, 2004; Antunes et al, 2005; Partridge and Hall, 2005). Recent studies have shown that gene cassettes other than antibiotic resistance are also present on the Integrons of class I type (Stokes et al, 2001 and Stokes et al, 2006). Genes involved in heavy metal resistant is an example of non-antibiotic resistant class-I integrons (Silver, 1996). These Integrons are not mobile *per se*; however they are mobile as they are associated with transposons such as Tn*21* and Tn-*402* like transposons (Liebert et al, 1999; Bennett, 1999).

1.4 Xenobiotic-degradation islands

Industrial activity over the past century resulted in the release of various quantities of xenobiotic chemicals into the environment. Xenobiotics are defined in a broad sense as guest chemicals that are not natural to the environment (Leisinger, 1983). Many microbes have adapted well to utilized these xenobiotics as sources of carbon, nitrogen and energy (Reviewed in Nojiri et al, 2004). There have been striking similarities, in metabolic pathways and enzymes involved in degradation of xenobiotics even in phylogenetically unrelated bacterial strains isolated from geographically distinct areas (Nojiri et al, 2004). Such similarities in the biochemical features can in part be attributed to the similarity in the levels of gene organizations and nucleotide sequences (Top et al, 1995; Herrick et al, 1997; Vallaeys et al, 1999; Poelarends et al, 2000b; Sentchilo et al, 2000). This scenario is also seen due to intergenomic gene acquisition by horizontal (lateral) gene transfer (Arber, 2000). Horizontal gene transfer between environmental bacteria is mediated by mobile genetic elements (MGEs), such as transferable plasmids, (conjugative) transposons, integrons, genomic islands, or phages, which are all able to move within and/or between genomes, thus allowing “evolution in quantum leaps”

(Hacker and Carniel, 2001). IS elements are also seen associated with catabolic genes and various catabolic genes and gene clusters often present as discrete MGEs (Tan, 1999).

1.5 Catabolic plasmids

The ability of bacteria to degrade and mineralize xenobiotic compounds is often facilitated by expression of the genes residing on the plasmids. Plasmids having genes responsible for catabolism of xenobiotic compounds are called catabolic plasmids. Plasmids involved in degradation of organic compounds are the first one to be described (Chakrabarthy, 1972; Chakrabarthy, 1973). Later many plasmids were reported that were involved in degradation of various xenobiotics from strains of *Rhodococcous, Pseudomonas, Alacligenes, Sphingomonas* and *Terrabacter* (Nojiri et al, 2004). Sometimes identical catabolic genes are present on phylogenetically distinct bacteria which might be due to intergenomic gene acquisition by horizontal gene transfer (Arber et al, 2000). The presence of complete sets of transfer and catabolic genes on large indigenous plasmids (>50 kb) along with the presence of IS elements, transposons and ICEs strengthen the view that HGT plays a major role in dissemination of these catabolic plasmids (Top et al, 2000). Genes for degrading naturally occurring compounds seem to be located on IncP-2 or IncP-9 plasmid, like OCT, CAM, and pWW0 (Chakarabarthy 1973; Rheinwald et al, 1973; Tan, 1999; Williams and Murry, 1974; Greated et al, 2002). In contrast, genes for degrading man-made xenobiotics are frequently found on a broad-host-range IncP-1 plasmid, such as pJP4, pSS60, and pBRC60 (Top et al, 2000). IncP-1 plasmids are the most promiscuous self-transmissible plasmids characterized to date, with a host range that is much wider than that of IncP-2 and IncP-9 plasmids which act as shuttle vectors for locally adapted genes, resulting in the rapid local adaptation of both

phylogenetically related and distinct populations present in the same bacterial community.

1.6 Catabolic transposons

Apart from Xenobiotic degradative islands and catabolic plasmids, catabolic transposons play a major role in dissemination of the catabolic genes. Catabolic genes when present on transposons or when associated with transposable elements are prone to undergo genetic rearrangement contributing for evolution of novel catabolic pathways. Most of the catabolic transposons have structural features of the class I (composite) elements. Chlorobenzoate (Tn*5271*), chlorobenzene (Tn*5280*) and benzene catabolic transposons are some of the examples of the class I transposons.

1.6.1 Chlorbenzoate transposon (Tn*5271*)

The chlorobenzoate transposon (Tn*5271*) harbouring the catabolic genes (*cbaABC*) for degradation of chlorobenzoate are present on the plasmid pBRC60 of *Alcaligenes* sp. strain BR60 (Nakatsu et al, 1991). Tn*5271* is a 17kb in length flanked by two copies of *IS1071*, which are 3.2 kb direct repeats (Nakatsu and Wyndham, 1993; Wyndham et al, 1994a; Nakatsu et al, 1997). Tn*5271* was observed to be able to transpose on to the chromosomes of *Comamonas testosteroni* (ATCC 11996) and *Comamonas acidovorans* (ATCC 15668), always a portion of a larger encompassing element, which suggests that Tn*5271* may be part of a larger transposon (Wyndham et al, 1994b).

1.6.2 Chlorobenzene transposon (Tn*5280*)

The chlorobenzene transposon (Tn*5280*) harbours the chlorobenzene dioxygenase and dehydrogenase genes which form the upper pathway for chlorobenzene

degradation. The Tn*5280* is 5.1 kb and is present on the 100kb plasmid pP51 of *Pseudomonas* sp. strain P51 (van der Meer et al, 1991a). The Tn*5280* is flanked by 1142 bp insertion sequences IS*1066* and IS*1067*, which differ in 1bp in the 13bp inverted repeats present on plasmid pP51. A 1068bp open reading frame present in the IS element is having features of a putative transpose (Galas and Chandler, 1989). When Tn*5280* was tagged with a kanamycin-resistance gene it was found to be functional, transposing into the genome of *Pseudomonas putida* KT2442, Tn*5280* (van der Meer et al, 1991b).

1.6.3 Class II transposons

The Class II or Tn*3*-like transposons are characterized by having short (fewer than 50 bp) terminal inverted repeat (IR sequences) and the involvement of a transposase (TnpA) and resolvase (TnpR) in a two-step transposition process that leads to a 5-bp duplication of the target sequence (Sherratt, 1989; Grinsted et al, 1990). Catabolic genes involved in degradation of toluene and naphthalene are shown to be present on class II transposons (Tsuda and Iino, 1987; Tsuda and Iino, 1988; Tsuda and Iino, 1990; Williams et al, 1992).

1.6.4 Toluene degrading transposons (Tn*4651*, Tn*4653*, Tn*4656*)

Catabolic genes for degradation of toluene are reported from *Pseudomonas putida* mt-2. These toluene degrading genes (*xyl* operon) are present on plasmid pWWO and pWW53. On these plasmids the toluene degrading genes are localized on transposons Tn*4651*, Tn*4653* and Tn*4656* (Jacoby et al, 1978; Nakazawa et al, 1978; Chakrabarty et al, 1978). In 110 kb plasmid pWWO, the genes involved in toluene degradation are present on the 56 kb transposon, Tn*4651.* Interestingly, Tn*4651* is

further found part of a 70 kb larger transposon, Tn*4653* (Tsuda and Iino, 1987; 1988 and Tsuda et al, 1989). Similarly, Tn*4652, a* 17 kb derivative of transposon Tn*4651* has been reported to be present on the chromosome of *Pseudomonas putida* PaW85 (Tsuda and Iino, 1987). It appears that the transposons Tn*4651* and Tn*4653* are largely responsible for many transposition events of the *xyl* gene cluster contributing for their existence in chromosomes and plasmids of bacteria (Sinclair et al, 1986; Sinclair and Holloway, 1991; Assinder and Williams 1990; Jahnke et al, 1993). Transposons 4651 and 4653 have the complete set of modules such as res site*, tnpA* and *tnpR* that contribute for their transposition.

Similar to the *xyl* operon of pWWO, the *nah* gene cluster involved in naphthalene catabolism was found to be part of a 38 kb class II transposon, Tn*4655* present on 83 kb plasmid NAH7 from *P. putida* (Yen and Serdar, 1988; Tsuda and Iino, 1990). Tn*4655* is defective in transposition as it has a non-functional transposase and requires a complementing transposase for transposition (Tsuda, 1996). The TnpR of Tn4566 was shown to be a site specific integrase (Tsuda, 1996) and was able to catalyze both integration and resolution reactions (Berg and Howe, 1989; Abremski and Hoess, 1992).

In addition to the aromatic catabolic transposons, there are other catabolic transposable elements that carry genes encoding the metabolism of sugars (Tn*951,* Tn*5276,* Tn*5301*) (Cornelis et al, 1976; Dodd et al, 1990) and citric acid (Tn*3411*)(Ishiguro et al, 1982).

1.7 Organophosphates

Organophosphates (OP), the potent neurotoxic compounds, used as insecticides are the recent additions to the environment. Though OP compounds are made during Second World War, their use as insecticides started only in 1943 (Taylor et al, 2007). Today they are the most prominent pest control agents contributing to nearly 70 % of the total insecticides consumed in the world. (www.agrochemex.net). Organophosphorus compounds are potent inhibitors of serine esterases, of which the critical enzyme is acetylcholinesterase (AChE). Chemically, OP compounds are esters of phosphoric, phosphonic or phosphoramidic acid and their derivatives (Sogorb and Vilanova, 2002). The OP-compounds are also used in industry as plasticizers and petroleum additives. The OP Chemical Warfare Agents (OP CWAs), also known as nerve agents, are a group of extremely toxic compounds and represent major portion of total CWAs stock piled around the world (Karpouzas and Singh, 2006).

1.7.1 OP poisoning

The primary mechanism of action of organophosphate compounds is inhibition of carboxyl ester hydrolases, particularly acetylcholinesterase (AChE). Acetylcholine (Ach) serves as a neurotransmitter at the endings of nerve fibers and hydrolysis of ACh in the synapse results in the decrease in the concentration of ACh and thus terminates cholinergic transmission (Taylor, 2001). AChE is a serine esterase that is involved in the rapid hydrolysis of acetylcholine (Ach) forming choline and acetate (Somani, 1992). OP compounds act as neurotoxins by binding irreversibly to the active site serine residue AchE causing accumulation of acetylcholine at the synapses. Such accumulation of Ach contributes for constant stimulation of post-synaptic receptor leading to loss of muscular

coordination. This would manifest for convulsions and eventually to death by asphyxiation, as control is lost over respiratory muscles (Gunderson et al, 1992).

1.7.2 OP-compounds and environmental pollution

Organophosphorus compounds are used extensively in agriculture as pesticides and in industry as plasticizers, fire-retardants etc. Persistent and indiscriminate use of OP pesticides has resulted in accumulating OP residues in various components of environment (Wang et al, 2007, Ragnarsdottir, 2000). It has been estimated that of the total quantity of pesticides used, less than 1% is utilized by the plants while the remaining 99% goes directly into soil, indirectly into biota and food chain leading to pollution of the environment (Thomson and Abbot 1966; Pimentel and Leviton, 1986). Apart from that, pollution of the environment also arises during manufacture and transport of these organophosphorus compounds (Munnecke, 1979a; Kearney et al., 1987; Copell et al., 1990; Smith et al., 1992). A number of OP poisoning cases have been reported around the world (Bhat and Krishnamachari, 1977; Jeyaratnam, 1990; van der Hoek, 1998; Eddleston and Phillips, 2004). A report presented by Center for Science and Environment (CSE) published in 2007 highlighted presence of OP pesticide residues in the soft drinks sold in Indian sub-continent. Further investigations have revealed usage of contaminated water during the manufacture of soft drinks as main source of OP contamination (cseindia.org). Elsewhere in the world, the studies conducted by Regnery and Püttmann have revealed presence of OP compounds in ground water, rain water and in snow (Regnery and Püttmann, 2009). In fact, poisoning by OP compounds is a global clinical problem as approximately, 3 million poisonings and 300,000 human deaths occur per year due to OP ingestion (Bird et al, 2008).

1.7.3 Microbial Degradation of Organophosphorus compounds

Flavobacterium sp. ATCC 27551 is the first OP compound degrading bacterium isolated from agricultural soils of IRRI, Philippines (Sethunathan et al, 1973). Subsequently, parathion degrading *Brevundimonas diminuta* MG, was isolated from agricultural soils of Texas, USA using parathion as carbon source (Serdar et al, 1982). After these two reports, a number of OP compound degrading bacteria belonging to different taxonomic groups were isolated from diverse geographical regions (Zhongli et al, 2001; Horne et al, 2002; Liu et al, 2005; Karpouzas and Singh, 2006; Singh, 2009). In all these OP degrading bacterial strains, a phosphotriesterase (PTE) is shown to be responsible for the hydrolysis of triester linkage found in structurally diverse groups of OP compounds. Though these triesterases are called with a number of different names based on the substrate used while assaying its activity, currently they are identified with a generic name called organophosphate hydrolase (OPH) (Benning et al, 1994; Cho et al, 2004).

1.7.4 Genetics of Organophosphate degradation

The phosphotriesterase sequences found in bacteria have been divided into three independent structural groups. The homologues of *opd* encoded organophosphate hydrolase (OPH) sequence of *Brevundimonas diminuta*. The second group consists of *mpd* gene coded methyl parathion hydrolase (MPH) mainly isolated from soil bacteria isolated from Chinese agricultural soils (Zhongli et al, 2001, Liu et al, 2005). The third group are *opaA* gene coded organophosphate acid anhydrolases sequences. There is no homology among these three structural groups. They are all kept under

phosphotriesterase super family due to their ability to degrade tri-ester linkage found in structurally diverse group of OP compounds.

1.7.5 The *opd* genes

In *Brevundimonas diminuta* and *Flavobacterium* sp. ATCC 27551, the OPH is coded by an identical organophosphate degrading (*opd*) gene present on large indigenous plasmids, pCMS1 and pPDL2 respectively (Serdar *et al*, 1982; Mulbry and Karns, 1986). The DNA region having *opd* gene has shown to be highly conserved between these two plasmids (Mulbry et al, 1987). The sequence identity extends to 2.6 kb upstream and 1.7 kb downstream of the *opd* gene (Mulbry et al, 1987) (Fig. 1. 4). Beyond this region no detectable homology was found between these two indigenous plasmids (Mulbry et al, 1987; Harper et al, 1988). Further analysis of this conserved sequence has shown transposon like organization for the *opd* gene present in plasmid pPDL2 (Siddavattam et al, 2003). However, no event of transposition event is shown for the transposon-like *opd* gene cluster. Contrary to these findings, the *opd* gene designated as *opdA* found in the chromosome of *Agrobacterium* radiobacter P230 has been shown to be an active chromosome (Horne et al, 2002, Horne et al, 2003).

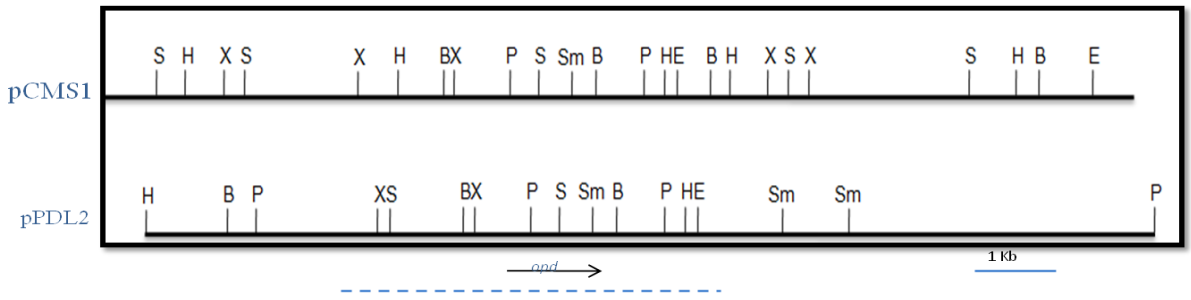


Fig. 1. 4. Physical map of DNA region of *opd* plasmids having identical *opd* genes.

1.7.6 Organization of *mpd* genes

The *mpd* gene encoding methyl parathion hydrolase has been reported from *Plesiomonas* sp. M6 (Zhongli et al, 2001), *Pseudomonas* sp. strain WBC-3 (Liu et al, 2005), and in seven other methyl parathion degrading strains belonging to four different genera (Zhang et al, 2006). In *Pseudomonas* sp. strain WBC-3, the *mpd* gene was present on a 6.5 kb *Kpn*I-*BamH*I fragment. The *mpd* gene was flanked by 2 copies of the insertion sequences (IS*6100*), which is typical character of class I transposon (Fig. 1. 5). The 2.19 kb sequence reported for *mpd* region of *Plesiomonas* sp. M6 shows high homology with the sequence of *Pseudomonas* sp. strain WBC-3. The *mpd* gene is chromosome-borne in *Plesiomonas* sp. M6 whereas in *Pseudomonas* sp. strain WBC-3 it is plasmid-borne (Liu et al, 2005).

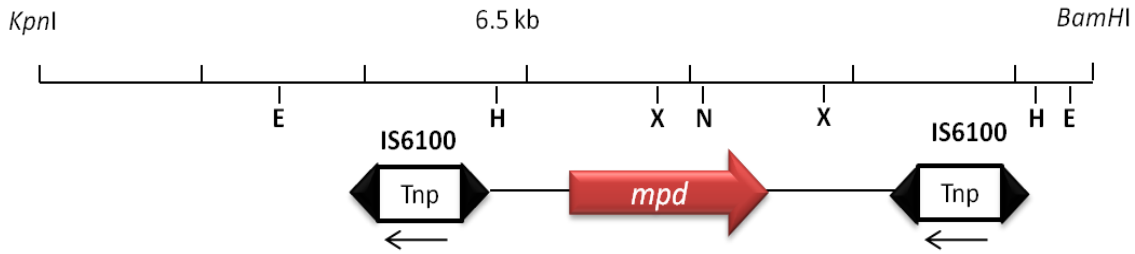


Fig. 1.5. Physical map of transposon carrying *mpd* gene in *Pseudomonas* sp. strain WBC-3.

1.7.7 The *opaA* genes

The organophosphate acid anhydrolase coding *opaA* genes were isolated from *Alteromonas undina* (Cheng et al, 1993). Though initially they were considered to be OP hydrolyzing enzymes further research to identify their function indicated that they are dipepetidases involved in hydrolysis of peptide bond found with a prolyl residue at the carboxy-terminal position (Cheng et al, 1999). Their involvement in OP compound hydrolysis has been shown to be due to structural similarity of OP compounds with their

native substrates. Therefore the *opaAs* have not been taken as part of PTE group while discussing the evolutionary aspects of these enzymes.

1.7.8 HGT of phosphotriesterase (*pte*) coding sequences:

As stated earlier all organophosphate hydrolases are kept under one functional class due to their ability to hydrolyse tri-ester linkage found in structurally diverse group of op compounds. This group of enzymes are kept in a super family designated as phosphotriesterases. The PTEs are found both in prokaryotes and eukaryotes. If discussion is restricted to prokaryotes the *opd* and *mpd* sequences form part of PTE coding sequences. As reported in earlier sections the *mpd* sequences have been shown to be functional transposable elements. There is ample evidence to suggest that HGT contributed for spreading of *mpd* information among soil bacteria. However such information pertaining to *opd* genes are scarce. The event of transposition is seen only for the chromosomally located *opd* element cloned from *Agrobacterium radiobacter* P230 stain isolated from Australian agricultural soils (Horne et al, 2003). However, its chromosome location may not be a favourable situation for horizontal transfer of *opd* gene. Plasmid borne *opd* elements are rather highly suitable for transferring *opd* information through lateral gene transfer mechanisms. Existence of transposon-like *opd* element on mobilizable/self transmissible plasmids adds further strength for its mobility (Siddavattam et al, 2003, Pandeeti et al, 2011). However no studies are yet conducted to assess HGT of *opd* plasmids pPDL2 and pCMS1 isolated from *Flavobacterium* sp ATCCC and *B*. *diminuta*, respectively. The present study is designed with the following objectives to assess the lateral mobility of *opd* plasmids and to evaluate transposition of plasmid borne *opd* elements found in *Flavobacterium* sp. ATCC 27551 and *B*. *diminuta*.

Objectives

1. To determine the complete sequence for the 39.75 kb indigenous plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551.

2. To asses horizontal mobility of *opd* plasmids pPDL2 and pCMS1 found in *Flavobacterium* sp. ATCC 27551 and *B. diminuta* respectively.

3. To determine transposition of *opd* element from plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551.

4. To use of *opd* plasmid, pPDL2 for manipulating *Acinetobacter* sp. DS002 having novel degradative properties.

MATERIALS AND METHODS

Table 2.1. Antibiotics

Name of the antibiotic	**Name of the Supplier**
Ampicillin sodium salt	HIMEDIA
Kanamycin Sulfate	HIMEDIA
Tetracycline hydrochloride	HIMEDIA
Chloramphenicol	HIMEDIA
Streptomycin	HIMEDIA
Gentamycin	HIMEDIA

Table 2.2 Chemicals

Name of the Chemical	**Name of the Supplier**
Acrylamide	SRL
Ammonium persulphate	Sigma Aldrich
Ammonium nitrate	Qualigens
Bovine serum albumin	GE Healthcare Lifesciences, USA
Bromophenol blue	GE Healthcare Lifesciences, USA
Butanol	Qualigens
Calcium chloride	Qualigens
Calcium nitrate	Qualigens
Chloroform	Qualigens
Coomassie Brilliant blue R-250	SRL
Cobalt chloride	SRL
α-cyano-4-hydroxycinnamic acid	Sigma Aldrich
Deoxynucleotide triphosphates	MBI Fermentas
Dipotassium hydrogen orthophosphate	Merck

N,N-Dimethylformamide	SRL
Ethidium bromide	SRL
Ethylene diamine tetraacetic acid (EDTA)	SRL
Ethyl acetate	SRL
Ferrous Sulphate	Qualigens
Glucose	Qualigens
Glycerol	Qualigens
Glycine	SRL
Hydrochloric acid	Qualigens
Isopropanol	SRL
Isopropyl thiogalactopyranoside (IPTG)	SRL
Lysozyme	Bangalore GeneI
Magnesium chloride	SRL
Magnesium sulphate	SRL
Methanol	SRL
N,N'-Methylene bis acrylamide	SRL
β-mercaptoethanol	Sigma Aldrich
Phenol Saturated	Bangalore GeneI
Potassium chloride	Qualigens
Potassium dihydrogen ortho phosphate	Merck
Sodium citrate	SRL
Sodium chloride	SRL
Sodium dodecyl sulphate	SRL
Sodium hydroxide	SRL
Sodium Sulphate	Merck
Sucrose	SRL
RNase A	Banglore GeneI
Tetra ethyl methylene diamine (TEMED)	Sigma
Tris	SRL
Tryptone	Himedia

X-gal	Calbiochem
Yeast extract	Himedia
Zinc chloride	SRL
Trifluoroacetic acid	Sigma Aldrich

Table 3. Restriction and other DNA modifying enzymes

Name of the enzyme	**Name of the Supplier**
*Bam*HI	MBI, Fermentas
*Pst*I	MBI, Fermentas
*Eco*RI	MBI, Fermentas
*Sma*I	MBI, Fermentas
*Xho*I	MBI, Fermentas
*Sal*I	MBI, Fermentas
T_4 DNA Ligase	MBI, Fermentas
Taq DNA polymerase	MBI, Fermentas
Tn*5* Transposase	Epicenter Biotechnologies, USA

Table 2.3 Bacterial strains

Strains	**Genotype or Phenotype**	**Reference or Source**
E. coli DH5α	*supE*44 Δ*lacU*169 (ϕ80 *lacZ* Δ*M*15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*1 *relA*1	Hanahan et al, 1983
E.coli pir-116	*F-* mcr Δ*(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (StrR) nupG pir-116(DHFR)*	Epicentre Biotechnologies, USA
E.coli HB101	F^- mcrB mrr hsdS20(r_B^- m_B^-) recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20(Sm^R) glnV44 λ^-	Boyer & Dussoix, 1969
Flavobacterium sp. ATCC27551	Wild type strain, OPH^+	Sethunathan & Yoshida 1973; Kawahara et al, 2010
Brevundimonas	Wild type strain, Sm^r, Pm^r, OPH^+.	Serdar et al, 1982

diminuta		
Pseudomonas putida KT2440	*hdsMR*, Cm^r	Franklin *et al*. 1981
Acinetobacter sp. DS002	Cm^r, Sm^r, ben^+	This work

Table 2. 4. Plasmids

Plasmids	**Genotype or Phenotype**	**Reference or Source**
pPDL2	OPH^+, indigenous plasmid	Mulbry et al,1986
pCMS1	OPH^+, 66 kb indigenous plasmid	Serder et al, 1982
pCMS1::*tet*	Tc^r, *tet* gene inserted in the *opd* coding region of pCMS1	This work
pCC1FOS	Cm^r, copy control fosmid vector	Epicenter Biotechnologies, USA
pCMSA	Cm^r, 35 kb pCMS1 fragment cloned in fosmid vector pCC1FOS.	This study
pCMSB	Cm^r, 40 kb pCMS1 fragment cloned in fosmid vector pCC1FOS.	This study
pCMSB31	Tn*5*<R6Kγori/KAN-2> inserted into pCMSB	This study
pCMSB32	Tn*5*<R6Kγori/KAN-2> inserted into pCMSB	This study
pCMSB48	Tn*5*<R6Kγori/KAN-2> inserted into pCMSB	This study
pCMB314	Tn*5*<R6Kγori/KAN-2> inserted into pCMSB	This study
pSM2	Ap^r, complete *opd* gene cluster cloned in pUC19	Siddavattam et al, 2003
pTrans::tet	Ap^r, Tc^r, tet gene cloned in the *opd* coding region of pSM2	Siddavattam et al, 2003
pBluescript II KS	Ap^r, $lacZ^+$, cloning vector	Alting-Mees and Short, 1989
pGEMT-Easy	Ap^r, $lacZ^+$, TA cloning vector	Kobs, 1995

pPDL2-Tn*5*<R6Kγori/KAN-2>	Tn*5*<R6Kγori/KAN-2> inserted into indigenous plasmid, pPDL2	This study
pE5	2.2 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS	This study
pE12	2.2 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS	This study
pP33II	12 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS	This study
pP4I	8 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS	This study
pP4II	5kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS	This study
pP3II	5.5 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS	This study
p33EP	8 kb *Eco*RI-*PstI* fragment of pP33II cloned in pBluescript II-KS	This study
p33SP	3 kb *Sal*I-*PstI* fragment of pP33II cloned in pBluescript II-KS	This study
pRK2013	20-kb Km^r ColE1-based restricted-host-range helper plasmid	Ditta et al, 1980
pMMBTnpA	Generated by cloning *tnpA* as *Bgl*II –*Sal*I fragment in the multiple cloning site of a broad host range expression vector pMMB206. Expression of TnpA in pMMBTnpA is driven from an inducible *tac* promoter.	This study
pMMBIstA	Generated by cloning *istA* as *Bam*HI -*Bgl*II fragment in the multiple cloning site of a broad host range expression vector pMMB206. Expression of IstA in pMMBIstA is driven from an inducible *tac* promoter.	This study
pCCFOS912	Cm^r 40 kb genomic DNA fragment of	This study

	Acinetobacter sp. DS002 cloned in fosmid vector pCC1FOS	
pCat	Ap^r, 8kb SalI fragment of pCCFOS912 cloned in pBluescript II-KS	

2.1Preparation of stocks, working solutions and buffers

2.1.1 Ampicillin

Ampicillin stock solution was prepared by dissolving 500 mg of ampicillin in 5 ml of milliQ water and sterilized by filtration. The sterilized solution was stored in 1 ml aliquots at -20°C. When required 100 µl of ampicillin stock solution is added to 100 ml of medium after cooling it to 45°C to get a final concentration of 100 µg/ml.

2.1.2 Chloramphenicol

Chloramphenicol stock solution was prepared by dissolving 300 mg of chloramphenicol in 10 ml of 100% ethanol. The stock solution was stored in 1 ml aliquots at -20°C. When required 100 µl of chloramphenicol stock solution was added to 100 ml of medium after cooling it to 45°C to get a final concentration of 30 µg/ml.

2.1.3 Tetracycline

Tetracycline stock solution was prepared by dissolving 250 mg of tetracycline hydrochloride in 10 ml of 70% (v/v) ethanol/water and sterilized by filtration. The stock solution was stored in 1ml aliquots at -20°C. When required 100 µl of tetracycline stock solution was added to 100 ml of medium after cooling it to 45°C to get the working concentration of 25 µg/ml.

2.1.4 Kanamycin

Stock solution of kanamycin was prepared by dissolving 250 mg of kanamycin sulfate in 10 ml of milliQ water and sterilized by filtration. The stock solution was stored in 1 ml aliquots at -20°C. When required 100 µl of kanamycin stock solution is added to 100 ml of medium after cooling it to 45°C to get the working concentration of 25 µg/ml.

2.1.5 Gentamycin

Stock solution of gentamycin was prepared by dissolving 200 mg of gentamycin in 10 ml of milli Q water and sterilized by filtration. The stock solution was stored in 1 ml aliquots at -20°C after filter sterilization. When required 100 µl of gentamycin stock solution is added to 100 ml of medium after cooling it to 45°C to get the working concentration of 20 µg/ml.

2.1.6 Streptomycin

Stock solution of streptomycin was prepared by dissolving 200 mg of streptomycin in 10 ml of milliQ water. The stock solution was stored in 1 ml aliquots at -20°C after filter sterilization. When required 100 µl of streptomycin stock solution was added to 100 ml of medium after cooling it to 45°C to get the working concentration of 20 µg/ml.

2.1.7 IPTG

1M IPTG stock solution was prepared by dissolving 236.8 mg IPTG in 1 ml of autoclaved milli Q water and stored as aliquots of 100 µl at -20°C. When required the stock solution was thawed on ice bath and 100 µl of stock solution was added to the cooled medium (45°C) to get 1 mM working concentrations of IPTG.

2.1.8 X-Gal

4% of X-gal stock solution was prepared by dissolving 40 mg of X-gal in 1 ml of N, N' -dimethylformamide. When required 100 µl of stock solution of 4% (w/v) X-gal was added to 100 ml of medium after cooling it to 45°C.

2.2 Preparation of buffers and solutions for SDS-PAGE

2.2.1 Acrylamide solution (30%)

100 ml acrylamide solution was prepared by dissolving 30 g acrylamide, 0.8 g N, N-methylene-bis-acrylamide in 70 ml of distilled water. The contents were then filtered and the solution was finally adjusted to 100 ml. The stock solution prepared in this manner is stored at 4^{O}C until further use.

2.2.2 Staking gel buffer for SDS-PAGE

3.93 g of Tris was dissolved in 50 ml of double distilled water and the pH of the solution was adjusted to 6.8 using 1N HCl and finally the volume of the buffer was made up to 100 ml with distilled water.

2.2.3 Running gel buffer for SDS-PAGE

59.93g of Tris was dissolved in 400ml of distilled water and pH of the solution was adjusted to 8.8 using 1N HCl. Finally the volume of the buffer solution was made up to 500 ml using distilled water.

2.2.4 Tank buffer (pH 8.5) for SDS-PAGE

Tank buffer (1X) concentration was made by 3.03 g of Tris, 14.4 g of glycine and 1 g of SDS was dissolved in 500 ml of distilled water. Finally, the volume of the buffer was made up to 1000 ml. The buffer at 1X concentration contains 0.025M Tris, 0.192M glycine and 0.1% SDS.

2.2.5 Sample loading buffer for SDS-PAGE

1.2 ml of 0.5 M Tris (pH 8.0), 2 ml of 10% SDS, 1 ml of 10% glycerol, 0.5 ml of β-mercaptoethanol, 0.001 g of bromophenol blue were taken in a 10 ml reagent bottle. The contents were mixed well before the volume was made up to 10 ml and solution was stored at4°C until further use. When necessary adequate amounts of loading buffer was used to prepare protein samples.

2.2.6 Staining solution

0.2 g of coomassie brilliant blue was dissolved in 30 ml of methanol. To this 10 ml of acetic acid was added and finally the volume was made up to 100 ml using distilled water. The contents were stored at room temperature in amber color bottle until further use.

2.2.7 Destaining solution

30 ml of methanol is mixed with 10 ml of glacial acetic acid before making up the volume to 100 ml using distilled water. The contents were freshly made and used once to destain the SDS-PAGE gels.

2.2.8 Protein Markers

Low molecular weight protein markers supplied by MBI Fermentas were used. Size of the protein markers include, 116.0, 66.2, 45.0, 35.0, 25.0, 18.4 and 14.4 kDa.

2.3 Preparation of buffers for Western blotting

2.3.1 Protein transfer buffer

3.03 g of Trizma-base and 14.4 g of glycine was dissolved in 650 ml of distilled water. To this 200 ml of methanol was added and final volume was made up to 1000 ml with distilled water. The buffer was stored at 4oC until further use.

2.3.2 TBS-T Buffer

2.42 g of Tris and 8 g of sodium chloride were dissolved in 500ml of distilled water. The pH of the solution was adjusted to 7.6 using 1N HCl before adding 1ml of 0.1% Tween-20. Finally the volume of the buffer solution was made up to 1000 ml with milli Q water.

2.3.3 Blocking Buffer

Blocking buffer was prepared by adding 10 g of skimmed milk powder to 10 ml of 1X Tris buffered saline with 0.1% Tween-20 and the contents were stirred well. The solution was prepared fresh before blocking the membrane.

2.4 Preparation of buffers for Isoelectro focusing (IEF)

2.4.1 Urea rehydration stock solution

Urea rehydration stock solution was prepared by dissolving 12 g of urea, 0.5 g of CHAPS, 500µl of Pharmalyte (pH 3-10), 50 µl of 1% bromophenol blue solution and the final

volume was made up to 25 ml with miliQ water. Aliquots of 2.5 ml were made and stored at -20°C until further use. When necessary 7 mg of DTT was added to the 2.5 ml aliquot and used immediately.

2.4.2 SDS equilibration buffer solution

72.1 g of urea, 10 ml of Tris-Cl 1.5M (pH 8.8), 69 ml of glycerol, 4 g of SDS and 4 ml of 1% bromophenol blue stock solution were added and the final volume was made upto 200 ml with milliQ water. Equilibration solution-I was made by adding 100 mg of DTT to 10 ml of the stock solution. Similarly, equilibration solution II was made by adding 250 mg of iodoacetamide to 10 ml of stock solution.

2.5 Preparation of solutions for Agarose gel electrophoresis

2.5.1 Tris Borate EDTA (TBE) buffer

A stock solution of 10 X TBE buffer was prepared by adding 108 g of Tris, 55 g of boric acid and 40 ml of 0.5 M EDTA (pH 8.0) to 900 ml of distilled water and dissolved properly before adjusting the volume of the contents to1000 ml with distilled water. The buffer was then stored in a container fixed with a tap. When necessary adequate volume of the stock buffer was diluted to get 1X TBE with distilled water and used for preparing agarose gels.

2.5.2 TE buffer

TE buffer was prepared by dissolving 121 mg of Tris and 37.2 mg of EDTA (pH 8.0) in 80 ml of distilled water. The pH of the buffer was adjusted to 8.0. Finally the volume was

made up to 100 ml with distilled water to get the concentration of 10 mM Tris and 1 mM EDTA.

2.5.3 Sample Loading buffer (6X) for Agarose gel electrophoresis

5 mg of bromophenol blue, 25 mg of xylene cyanol and 4 g of sucrose was dissolved in few ml of distilled water and finally volume was made up to 10 ml and stored at 4°C.

2.6 Substrates for growth / enzyme assays

2.6.1 Benzoate (1M)

1M benzoate stock was prepared by dissolving 12.1 g of benzoate in 100 ml of double distilled water and the solution was filter sterilized using 0.2µm sartorius filter and stored at 4^{o} C. Whenever needed adequate amount of stock was added to minimal media to get a final concentration of benzoate to 5-50mM in the culture medium. While performing oxygraph studies appropriate amounts of benzoate was used to get a final concentration of 50µM in the reaction mixture.

2.6.2 Sodium succinate (1M)

1M sodium succinate stock solution was prepared by dissolving 27 g of sodium succinate in 100 ml of distilled water and the solution was filter sterilized using 0.2µm sartorius filter and stored at 4^{o} C. Whenever needed adequate amount of stock was added to minimal media to get a final concentration of 10 mM.

2.6.3 Catechol (50mM)

50 mM catechol stock was prepared by dissolving 110 mg of catechol in 20 ml of double distilled water and the solution filter was sterilized using 0.2µm sartorius filter and

stored at 4^{o} C. While performing oxygraph studies appropriate amounts of catechol was used to get a final concentration of 50µM in the reaction mixture.

2.6.4 Methyl parathion (MP) (*O,O*-dimethyl-*O*-paranitrophenyl phosphorothioate)

0.2 M stock solution of methyl parathion was prepared by dissolving 263.2 mg of methyl parathion in 5 ml methanol and was stored at -20^{o}C until further use. While performe enzyme assay appropriate amount of methyl parathion was used to get a final concentration of 200 µM in the reaction mixture.

2.6.5 Paraoxon (Diethyl *p*-nitrophenyl phosphate)

14.5 mM Paraxon stock was prepared by dissolving 200mg of paraoxon was dissolved in 50 ml of ddH_2O. While performing enzyme assays appropriate amount of paraxon was used to get a final concentration of 2mM in the reaction mixture.

2.6.6 *p*-nitrophenol (50 mM)

50 mM *p*-nitrophenol stock was prepared by dissolving 347.7 mg of *p*-nitrophenol in 50 ml of double distilled water and the solution was filter sterilized using 0.2µm Sartorius filter and stored at 4^{o} C. While performing oxygraph studies appropriate amounts of *p*-nitrophenol was used to get a final concentration of 50µM in the reaction mixture.

2.6.7 4-Nitrocatechol (50 mM)

50 mM 4-nitrocatechol stock was prepared by dissolving 155 mg of 4-Nitrocatechol in 20 ml of double distilled water and the solution was filter sterilized using 0.2µm Sartorius filter and stored at 4^{o} C. While performing oxygraph studies appropriate amounts of 4-nitrocatechol was used to get a final concentration of 50µM in the reaction mixture.

2.6.8 1, 2,4-Benzenetriol(50 mM)

50 mM Benzenetriol stock was prepared by dissolving 126 mg of Benzenetriol in 20 ml of double distilled water (pH 2.0) and the solution was filter sterilized using 0.2µm sartorius filter and stored at 4^{o} C. While performing oxygraph studies appropriate amounts of 1, 2, 4-benzenetriol was used to get a final concentration of 50µM in the reaction mixture.

2.7 Media

The following media were used to propagate the bacteria. All media were sterilized by autoclaving for 20 min at 15 lb/Sq. When required appropriate amounts of antibiotics of Ampicillin (100 µg/ml), Tetracycline (25 µg/ ml), Chloramphenicol (30 µg/ ml), Streptomycin (20 µg/ml), Spectinomycin (20 µg/ml) Gentamycin (20 µg/ml) and Kanamycin (25 µg/ml) were added after cooling the media to 45°C.

2.7.1 Modified Wakimoto Medium

Modified Wakimoto medium was prepared by dissolving 0.5 g of $Ca(NO_3)_2 . 4H_2O$, 2.0 g of $Na_2HPO_4.12H_2O$, 5.0 g of peptone, 15.0 g of Sucrose,0.5 g $FeSO_4 . 7H_2O$, 15.0 g of agar in 1 liter of distilled water and sterilized by autoclaving at 121^{o}C for 15 min.

2.7.2 Luria Bertani (LB) broth

The LB medium was prepared by dissolving 10 g of peptone, 5 g of yeast extract and 10 g of NaCl in 500 ml of distilled water. The contents were stirred and finally made up to 1 liter with milli Q water. The pH of the medium was adjusted to 7.0 with 1N NaOH and then sterilized it by autoclaving. The LB agar plates were prepared by adding 2% agar to LB broth.

2.7.3 Minimal Medium

The minimal medium was prepared by dissolving 4.8 g of K_2HPO_4, 1.2 g of KH_2PO_4 and 1 g of NH_4NO_3 in a few ml of milli Q water and finally the volume was made up to 978.5 ml. Then the salt solution was sterilized by autoclaving for 15 minutes at 15 lb pressure. $MgSO_4.7H_2O$ (1 g/10 ml), $Fe_2(SO_4)_3$ (1 g/100 ml) and $Ca(NO_3)_2.4H_2O$ (4 g/20 ml) were prepared separately and sterilized by autoclaving as described above. 2 ml of $MgSO_4.7H_2O$, 100 µl of Ca $(NO_3)_2.4H_2O$ and 50 µl of Fe_2 $(SO_4)_3$ were added to sterile 1 litre of minimal medium under aseptic conditions. When required filter sterilized benzoate (5 mM or 50 mM) or succinate (10 mM) were added to minimal medium as growth substrates.

2.8 Isolation of plasmids by alkaline lysis method

Mini preparations of plasmid DNA was carried out by the following procedures of Birnboim and Doly (1979), and Ish-Horowicz and Buker (1981) with slight modifications. A single bacterial colony carrying plasmid to be isolated was inoculated into 3 mL of LB medium containing appropriate antibiotic and was incubated overnight at 37 °C with vigorous shaking. 1 ml of overnight culture was centrifuged at 13400 rpm for 1 min and supernatant was discarded. Bacterial pellet was resuspended in 100 µL of ice-cold solution I (50 mM glucose, 25 mM Tris-Cl, pH 8.0; 10 mM EDTA, pH 8.0) by vigorous vortexing. To the above bacterial suspension 200 µL of freshly prepared solution II (0.2 N NaOH, 1% SDS) was added and the contents were mixed by inverting the tube 5-6 times. Then 150 µL of ice-cold solution III (3M potassium acetate, pH 4.8) was added to the above bacterial lysate and mixed by inverting the tube 4-5 times. Then tube was kept on ice for 3-5 minutes. Precipitate formed in the above mixture was centrifuged by spinning the contents at 13400 rpm for 10 minutes.Then the supernatant was

transferred into a fresh tube and equal volumes of phenol: chloroform mixture was added. The contents were mixed by vortexing and followed by centrifugation at 13400 rpm for 5 minutes. Aqueous phase was transferred to a fresh tube and 1/10th volume of 3 M sodium acetate (pH 4.8) and 2 volumes of ethanol were added and tubes were kept at –20°C for 30 minutes. Then the tubes were centrifuged at 13400 rpm for 20 min at 4°C in a microfuge to pellet down the plasmid DNA. The DNA pellet was further washed with 70% ethanol to remove traces of salts associated with plasmid. Subsequently, the plasmid DNA was dried before dissolving it in 50 µl of TE (pH 8.0) and stored at -20 °C until further use.

2.9 Purification of plasmids using QIAgen Mini preparation kit method

Plasmids were purified using QIAgen mini preparation kit especially when used for cloning and sequencing reactions. A single bacterial colony carrying plasmid was inoculated into 3 ml of LB broth containing appropriate antibiotic and was incubated overnight at 37 °C with vigorous shaking (~150 rpm). The overnight culture was centrifuged at 13000 rpm for 1 min and supernatant was discarded. The bacterial cell pellet was resuspended in 250 µl of buffer P1 and were lysed by adding 250 µl of buffer P2 prior to mixing of the tubes by inverting them 4-6 times. After lysis of the cells the contents were neutralized by adding 350 µl of buffer N3 prior to mixing of the contents by inverting immediately for 4-6 times. Then tubes were centrifuged at 13000 rpm for 10 minutes to pellet down the cell debris. After centrifugation the supernatant was directly transferred to a QIAprep column placed in a collecting tube. The entire assembly was placed in a microfuge and the supernantant was allowed to pass through the column for a minute by centrifuging at 13000 rpm. The column was then washed with 750 µl of buffer PE. To remove the residual wash buffer PE, column

placed in the collection tube was centrifuged at 13000 rpm for 1 minute. Finally, plasmid DNA was eluted from the column by adding 50 µl buffer EB (10 mM Tris-HCl, pH 8.5) or H_2O to the centre of QIA preparation column followed by brief centrifugation at 13000 rpm. The plasmid DNA was stored at –20°C until further use.

2.10 Agarose gel electrophoresis

Agarose gel electrophoresis was performed by the following standard procedures described by Sambrook *et al*., 1989. Required amount of agarose was dissolved in TBE by heating in micro oven. The solution was then cooled to 50 - 55°C and poured on a clean sterile gel tray fitted with combs. The gel was allowed to solidify at room temperature. After solidification the gel along with the gel tray was immersed in the TBE buffer poured in the electrophoretic tank. Buffer level was adjusted to cover the gel to a depth of about 1 mm. Appropriate amount of DNA was mixed with 4 µl of 10 X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 40% sucrose) and loaded into the wells of submerged gel. Electrophoresis was carried out at 100 volts till the bromophenol blue reaches the end of the gel. The gel was then stained (0.5 µg/ml ethidium bromide in water) destaining solution (distilled water) each step for 15 minutes before visualizing under UV transilluminator. The electrophoretic mobility of DNA was recorded by taking the photograph of the gel using gel documentation system. The size of the unknown DNA fragments were determined by comparing their mobility pattern with that of the molecular size markers obtained from commercial companies.

2.11 Southern blotting

Colony Hybridization:

E. coli colonies containing recombinant fosmids (*Acinetobacter* sp. DS002 fosmid library) was grown in LB media supplemented with chloramphenicol (30 µg/ml). Colony hybridization was performed by following the procedure described elsewhere (Sambrook et al. 1989). Colonies of the library were allowed to grown to a size of 1.5 mm and the plates were transferred to a cold room. A dry nitrocellulose filter paper was placed on the LB plate having the colonies so that it makes contact with all the colonies. The filter paper was removed and the colonies on the filter paper were lysed by placing in a denaturation solution (1.5M NaCl, 0.5M NaOH) for 5 min. Neutralization was done for 30 min with 0.5M Tris-HCl buffer (pH 7.5) having 0.5M NaCl and 2X SSPE (0.3 M NaCl, 0.02M NaH2PO4, 0.002M EDTA). After transfer the membrane was UV cross-linked and treated with pre-hybridization buffer (0.5M phosphate buffer pH 7.2, 7% SDS and 1mM EDTA pH 8.0) for 4 hours at 65 ^{o}C. The probe was hybridized for 16-20 hours in hybridization buffer (0.5M phosphate buffer pH 7.2, 7% SDS and 1mM EDTA pH 8.0). After hybridization the membrane was washed with wash buffer for 4 times. Excess wash buffer was removed with a pad of paper towels and the damp membrane was sealed in a polythene bag. The membrane was exposed to X-ray film for 16 hrs at -80 ^{o}C to obtain the autoradiographic image. After exposure, the film was kept in a dark room and allowed to warm up to room temperature. The X-ray film was developed by immersing it in developer for 5min, washed with water and fixed with a fixer for 5 min.

2.12 SDS-polyacrylamide gel electrophoresis

The protein samples were separated for SDS-Polyacrylamide gel electrophoresis

following the procedures of Laemmli (1970). 12.5% running acrylamide gel solution was prepared by mixing 3.6 ml 30% acrylamide solution, 4.5 ml running buffer (0.98 M Tris-Cl, pH, 8.8), 0.9 ml distilled water, 90 µl of 10% SDS, 4.5 µl of TEMED and 45 µl of freshly prepared 10% ammonium persulphate. The solution was mixed well and degassed before adding ammonium persulphate. The contents were then poured in between two sealed glass plates containing 1.5 mm spacers to form a slab. These contents were over-layered with 0.1 ml water saturated n-butanol and allowed to polymerize for 40 minutes at room temperature. After polymerization of running gel, butanol was removed by washing with water and the traces of water were removed by wiping with filter paper strips. The 7.5 % stacking acrylamide gel solution was prepared by taking 1.25 ml of 30 % acrylamide solution, 2.5 ml of stacking gel buffer (0.325 M Tris-Cl, pH, 6.8), 1.25ml of water, 50 µl of 10% SDS, 5 µl of TEMED and 50 µl ammonium persulphate. The stacking gel solution was carefully over layered on the running gel. Immediately a comb of required size was placed to form the wells. The stacking gel was allowed to polymerize for 30 min. After 30 min the comb was removed and the wells were washed with tank buffer (50 mM Tris-Cl, 0.384 M glycine, 0.1% SDS, pH 8.5). Protein samples were mixed with equal volumes of sample loading buffer (50 mM Tris HCl, pH 6.8, 2% SDS, 0.1% bromophenol blue, 10% glycerol, 100 mM 2-mercaptoethanol) and kept in a boiling water bath for 5 min. The contents were then briefly centrifuged before loading the sample into the wells. The electrophoresis was carried at 100 volts till the tracking dye reached the anode end of the gel. The gel was removed from the glass plates and the protein bands were stained with staining solution containing 0.2% W/V coomassie brilliant blue in methanol : water : acetic acid (30 : 60 : 10) for 6-8 h. The gel was then destained in the destaining solution containing methanol : water : aceticacid (30 : 60 : 10) for another 6-8 h.

2.13 Western blotting

Western blotting was performed using ECLPlus western blotting detection system supplied by Amersham Pharmacia Biotech, USA. Left-hand corner on the top of the gel was cut to mark the orientation of the gel. A thin sheet of Whatman filter paper soaked in towbin buffer (0.3M Tris-Cl pH 10.4; 20% methanol) and placed on anode graphite plate. Polyvinylidene fluoride (PVDF) membrane (Highbond-P) was kept in methanol for 5 seconds to prewet the membrane and immediately rinsed with towbin buffer and placed on top of the filter papers. Then the SDS-PAGE gel submerged in towbin buffer was placed on top of the membrane. On top of the gel, a thick sheet of filter paper soaked in towbin buffer was placed. Transfer of proteins onto PVDF membrane was carried out at constant voltage of 15V for 45 min.

Once the transfer process was finished, membrane was stained with ponceau reagent to check the transfer of proteins on to the PVDF membrane. The Ponceau was removed by rinsing the membrane with 1X TBS-T buffer. Then the membrane was blocked with blocking reagent such as 10% skimmed milk powder in 1X TBS-T buffer (20 mM Tris-base; 137 mM NaCl; pH 7.6). Blocking was continued for 1 hr at room temperature with constant shaking. Then the membrane was washed three times for 15 min in large volumes of 1X TBS-T buffer. Primary antibody (mouse IgG) raised against OPH antibodies were diluted in a ratio of 1:5000 in 1X TBS-T buffer containing 10% blocking agent and the membrane was incubated at room temperature with constant shaking for 3 hrs. After incubation the membrane was washed with 1X TBS-T three times for 15 min each wash to remove unbound primary antibody. After primary antibody incubation, the membrane was incubated for 45 min at room temperature with a secondary antibody (anti-mouse IgG supplied by ECL+Plus kit, Amersham Pharmacia Biotech, UK) diluted in a ratio of 1:5000 in

1X TBS-T buffer containing 10% blocking reagent. Then the membrane was washed three times with 1X TBS-T as mentioned above and protein signals were detected by following the manufacturer's instructions.

2.13.1 Detection

Detection was performed using ECL+Plus kit procured from Amersham Pharmacia Biotech, UK by taking Solution A and B in a ratio of 1 : 40 respectively and mixed them gently in dark room. Membrane was then incubated in the reagent mix for exactly one min and dried with tissue papers to remove excess detection reagent before wrapping in a clean cling film. Then the membrane was exposed to Biomax X-ray film for 30 seconds. The film was developed by immersing it in a Kodak developer solution and fixed by transferring it to Kodak - fixer solution for 5 min. The film was then washed in water and dried before analyzing the results.

2.14 DNA ligation

Concentration of vector and insert were estimated and they were taken in a ratio of 1:3 (vector : insert) in a sterile eppendorf tube. One unit of T_4 DNA ligase and 2µl of 10X ligation buffer were added and the reaction volume was adjusted to 20µl. Ligation reaction was incubated at 16 ^{o}C for 16 hours. A portion of ligation mixture was taken to transform *E. coli* competent cells.

2.15 Preparation of competent cells

Competent cells were prepared following the procedures of Mandel and Higa, (1970) with slight modifications. *E. coli* cells required to be made competent were grown in LB broth of 100 ml at 37°C with an initial inoculum of 1%. The cultures were allowed to grow till the cell density reached to 0.3-0.4 OD at 600 nm. The culture was chilled on ice for 30

minutes and centrifuged at 6000 rpm for 10 minutes to harvest the cells. The cell pellet was then suspended in 50 ml of ice cold 0.1 M $CaCl_2$ and incubated on ice for 30 minutes. After incubation the cells were collected by centrifugation and gently resuspended in 10 ml of ice cold 100mM $CaCl_2$ containing 15% glycerol and stored in aliquots of 200 µl at -70°C untill further use.

2.16 Transformation

The frozen competent cells were thawed by placing them on ice bath. The ligation mixture/plasmid of interest was added and incubated on ice for 30 minutes. After 30 minutes, the cells were subjected to heat shock at 42 °C for exactly 90 sec and immediately chilled on ice for 2 min. Further, 800 µl of LB broth was added and incubated at 37 °C for 45 min. The cells were collected by centrifugation and resuspended in 200 µl of LB broth before plating. When needed 60µl of 2% X-gal or 60 µl of 2% IPTG was added and plated on LB agar plates containing appropriate antibiotic. The plates were then incubated at 37°C for 12 hr for colonies to appear.

2.17 Preparation of electro-competent cells

Electro-competent cells were prepared following the procedures of Sharma and Schimke, 1996. The *E. coli* / *Acinetobacter* sp. DS002 were inoculated with 1% of overnight cultures and grown in LB broth of 1000 ml at 37°C and 30°C respectively. The culture was then allowed to grow till the cell density reached to 0.3-0.4 OD at 600 nm. The culture was then chilled on ice for 30 minutes and centrifuged at 4000 rpm for 10 minutes to harvest the cells. The cell pellet was then resuspended in 50 ml of ice cold, 10% glycerol and was centrifuged to collect the cells. The process was repeated thrice to wash the cells. Finally,

the cells were resuspended in 10 ml of sterile 10% glycerol and distributed into 100 µl aliquots before flash freezing them by dipping them in liquid nitrogen. The electro-competent cells prepared in this manner were then stored in -80 ^{o}C until further use.

2.18 Electroporation

Electro-competent cells were thawed on ice before distributing them into two equal portions (50 µl) using prechilled 1.5 ml tubes. About 1 to 2 µg of DNA was added to one of the tubes and the second tube containing 50 µl of competent cells served as control. The suspension was mixed vigorously by flicking the tube. The cell/DNA mixture was placed in a pre-chilled cuvette, between the electrodes and a pulse of 2.5 kV was applied for 4.5 sec using Genepulser (BioRad, USA). Following the pulse, immediately 1 ml of SOC medium was added to the cells (SOC: 2% Bacto tryptone, 0.5% Bacto yeast extract, 1 mM NaCI, 2.5mM KCI, 1 mMMgCI2 10 mM MgSO4 20 mM glucose). The cells along with broth were taken in a sterile 1.5 ml tube and the samples were incubated, with shaking at 250 rpm, for 1 hour at 37^{o}C/ 30 C. After the incubation, the cells were diluted using SOC and appropriate dilution was plated on LB agar plates containing suitable antibiotics.

2.19 Isolation of plasmid pPDL2 from *Flavobacterium* sp. ATCC 27551

Large indigenous plasmids of *Flavobacterium* sp. ATCC 27551 were isolated by following the Currier Nester protocol with the following modifications (Currier and Nester, 1976). *Flavobacterium* sp. ATCC 27551 were grown at 30 ^{o}C on modified Wakimoto plates for 2 days. The cells were scrapped from the fresh plates using sterile 1 ml tips and a suspension of bacterial cells was made in 5 mL 10 mM Tris (pH 8.0) such that the cell density is A_{425nm} = 0.75-0.8. The suspension was centrifuged at 10000 rpm for 10 min at

4 ^{o}C and the supernatant was discarded. Bacterial pellet was then resuspended in 5.4 mL of TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) containing lysozyme (10 mg/mL). The contents were thoroughly mixed before incubating in a water bath maintained at 37 ^{o}C for 15 min. Further, 300 µL of 20% SDS prepared in TE buffer was added and the tubes were mixed gently by inversion. The lysate was vortexed at maximum speed for 2 min to shear chromosomal DNA. The sheared DNA was denatured by mixing the contents after adding 200 µL of 3 N NaOH. The samples were then incubated at room temperature for 15 min with occasional shaking. After incubation, the reaction mixture was neutralized by adding 1 mL of 2 M Tris (pH 7.0) with constant mixing by inverting the tube. Subsequently, 1.5 mL of 5 M NaCl was added to the contents and thoroughly mixed before extracting it with equal volume of salt saturated phenol. A clear top phase obtained after extraction was collected in a fresh tube and the plasmid DNA was precipitated by adding 2 volumes of ice cold ethanol and incubating at -80 ^{o}C for 30 min. Then the tubes were centrifuged at 8000 rpm for 10 min at 10 ^{o}C to pellet down the precipitated plasmid DNA. The DNA pellet was then washed with 70% ethanol to remove traces of salts. Subsequently, the plasmid DNA was air-dried before dissolving it in 50 µl of TE (pH 8.0). The plasmids isolated in this manner were stored at -20°C until further use.

2.20 Rescue of pPDL2 from *Flavobacterium* sp. ATCC 27551

The plasmid preparation made from *Flavobacterium* sp. ATCC 27551 was directly used to rescue pPLD2 from other plasmids by using the transposon EZ-Tn*5*™<R6Kγori/KAN-2>. The strategy of rescue cloning is shown in Figure 2. 1. In a reaction mixture containing 1 µl of 10 X buffer, 1U of transposase, the plasmid pPDL2 and mini-transposon were taken in

equi-molar concentration before incubating the contents for 2 h at 37 ^{o}C. After incubation

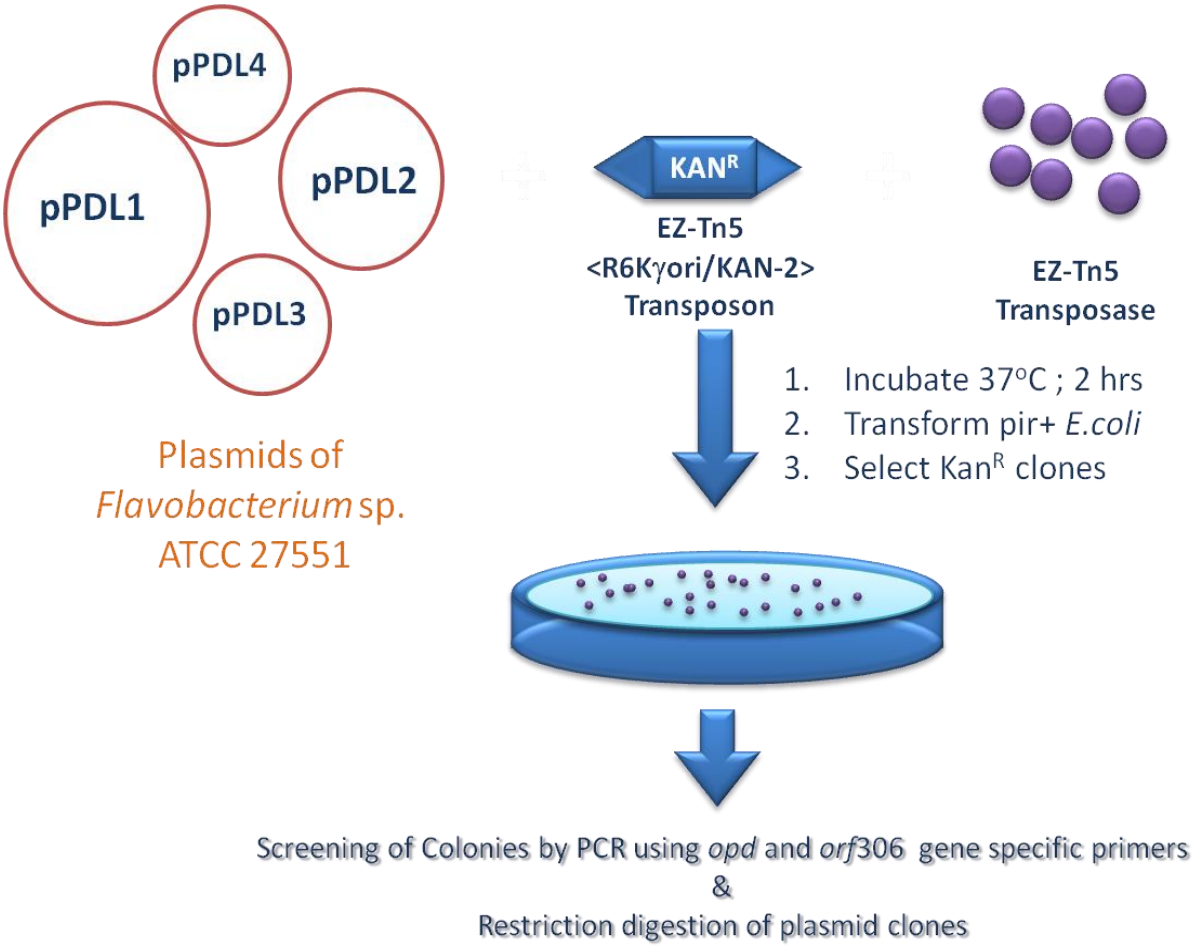


Fig. 2. 1. Strategy for rescue cloning of pPDL2 from *Flavobacterium* sp. ATCC 27551

the transposase was inactivated by adding 1 µl of stop solution followed by incubation at 70 ^{o}C for 10 min. About 1 µl of transposition mixture was then electroporated into *E.coli* EC100D *pir*-116 cells at 2.5 kV, 200 Ω for 4.5 sec using a GenePulser (Bio-Rad Laboratories, USA). After electroporation, 1 ml of SOC medium was added and the cells were immediately incubated for 1 h at 37 ^{o}C with vigorous shaking. About 100 µl of electroporated cells were taken to spread on LB plates supplemented with kanamycin prior to incubation for 12 hours at 37 ^{o}C. Colonies having plasmid pPDL2 were identified by performing colony PCR using *opd* specific primers.

2.21 Sub-cloning of pPDL2-Tn™<R6Kγori/KAN-2>

The overnight cultures of E.*coli-pir116* cells containing pPDL2- Tn*5*™<R6Kγori/KAN-2> were grown in LB medium having kanamycin and the copy number of the pPDL2::Tn5<R6Kγori/KAN-2> was induced by adding appropriate amounts of copy number induction solution supplied by the manufacturer. The plasmid pPDL2::Tn*5*™<R6Kγori/KAN-2> was isolated from E.*coli-pir116* cells using the BAC isolation protocol (Sambrook et al., 1989). The restriction profile of pPDL2- Tn*5*™<R6Kγori/KAN-2> was created by using the following restriction enzymes *Sma*I, *BamH*I, *Pst*I, *Sal*I, *Xho*I, *Hind*III and *EcoR*I. Further, all fragments of pPDL2- Tn*5*<R6Kγori/KAN-2> obtained after digestion with *Pst*I and *Eco*RI were ligated independently into pBluescript-KS digested with similar enzymes.

2.22 Sequencing of plasmid pPDL2

Recombinant plasmids having *Pst*I and *Eco*RI fragments of pPDL2-Tn*5*<R6Kγori/KAN-2> were isolated using QIAgen mini-prep columns and were used as templates. BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, USA) was used for sequencing reactions and an ABI PRISM® 3100 Genetic Analyzer was used for generation of the DNA sequence. Initially, sequencing was done by using M13/pUC18 universal forward and reverse primers taking the recombinant plasmids as templates, later primer walking strategy was employed for generation of the remaining sequence. External primers were used on the pPDL2-Tn*5*<R6Kγori/KAN-2> to get the sequence overlaps for two adjacent fragments. The large *Pst*I fragment of pPDL2-Tn*5*<R6Kγori/KAN-2> was sub-cloned into pBluescript as *Pst*I-*Sal*I and *Pst*I-*EcoR*I fragments and these sub-clones were used as templates for sequencing.

2.23 Sequence Assembly

All sequences were viewed and edited to remove vector sequences by using Chromas 2.13 (www.technelysium.com.au/chromas) software. Sequences were assembled into contigs by using the program ContigExpress of VectorNTI software (Invitrogen Technologies, USA).

2.24 Annotation of pPDL2 sequence

The sequence of plasmid pPDL2 was annotated using Artemis sequence annotation tool (http://www.sanger.ac.uk/resources/software/artemis/) (Rutherford et al, 2000). ORFs (open reading frames) were identified by using the inbuilt tool of the Artemis software and the start codon in the predicted ORFs was fixed with the help of BLAST searches. BLAST searches were made against the non-redundant database of NCBI using BLASTx programme of NCBI (www.ncbi.nlm.nih.gov/BLAST). IS elements, transposons and their repeat elements were identified by doing a pairwise alignment using BLASTn programme against the ISfinder database (http://www-is.biotoul.fr/).

2.25 Prediction of *oriV*

OriV is predicted based on sequence homology to other annotated or predicted *oriV*s and also by performing GC Skew analysis. GC skew was calculated according to the method described by Grigoriev (1998). For the purpose of calculating GC skew the total sequence of pPDL2 is divided into sub-sequences of equal lengths and GC Skew in a sub-sequence is calculated as the ratio of (G - C) / (G + C), where G and C is the number of G's and C's present respectively in a sub-sequence of pPDL2. The GC-skew is multiplied by w/c in order to its dependence on sub-sequence length where w and c are lengths of the sub-sequence

and total sequence of pPDL2 respectively. The maximum and minimum value of GC skew are usually associated with termination and origin of replication.

2.26 Prediction of *att* sites

While predicting *att* sites two independent approaches were used. In the first approach sequences of all available plasmids having identical integrases were collected from NCBI database. From bacteria having these plasmids, a dataset of all tRNA sequences along with their upstream and downstream sequences was created. Pairwise alignments were made between the plasmids and the tRNA sequences to identity the *att* sites.

In a second approach, the nucleotide sequences of hosts having the integrase genes were collected from NCBI database. Using these plasmid and genomic sequences, Genomic Islands (GIs) were predicted using Islandviewer software (http://www.pathogenomics.sfu.ca/islandviewer/query.php). The predicted GI sequences were then used to make pairwise alignments with sequences of tRNA of *Sphingomonas* and *Sphingobium* to identify short exact repeats flanking the GIs. Repeats present on the plasmids and genomes having low E-value were considered as the *att*P and *att*B sites respectively. The *att* sites predicted using these datasets were then identified on the plasmid pPDL2 using BlastN program. Alternatively, tRNA sequences were directly compared with pPLD2 sequence using BlastN program to find short sequences of tRNA on the plasmid.

2.27 Promoter and terminator predictions

Promoters were predicted using BPROM (http://www.softberry.com), Promoter Prediction by Neural Network and PROMSCAN (http://www-

bimas.cit.nih.gov/molbio/proscan/) softwares. Prediction of terminators was done with TransTermHP (Kingsford, 2007).

2.28 Horizontal transfer of *opd* plasmids

2.28.1 Mobilization of pPDL2::Tn5<R6Kγori-Kan2>

2.28.1.1Biparental mating

Biparental mating of *E.coli* pir-116 harbouring pPDL2::Tn5<R6Kγori-Kan2> (donor) and *Acinetobacter* sp. DS002 (recipient) was performed by following protocols described by Figurski and Helinski (1979). Cultures were grown to log phase in LB media supplemented with kanamycin and chloramphenicol respectively. Cultures grown to mid log phase were collected to harvest the cells. The cells were then thoroughly washed with citrate saline buffer before equalizing the OD of the culture. The donor and recipient cells were then mixed in 3:1 ratio and the cells were collected by centrifugation at 6000 rpm for 5 min at 4 ^{o}C. The cell pellet was then resuspended in minimal amount of citrate saline buffer and plated on LB plates having no antibiotic. The plates were then incubated for 12 h at 30 ^{o}C. The colonies were scraped from the LB plate and resuspended in 1 ml of citrate saline buffer. The cells were serially diluted and 100 µl of each dilution were independently plated on selection plates having of minimal media supplemented with 5 mM benzoate and kanamycin. Donor and recipient cells treated in similar manner were also plated on the MM plates supplemented with 5 mM benzoate and kanamycin.

2.28.1.2 Triparental mating

Triparental mating experiments were done taking the *E. coli* strain HB101 harbouring plasmid pRK2013 (helper) along with *E.coli* pir-116 harbouring pPDL2::Tn5<R6Kγori-Kan2> (donor) and *Acinetobacter* sp. DS002 (recipient). All three parental strains were grown to

mid log phase and were mixed in a ratio of 1:3:1 of donor, recipient and helper before plating them on LB plates having no antibiotics. The plates were incubated at 37 ^{o}C for 12 h and the cell mass grown on LB was scraped to resuspend in 1 ml of citrate saline buffer. The suspended cells were serially diluted and each dilution was independently plated on selection plates having of minimal media supplemented with 5 mM benzoate and kanamycin. Parents strains treated in similar manner were plated on selective plates and were served as negative controls.

2.28.2 Horizontal transfer of pCMS1

B. diminuta (pCMS1::*tet*) (donor) and *P*. *putida* KT2440 (recipient), were grown at 30 ^{o}C for log phase in LB media supplemented with tetracycline and chloramphenicol respectively. Cultures grown to mid log phase and were collected to harvest the cells. The cells were then thoroughly washed with citrate saline buffer before equalizing the OD of the culture. The donor and recipient cells were then mixed in 3:1 ratio and the cells were pelleted by centrifugation at 8000 rpm. The cell pellet was then resuspended in minimal amount of citrate saline buffer and plated on LB plates having no antibiotic. The plates were then incubated for 72 hrs at 30 ^{o}C. The cell mass collected from LB plate was serially diluted and were independently plated on LB plates having chloramphenicol and tetracycline. Donor and recipient cells treated in similar manner were also plated on the LB selection plates (Tc and Cm) and on LB plates having either tetracycline or chloramphenicol.

2.29 *In vivo* transposition assay

In vivo transposition assay was developed by using three compatible plasmids. Of the three plasmids one of them is an expression plasmid which codes for transposase (pMMB-TnpA). The second plasmid was a reporter plasmid which contained *sacB* gene. Disruption of *sacB* due to transposition can be monitored by growing the culture on LB plates having

sucrose. Therefore the second plasmid is designated as a reporter plasmid and in the present study pJQ210 is used as a reporter plasmid (Quandt and Hynes, 1993). The third plasmid has transposon-like *opd* cluster, where the *opd* gene is insertionally inactivated by inserting a tetracycline resistance gene (Siddavattam et al, 2003). The cultures have these three plasmids were grown in LB broth in presence of appropriate antibiotics till log phase and 0.2 mM IPTG was added to induce the transposase for 1 hours. After the induction period the cells were harvested and plated on LB plates having 5% sucrose and tetracycline, to select colonies having *sacB* gene disruptions due to transposition of *tet*-resistant transposon-like *opd* cluster (Siddavattam et al, 2003).

2.30 Identification of catabolic intermediates of benzoate

Catabolic intermediates of benzoate were determined by using LC/MS. The spent medium (1 ml) from the *Acinetobacter* sp. strain DS002 grown culture was for every 8 h and centrifuged at 15,000 g to get a clear supernatant. The supernatant thus obtained was acidified with 10µL of concentrated HCl and metabolites were extracted twice with 2 mL of ethyl acetate. The extract was air-dried and the residue dissolved in 30% (v/v) ACN containing 1% (v/v) acetic acid. A 20µL sample was injected by autosampler into Eclipse XDB-C18 column (5mm x 4.6 x 150 mm) using a binary pump HPLC (Agilent 1200 series). An isocratic buffer system of 30% (v/v) acetonitrile having 1% (v/v) acetic acid was used at a flow rate of 0.8 mL min^{-1} for 20 min. The spectrum was recorded for 10 min at 254 nm in the UV range and part of the elute was diverted to ESI source of Q-TOF (Bruker Daltonics). Mass spectra were acquired in negative mode with collision energy of 5ev, capillary temperature 200^{o}C and a source voltage of 3.8kV. Mass was acquired in the m/z range of 50 to 1000. The data was averaged for each peak and smoothed using Gaussian algorithm. Calibration of the instrument was done using ES Tunning mix (Agilent technologies) diluted 1: 60 v/v with 5% ACN. MS/MS of the base peaks at 1.7, 2.1 and 3.9 min were done in MRM mode with

10-15ev collision energy.

2.31 Determination of benzoate induced genome-wide expression profile

To determine benzoate induced genome-wide expression profile of *Acinetobacter* sp. DS002, proteome maps for the soluble proteins were generated by culturing DS002 in minimal media supplemented with either 5 mM benzoate or 50 mM benzoate or 10 mM succinate. Proteins were extracted sequential from these cultures by using a set of extraction buffers differing in their solubilizing capacity (described below). The extracted proteins were first resolved on IEF and then were analyzed on 12.5% SDS PAGE.

2.31.1 Two-dimensional electrophoresis

2.31.2 Preparation of protein sample

Acinetobacter sp. DS002 was grown in minimal media supplemented with either 5 mM benzoate or 50 mM benzoate or 10 mM succinate till they reach mid-log phase. The cell pellet was collected by centrifuging at 15000g for 10 min. The cell pellet was washed with minimal media and then with 20 mM Tris pH 8.0 to remove the salts. Proteins were extracted serially from 200 mg of cell pellet with 3 buffers, buffer I (40 mM Tris, pH 8.0), buffer II (40 mM Tris, 8 M Urea, 4% (w/v) CHAPS, 2 mM DTT, 0.2% (v/v) Ampholytes) and buffer III (40 mM Tris, 6 mM Urea, 2 mM Thiourea, , 4% (v/v) CHAPS, 2 mM DTT, 0.2% (v/v) Ampholytes).The proteins extracted using buffer II were precipitated with cold acetone and the pellet was washed with absolute alcohol. The protein pellet was air dried prior to its dissolution in 340 µL of rehydration solution (8 M Urea, 2% (v/v) CHAPS, 10 mM DTT and 0.2% (v/v) Ampholytes). Protein was estimated using Bradford method after TCA-acetone precipitation (Bradford, 1976).

2.31.3 Isoelectro focusing (IEF)

100 µg of protein extracted with the buffer II was taken in 350 µL of sample buffer and loaded onto IPG strips (18 cm, pH 3-10) and rehydrated. The strips were actively rehydrated for 12 hrs at 20 V and IEF was performed on Ettan IPGphor3 system using a four step programme [500 V for 30 min (gradient), 500 V for 30 min (step), 8000 V for 3 hrs (gradient) and continued till 40,000 Vhr].

After focusing, strips were equilibrated with equilibration buffer I (75 mM Tris, 6 M Urea, 2% (w/v) SDS, 20% (v/v) glycerol, 2% (w/v) DTT) and equilibration buffer II (75 mM Tris, 6 M Urea, 2% (w/v) SDS, 20% (v/v) glycerol, 2% (w/v) Idoacetamide) for 15 min in each. The second dimension electrophoresis was performed on 12.5% polyacrylamide gels in a Ettan DALTsix system (GE Healthcare) at a constant voltage of 200 V. The gels were stained with coomassie brilliant blue R-250 and image analysis was performed with ImageMaster 2D Platinum software (GE Healthcare). After the spot detection and the volume of each spot was taken for construction of the histograms. Protein spots whose volume ratios are more than 2 times were considered as over expressed and were picked for further analysis.

2.31.4 In-Gel Digestion

The coomassie brilliant blue stained protein spots were excised manually from the gel and the gel pieces were destained with a solution of ACN and 25 mM NH_4HCO_3 in 1:1 (v/v), dried with ACN and reduced with dithiothreitol (DTT) in 25 mM NH_4HCO_3 for 1 hrs at 50^oC, alkylated using a 55 mM iodoacetamide in 25 mM NH_4HCO_3 for 45 min at room temperature. The gel pieces were dehydrated with ACN, rehydrated with a minimum volume of 50 mM of NH_4HCO_3 containing Trypsin (10 ng/µL), and digested at 37^oC for 16h. The peptides were extracted twice with 50% (v/v) acetonitrile containing 1% (v/v) TFA. The peptide mixture was concentrated under vacuum in a Concentrator (Eppendroff) for 1h at room temperature.

2.31.5 MALDI-MS

The tryptic peptides were dissolved in 2 µL solution of 50% (v/v) ACN containing 1% (v/v) TFA and mixed with 2 µL of 1% cyano-4-hydroxycinnamic acid (HCCA) dissolved in 50% ACN and 1% TFA and 1µL of it was applied on the MALDI target plate. Peptides were analysed using MALDI TOF/TOF Autoflex (Bruker Daltonics) in reflectron mode. MS/MS of selected peptides were performed by LIFT. The spectra were calibrated by Pepmix (Bruker Daltonics).

2.31.6 Protein Identification

The spectral data were analyzed using Biotools software and searches were performed for protein identification using MASCOT search engine (www.matrixscience.com) against Swiss-prot (http://www.expasy.ch/sprot) and NCBInr (http://www.ncbi.nlm.nih.gov/) databases. The following search parameters were used: trypsin is the enzyme and one missed cleavage was allowed, the peptide tolerance was set at ± 0.5 Da, carbamidomethyl and oxidized methionine were set as fixed and variable modifications respectively. MS/MS data was analyzed using Biotools software and mass tolerance of ± 0.2 Da was used.

2.32 Protein estimation

An aliquot of sample was pipetted out into a 1.5 ml microfuge tube. The volume of the sample was adjusted to 0.1 ml with double distilled water. One ml of Bradford reagent was added and the contents were mixed by gentle inversion. After 15 min, absorbance at 595 nm was measured spectrophotometrically against a reagent blank prepared from 0.1 ml of appropriate buffer and 1 ml of Bradford reagent. Protein concentration of the sample

was calculated from a standard curve drawn using bovine serum albumin.

2.33 Catechol 1,2- dioxygenase assay

Catechol 1, 2- dioxygenases (C12O) was assayed spectrophotometrically by following the procedures described elsewhere (Briganti et al, 1997). Formation of the product, *cis-cis* muconic acid from catechol was monitored using spectrophotometer at 260nm. Reactions were performed in 20 mM Tris-Cl buffer; pH 8.0, with appropriate amount of protein and the reaction was initiated by the addition of 10 µl of 10 mM catechol. Further, substrate specificity experiments were done using a oxygraph with various aromatic compounds such as 1,4 benzene diol, 4-nitrocatechol, 3-chloro catechol, 4-chlorocatechol and PNP were used as substrates at final concentrations of 1mM.

2.34 Purification of catechol 1, 2 dioxygenase

2.34.1 Preparation of cell free extracts

The cell free extracts of *Acinetobacter* sp. strain DS 002 were prepared from the cultures grown in Minimal medium supplemented with 5 mM Sodium benzoate for 16 hours. The cells were harvested by centrifuging at 8000 rpm for 10 min and the cell pellet was washed twice 20 mM Tris buffer pH 8 and resuspended in the 7 volumes of 50mM Tris buffer (pH 8.0). After resuspension of the cells were disrupted by sonication for a period of 10 min with a pulse of 30 sec. The resulting homogenate was centrifuged at 15,000 rpm for 30min and the supernatant was again centrifuged at 45,000 rpm for 1 hour. The supernatant thus obtained was considered as the cytoplasmic fraction and was used for purification of catechol 1, 2- dioxygenase.

2.34.2 Ammonium Sulphate Fractionation

To the cytoplasmic fraction, ammonium sulphate was added slowly with constant stirring at 4 ^{o}C in order to bring the solution to 20% saturation. Later, the solution was

centrifuged at 15,000 rpm for 30 min to pellet out the precipitated proteins. Similarly, the supernatant obtained was further saturated to 40% and 60% respectively by adding appropriate amounts of ammonium sulphate. The proteins that are precipitated in the 40-60% fraction was used for purification of catechol 1, 2 dioxygenase.

2.34.3 Anion Exchange Chromatography

Anion exchange chromatography was done on manually packed DEAE Sepahrose column (XK 16/40) operated on an AKTA basic FPLC system (Amersham Biosciences, UK). The protein sample was loaded on to the column pre-equilibrated with 50 mM Tris-Cl (pH 8.0) buffer having 0.18M NaCl. The protein was eluted at a flow rate of 0.4 ml/min with a linear gradient of 1M NaCl in 50mM Tris buffer.

2.34.4 Hydrophobic Interaction chromatography

Hydrophobic Interaction chromatography was performed on a manually packed Phenyl sepharose column operated under gravity flow. The proteins from the anion exchange fractions having high C12O activity were pooled, and loaded on to the phenyl sepharose column pre-equilibrated with a 50 mM Tris-Cl buffer having 0.18 M NaCl and 15% $(NH_4)_2SO_4$. The proteins were eluted using a 50 mM Tris-Cl buffer having 0.18 M NaCl with a negative stepwise gradient of $(NH_4)_2SO_4$.

2.34.5 Gel Permeation Chromatography

Gel permeation chromatography was performed on a manually packed Sephacryl 200 HR columns (XK 16/100) operated at a flow rate of 0.5ml/min using an AKTA basic FPLC system (Amersham Biosciences, UK). The protein from the Phenyl sepharose fractions having high C12O activity were pooled and concentrated by ammonium sulphate precipitation. Further, the protein was dialysed to remove the ammonium salt and then

loaded on to the Sephacryl 200HR column pre-equilibrated with 50mM Tris-Cl (pH 8.0) buffer having 0.18M NaCl.

2.34.6 Organophosphorus hydrolase assay

Acinetobacter sp. DS002 cells and *E.coli pir*-116 cells harbouring plasmid pPDL2-Tn*5*<R6Kγori/KAN-2> were grown to log phase and the cells were harvested by centrifugation at 8000 rpm for 10 min. The cell pellets obtained was then washed twice with saline and was finally resuspended in 5 ml of 10mM Tris HCl (pH 8.0) buffer. The cells were broken by sonication at 4 ^{o}C with a pulse of 30 sec for 5 min. The lysate was centrifuged at 13000 rpm for 10 minutes. The supernatant and pellet fractions were used to assess paraxonase activity. 1 ml of reaction mixture contained 790 µl of double distilled water and 100 µl of 200 mM CHES and 10 µl of enzyme source (cytoplasmic fraction or pellet fraction). Whenever needed 10 µm of CoCl2 was added to the reaction mixture. The reaction was started by adding 100 µl of 10 mM paraoxon and mixed by inverting. An increase in the absorbance at 410 nm due to formation of *p*-nitrophenol was determined (Chaudhry et al, 1988). The concentration of *p*-nitrophenol formed in the reaction was determined using the extinction coefficient (16500 M-1 Cm -1). The specific activity of the enzyme was expressed as micromoles of PNP produced per mg of protein.

As mentioned in the introduction chapter, OP compound degrading *Flavobacterium* sp. ATCC 27551 and *Brevundimonas diminuta* MG were isolated from agricultural soils collected from IRRI, Philippines (Sethunathan et al, 1973) and Texas, USA (Serdar et al, 1982) respectively. After these two reports a number of OP compound degrading bacteria belonging to different taxonomic groups were isolated from diverse geographical regions (Zhongli et al, 2001; Horne et al, 2002; Liu et al, 2005; Karpouzas and Singh, 2006; Singh, 2009). In all these OP degrading bacterial strains, a phosphotriesterase (PTE) is shown to be responsible for the hydrolysis of triester linkage found in structurally diverse groups of OP compounds (Benning et al, 1994; Cho et al, 2004). In *Brevundimonas diminuta* and *Flavobacterium* sp. ATCC 27551, the phosphotriesterase is coded by an identical organophosphate degrading (*opd*) gene present on large indigenous plasmids, pCMS1 and pPDL2 respectively (Serdar *et al*., 1982; Mulbry and Karns, 1986). Further studies from our laboratory have shown transposon-like *opd* gene cluster in *Flavobacterium* sp. ATCC 27551 (Siddavattam et al, 2003). In both pCMS1 and pPDL2 almost identical *opd* sequences were found. The DNA region identified 2.6 kb upstream and 1.7 kb downstream of the *opd* gene was found (Mulbry et al, 1988; Siddavattam et al, 2003). Beyond this region no detectable homology was found between these two indigenous plasmids. Such observation suggests existence of horizontal mobility of *opd* genes among soil bacteria. Horizontal mobility of genetic information occurs through various genetic elements such as plasmids, bacteriophages, genomic islands, Integrons, transposons, conjugative transposons and IS elements (Dobrindt et al, 2004). Existence of *opd* genes on large indigenous plasmids strengthens the hypothesis of spreading *opd* information through horizontal mobility. However, till date no experiments were conducted to validate if lateral gene transfer is contributing

for distribution of *opd* information among soil bacteria. Existence of *cis*-elements that contribute for horizontal mobility of plasmids, as well as structural information pertaining to the organization of *opd* information will be known from the primary sequence of the plasmids. Therefore, as a basic requirement to understand the HGT of *opd* gene, complete sequence of pPDL2 isolated from *Flavobacterium* sp. ATCC 27551 is determined.

3.1. Isolation and rescue cloning of Indigenous plasmid pPDL2 from *Flavobacterium* sp. ATCC 27551

The plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551 is a low copy number, large indigenous plasmid of 39.75 kb in size. In order to obtain sequence of pPDL2 it has to be isolated and sub-cloned in multipurpose vectors. Further, in *Flavobacterium* sp. ATCC 27551 there are more than one plasmid. While sub-cloning pPDL2 it has to be isolated in a pure form. Therefore, plasmids pPDL2 was rescue cloned into *E. coli pir*-116 cells following procedures described in materials and methods. Initially all plasmids from *Flavobacterium* sp ATCC 27551 were isolated using a modified protocol of Courier and Nester method described in materials and methods section (Currier and Nester, 1976) and were analysed on agarose gels. As reported by Mulbry and his associates, plasmid preparations made from *Flavobacterium* sp. ATCC 27551 have revealed existence of four plasmids (Fig. 3.1) (Mulbry et al, 1986). Out of these four plasmids existence of *opd* was reported only in plasmid pPDL2. The *opd* plasmid pPDL2 was rescued from the rest of the three indigenous plasmids by rescue cloning technique.

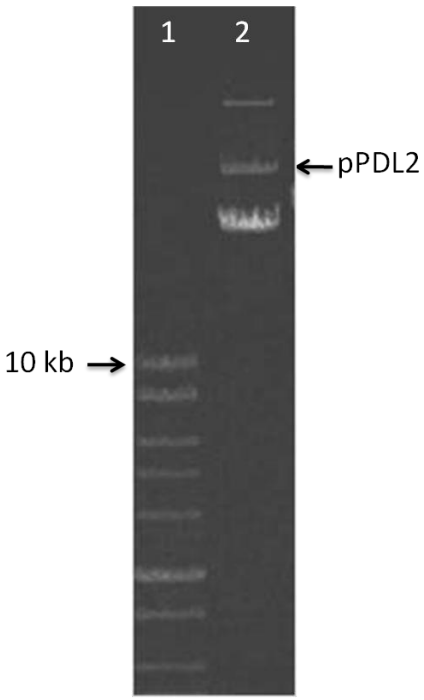


Fig. 3.1. Isolation of indigenous plasmids from *Flavobacterium* sp. ATCC 27551. Lane 1 represents 1 kb DNA ladder and lane 2 represents plasmid preparations from *Flavobacterium* sp. ATCC 27551. Plasmid pPDL2 is shown with an arrow mark.

3.2. Rescue cloning of plasmid pPDL2

The plasmid preparation containing mixture of four plasmids were tagged with R6Kγori replication origin containing mini-transposon EZ-Tn5<R6Kγori/Kan2>. These plasmids were then transformed into *E. coli pir*-116 cells and the kanamycin resistant colonies were then used for doing colony PCR using *opd* specific primers. Out of 100 colonies screened only 30 colonies gave amplification of *opd* gene indicating existence of pPDL2 in these plasmids. The authenticity and purity of pPDL2 was established by performing restriction analysis and by amplification of *orf306*, an ORF adjacently located to the *opd* gene (Siddavattam et al, 2003). Amplicons of *opd* and *orf306* obtained from the rescued clones coincided with the amplicon size obtained from *Flavobacterium* sp ATCC 27551 used as positive control, suggesting successful rescuing of pPDL2 into *E. coli*

pir-116 cells (Fig. 3.2). Further the restriction profile generated to the rescued plasmid has perfectly coincided with the similar profile reported by Mulbry and his associates (Mulbry et al, 1986),

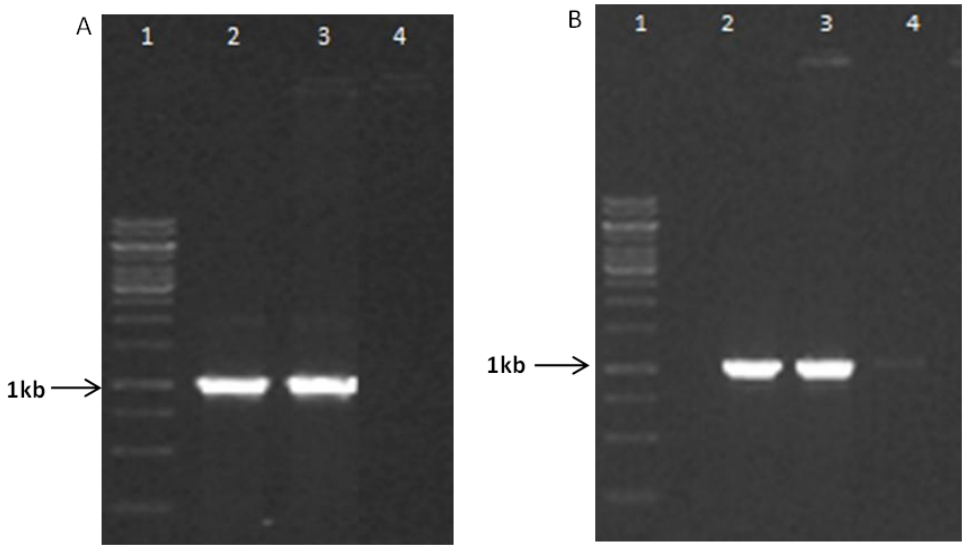


Fig. 3.2. Amplification of plasmid pPDL2 borne *opd* and *orf306* from *Flavobacterium* sp ATCC 27551 and *E. coli pir*-116 containing rescue cloned pPDL2:: Tn*5*<R6Kγori/Kan2>. Panel A, Lane 1 represents 1 kb DNA ladder. Lane2 and 3 represents amplicons of *opd* obtained from *E. coli pir*-116 (pPDL2::Tn*5*<R6Kγori/Kan2>) and *Flavobacterium* sp. ATCC 27551 respectively. Lane 4 represents negative control where a colony of *E.coli pir*-116 cells was used while performing colony PCR. Panel B indicated similar loading pattern except that *orf306* specific primers were used while performing colony PCR.

except that the size of the 14.7 kb large *Eco*RI fragment increased by 2kb. Further, a 5.7 kb *PstI* fragment found on plasmid pPDL2 has disappeared in pPDL2::Tn5<R6Kγori/Kan-2>. *In lieu* of that, two new bands with a size of 3 kb and 2.5 kb were seen after digestion with *PstI.* Obviously, this is due to existence of an internal *PstI* site in mini-transposon Ez-Tn5<R6Kγori-Kan2> (Fig. 3.3). With the exception of the increase in the size of large *Eco*RI fragment by 2 kb (A-14.7kb) and disappearance of the 3rd largest *Pst*I fragment (C-5.7kb), the restriction profile of pPDL2 perfectly matched with the restriction profile of pPDL2:: Tn*5*<R6Kγori/Kan2> (Fig. 3.3). The *Pst*I and *Eco*RI restriction profiles of

pPDL2::Tn5<R6Kγori/Kan2 gave 10 and 5 fragments respectively (Fig. 3. 3). Rescuing of pPDL2 was found to be advantageous in number of ways. One of them was apparent

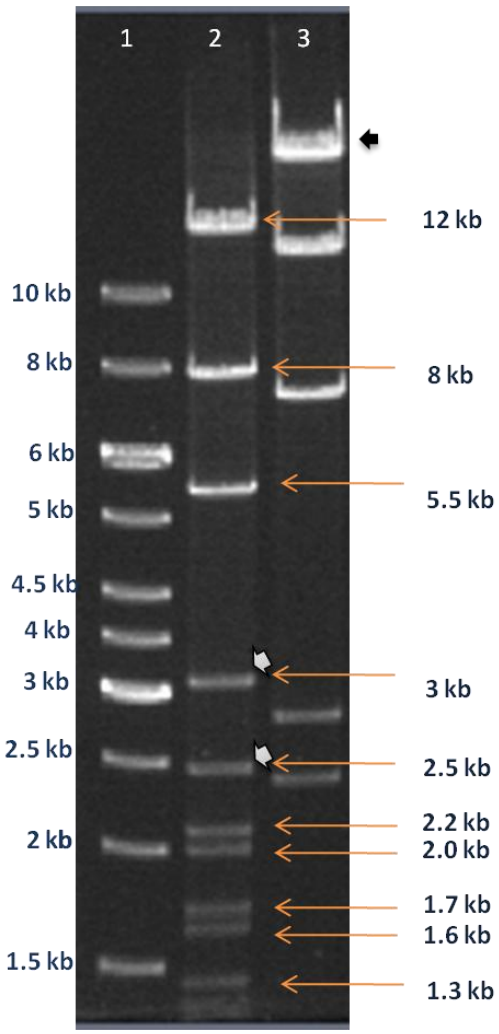


Fig. 3.3. Restriction profile of plasmid pPDL2::Tn5<R6Kγori/Kan2>.Lane 1 represents 1 kb DNA ladder. Lanes 2 and3 represents restriction profile of pPDL2::Tn5<R6Kγori/Kan2> generated by digesting with *PstI* and *Eco*RI respectively. Increase in size of *Eco*RI is indicated with a black arrow. Additional *PstI* fragments generated due to existence on mini-transposon specific *PstI* are indicated with open arrows

increase in copy number. Due to increase in copy number, pPDL2::Tn5<R6Kγori/Kan2> could be isolated from *E. coli pir*-116 cells by using mini-prep protocols optimized for isolation of high copy number plasmids from *E. coli* (Fig. 3.4). Such easy isolation of

plasmid pPDL2::Tn5<R6Kγori/Kan2> facilitated for easy sub-cloning of its fragments in multipurpose vectors.

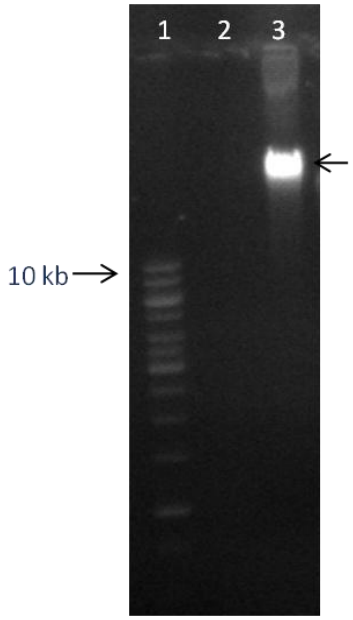
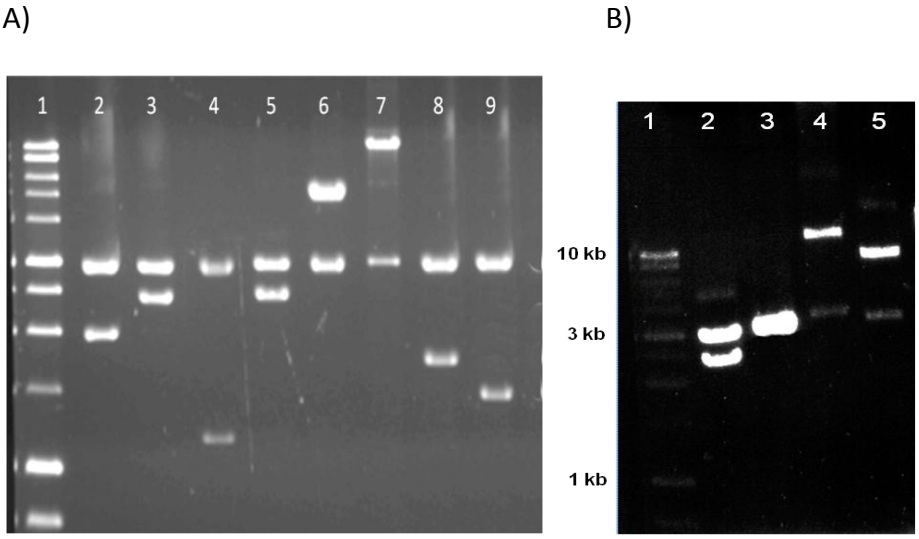


Fig. 3.4. Isolation of rescued plasmid pPDL2::Tn5<R6Kγori/Kan2> from *E.coli pir* 116 cells. Lane 1. Represents 1kb DNA ladder. Lanes 2 and 3 represent plasmid profile of *E.coli pir*-116 and *E.coli pir-116* (pPDL2::Tn5<R6Kγori/Kan2>). Plasmid pPDL2::Tn5<R6Kγori/Kan2> is shown with an arrow mark.

3.3. Sequencing of pPDL2 and annotation:

As stated before, 10 fragments were generated when plasmid pPDL2::Tn5<R6Kγori/Kan2> was digested with *Pst*I. All the 10 fragments with the following sizes 12 kb, 8 kb, 5.5 kb, 3 kb, 2.5 kb, 2.2 kb, 2.0 kb, 1.7 kb, 1.6 kb and 1.3kb were sub-cloned in pBluescript-II KS digested with similar enzymes (Fig. 3. 5). Similarly, 4 *Eco*RI fragments, with the sizes of 11 kb, 6.7 kb, 3.0 kb and 2.5 kb were cloned in pBluescript-II KS digested with *Eco*RI. Description of the recombinant plasmids containing different pPDL2 fragments used for sequencing is shown in Table 3.1.



3. 5. Sub-cloning of pPDL2. Lane 1 represents 1kb DNA ladder. Panel A shows shows sub-clones of pPDL2 generated by ligating *PstI* fragments in pBLuescriptII vector. Panel B shows similar sub-clones of *EcoRI* fragments. Refer table 3.1 for size description.

3.4. Sequence strategy

The detailed strategy used to obtain complete sequence of pPDL2 was shown in Fig. 3.6. Initially the sub-clones were directly used to generate sequence using universal forward and reverse primers. After obtaining plasmid pPDL2 specific internal sequence, fragment specific primers were designed to obtain entire sequence of the fragment through gene walking strategy. A detailed strategy used for obtaining the complete sequence of pPDL2 is shown in Fig. 3.6. Clones pE5 (2.2kb), pE12 (6.5kb), pP33II (12kb), pP4I (8kb), pP4II (5.5kb) and pP3II (5.7kb) were initially sequenced using vector specific primers and then were sequenced using primer walking strategy. Clone pP33II is 12kb in size. In order to reduce the size of the insert plasmid pP33II was further digested with both *PstI*, *SalI* and *Eco*RI and sub-cloned in pBluescript KSII as *PstI-Eco*RI (p33EP, 8kb) and *PstI-Sal*I (p33SP, 3kb) fragments. While obtaining the sequence of junction regions appropriate primers were designed and plasmids with overlapping fragments were used

Sub-clone of pPDL2	**Description**
pE10	11.0 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS
pE9	3.0 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS
pE5*	2.5 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS
pE12*	6.7 kb *EcoRI* fragment of pPDL2 cloned in pBluescript II-KS
pP33II*	12 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS
pP4I*	8 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS
pP3II*	5.5 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS
p33EP*	8 kb *Eco*RI-*PstI* fragment of pP33II cloned in pBluescript II-KS
p33SP*	4.5 kb *Sal*I-*PstI* fragment of pP33II cloned in pBluescript II-KS
pP1I*	1.5 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS
pP3I*	2.5 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS
pP7I*	2.2 kb *PstI* fragment of pPDL2 cloned in pBluescript II-KS

Table 3. 1. Details of *EcoRI* and *PstI* fragments of pPDL2::Tn5<R6Kγori-Kan-2> sub-cloned in pBluescript. Sub-clones used for sequencing are indicated with * mark.

as template to generate sequence reactions. The strategy followed to get the complete sequence was shown in Fig. 3.6.

3.5. Sequence assembly and analysis

Sequence assembly of pPDL2 was done using the contigexpress software of VectorNTI. The chromatograms were compared using contigexpress and contigs were created with chromatograms having good quality bases. All the contigs were then aligned using

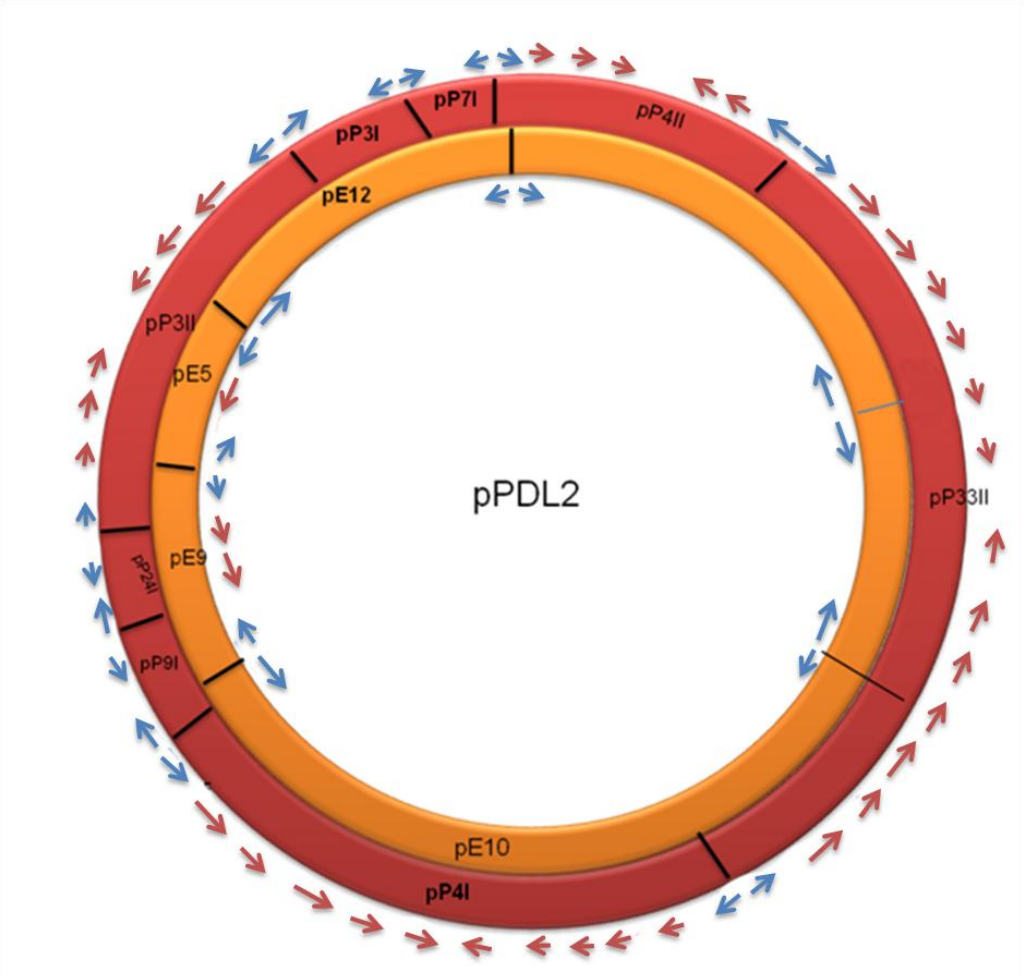


Fig. 3.6. Strategy used for sequencing of plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551. The outer and inner circles represent *Pst*I and *Eco*RI fragments of pPDL2. Arrows indicate position of primers used for sequencing of pPDL2. Blue coloured arrows show vector specific primers. Orange and maroon coloured arrows indicate the primer positions used to sequence *Pst*I and *Eco*RI fragments cloned in pBluescript vector.

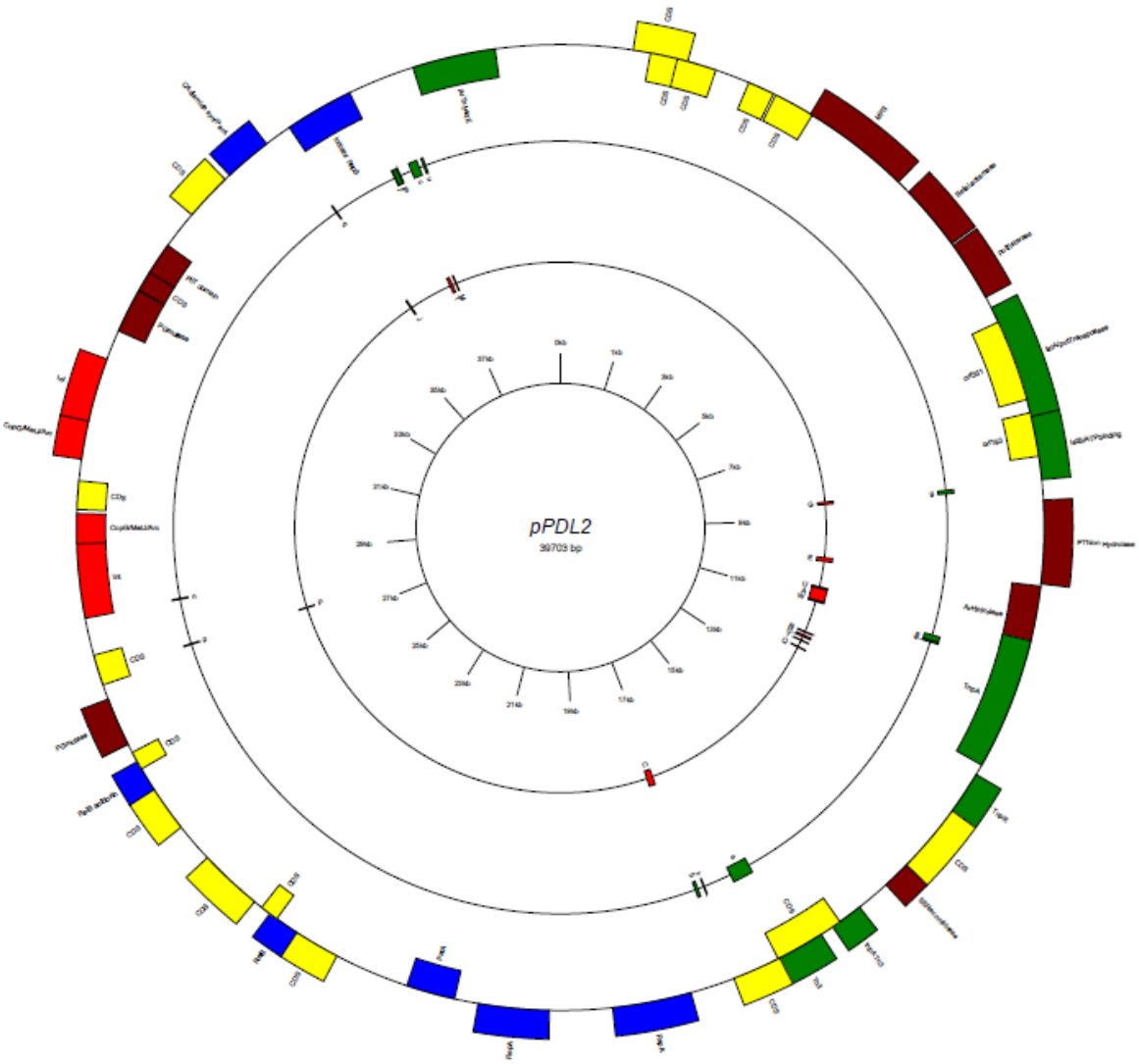
Primer Name	Sequence of the primer	Primer position
E5F	5’-TAAGGATAGTGGGACGTCGC-3’	3838 – 3819
E5R	5’-GTCGTGGGCGTCGTTAAGCTG-3’	2452 -2472
P3R	5’-TATCCTTGATGCCGAAGACC-3’	6509 – 6490
E13R	5’-GCCGGCAGGATATAGGTT-3’	28028 -28045
E13F	5’-CACCTCACCAGCAATTCGTA-3’	38306 -38287
E12F	5’-AGCCTGTTGACGCAGAAAGT-3’	4498 -4517
P33F	5’-GTTCGCGATCGTCAAGAACT-3’	29779 – 29760
P4I-Ext4	5’-GTTCTACAACACGCTGAAC-3’	38999 -39017
P4II-M	5’- GTCCGGGCTGATGAAATATG-3’	13915 -13934
P4IR4	5’-ACTCGCTGGCCTATGTGTTC-3’	32593-32574
P33-R4	5’-ATCCCGATCTGTTCATTTGC-3’	18240-18259
P4F	5’-ACAATTTCCAGGTCGTCACC-3’	15599 -15580
E12F-internal2	5-'GCCTCAATCTGGTGTTCGAT-3'	5157 -5176
E13F-internal2	5'-CGAATTGGTGGGATTTGTCT-3'	37795 -37776
E13R-internal2	5'-CCATCGCCTAGATCAACACC-3'	28576 -28595
P3RII-internal	5'-ACGATGTCGTCGATGTGTGT-3'	6248 -6229
p33FII-internal2	5'-CTGTTCCGCACGATCGCG-3'	31865 -31882
P33R-Internal	5'-TATCTCGCTCACCGGCGACT-3'	11523 11504
P4FII internal	5'-TGGCGTTGATCGGCTATG-3'	15341 -15324
P4RI external	5'-CACCAGGCCAACAAGAAATC-3'	30585 -30566
P4RI internal	5'-ATGGCCGATTGGCTGCTGGC-3'	29841 -29822
P4FI-Internal	5'-GCACTTTCGTGTAGTGACCCC-3'	16962 -16981
P4RI-Internal2	5'-ATCGCGCGCGGCACTAAGC-3'	31877 -31895
P4RII-Internal2	5'-AGATCCACCATCTATCGCGA-3'	13832 -13851
P3FII-Internal	5'-GAAGTCACCGAGGAGCACTT-3'	2209 -2228
H10_P33RII(External)	5'-CGGTGTAGTGCTCCTCGATT-3'	16219 -16200
D07_P4FI(External)	5'-CTTTGACTTCATCCGGCAGT-3'	37893 -37912
G09_P4FII(External)	5'- AGTTGTCGCATTCTCGATCC-3'	15852 -15871
E09_P3FII(External)	5'- GCACGATGTTCTTCGACCTT-3	1886 -1867
P33SPF-E	5'- CGCTGAATCTGAACTGACGA-3'	21983 -22002
p33F2-g	5'-CGGCTTTCGGCATCCAACCT-3'	28414 -28395

Table 3. 2. Primers used for sequencing of plasmid pPDL2 and their sequences

Pairwise alignments to get the complete sequence of pPDL2. The complete sequence of pPDL2 is available in web page (www.uohyd.ernet.in/uploads). After generating the complete sequence it was analyzed to indentify GC ratio, number of ORFs, inverted and direct repeats, promoter elements and other *cis*-elements that play a predominant role in integration excision and mobilization of plasmids.

3.6. The GC composition

The 39.75 kb sequence of pPDL2 has shown high similarity to either chromosomal or plasmid DNA sequences of *Sphingobium* and *Sphingomonas* sp. Since, plasmid pPDL2 has shown sequence similarity to the genome sequence of *Sphingomonas* and *Sphingobium,* the GC content of the total genome sequences were obtained from genome database. Based on the total genome sequence found in database (www.**ncbi**.nlm.nih.gov/) the GC content of *Sphingomonas wittichi* RW1 is 68.4%. It contains two indigenous plasmids designated as pSWIT01 and pSWIT02 which have a GC content of 64.1% and 61.2% respectively. Similarly there are two chromosomes 1 and 2 in *Sphingobium japonicum* UT26S and each of them have a GC content of 64.8 % and 65.9 % respectively. In addition to these two chromosomes there are three circular plasmids in *Sphingobium japonicum* UT26S (NBRC101211). These three plasmids designated as, pCHQ1, pUT1 and pUT2 have a GC % of 63.0%, 63.7% and 61.0% respectively (NBRC101211). The G+C content of pPDL2 of *Flavobacterium* sp. ATCC 27551 is 61.76% and found to be very close to plasmids pSWIT02 of *Sphingomonas wittichi* RW1 and pUT2 of *Sphingobium japonicum* UT26S. Consistent with GC content, the proteins coded by the open reading frames (Table 3. 3) of pPDL2 have shown homology to the proteins coded by the *Sphingomonas*. The 39.75kb plasmid pPDL2 codes for 41 open reading frames. Out of 42 predicted ORFs 18 of them are hypothetical proteins that show (31-95%) homology to the hypothetical proteins of *Sphingomonas* or *Sphingobium* sps. (Table 3.3, Fig. 3.7).



.3. 7. Circular map of plasmid pPDL2. The ORFs specified by sense or positive and antisense or negative strand are shown in outer and inner circles respectively. Colour boxes represent ORFs. Very distinct colours are used to indicate possible function of ORFs. Yellow, for all hypothetical proteins; dark green for transposons and IS elements; Dark blue, for proteins involved in plasmid replication and partition; Red to show phage integrases and CopG family transcription factors and brown colour for all ORFs which have accessory functions.

The ORFs identified in plasmid pPDL2 are given generic, conventional and functional names. The generic name indicates the serial order of the ORF on plasmid pPDL2. In the conventional names the ORF is followed by a number which is equivalent to the number

Table 3.3 Open reading frames present on plasmid pPDL2

Generic Name	Position	Conventional Name	Putative function	Homologous protein	Amino acid identity (%)	GenBank accession no.
ORF1	954-1682	*orf242*	Unknown	Hypothetical protein	31	ZP_01623203
ORF2	1185-2067 C	*orf294*	P45O	β- component of Protocatechuate dioxygenase of *Xanthomonas axonopodis* pv. citri str. 306	78	gb|AAM35766.1
ORF3	2527-2880	*orf117*	Hydrolase	alpha-beta hydrolase fold protein of *Commamonas testosteroni* S44		gb|EFF44050.1|
ORF4	2915-3460	*orf181*	Hydrolase	Hydrolase of *Xanthomonas fuscans* subsp. aurantifolii str. ICPB		
ORF5	3433..4887	*orf484*	Transporter	MFS of *Asticcacaulis excentricus* CB 48	72	gb|ADU14055.1
ORF6	5075..5956	*orf293*	Antibiotic resistance	Beta-lactamase domain protein of *Methylobacterium nodulans* ORS 2060	48	gb|ACL58162.1
ORF7	5986-6786	*orf266*	Unknown	Hypothetical protein Swit_1907 *Sphingomonas wittichii* RW1	41	gb|ABQ68267.1
ORF8	6941-8464	*orf507*	Transposase(*istA*)	Transposase *Aurantimonas manganoxydans* SI85-9A1	59	gb|EAS49905
ORF9	8407-8988 C	*orf193*	Acetyl transferase	dihydrolipoamide acetyltransferase *Legionella pneumophila* str. Paris	22	emb|CAH12611.1
ORF10	8461-9300	*orf279*	Resolvase (*istB*)	IstB domain protein ATP-binding protein Nitrosomonas sp. AL212	78	gb|EET30478
ORF11	9567-10664	*orf365*	OP Hydrolase (*opd*)	Parathion hydrolase of P. diminuta	99	gb|AAA24930

ORF12	10696-11613	*orf305*	Aromatic hydrolase	Putative aromatic hydrolase		
ORF13	11424-13175 C	*orf583*	Transpoase (*tnpA*)	TnpA transposase Sphingomonas sp. KA1	86	dbj|BAF03245
ORF14	13316-13885	*orf189*	Resolvase (*tnpR*)	Resolvaseof *Sphingomonas* sp. KA1	96	dbj|BAE75870
ORF15	13886..14833	*orf315*	Unknown	Hypothetical protein SphchDRAFT_3708 *Sphingobium chlorophenolicum* L-1	85	gb|EFN09107
ORF16	14812-15186	*orf124*	Recombinase (*tnpR*)	Putative site-specific recombinase	63	gb|ACJ63562
ORF17	15640-16077	*orf145*	Transposase (*tnpA*)	Tn3 family transposase of plasmid pLB1	98	YP_740316
ORF18	15958-16902C	*orf314*	Unknown	Hypothetical protein of *Acinetobacter baumannii* AYE	40	emb|CAM88370
ORF19	16255-16974	*orf239*	Transposase (*tnpA*)	Tn3 family transposase of plasmid pLB1	98	YP_740316
ORF20	16878-17534	*orf218*	Plasmid replication initiation (*repAa*)	RepA of *Sphingobium japonicum*	96	dbj|BAI99177
ORF21	18125-19168	*orf347*	Plasmid replication initiation (*repAb*)	RepA of *Sphingobium japonicum*	97	dbj|BAI99177
ORF22	21246-21893 C	*orf215*	Partitioning (*parA*)	ParA of *Sphinogobium japonicum*	99	dbj|BAI99179
ORF23	22887-23507	*orf206*	Hypothetical protein	Hypothetical protein of *Sphingomonas wittichi* RW1	35	gb|ABQ66545
ORF24	23494-23922	*orf142*	Plasmid stability (*relB*)	RelB antitoxin of *Sphingobium japonicum*		dbj|BAI98982
ORF25	23810-24058 C	*orf82*	Unknown	Hypothetical protein		dbj|BAI99186
ORF26	25546-26166	*orf206*	Unknown	Hypothetical protein	95	dbj|BAI99186
ORF27	26469-26717 C	*orf82*	Hypothetical protein	Hypotheitical protein XAUC 31650	No significant	

					homology	
ORF28	26836-27483	*orf215*	Enzyme of glycolytic pathway (*pgm*)	Phosphoglycerate mutase family protein	95	dbj|BAI99186
ORF29	27679-28086 C	*orf135*	Unknown	Hypothetical protein	70	dbj|BAI99187
ORF 30	26153-26581	*orf142*	Antitoxin (*relB*)	RelB antitoxin of *S. japonicum* UT 26	96	dbj|BAI98982
ORF31	28583-29569 C	*orf328*	Invertase/ recombinase like protein (*int*)	Phage integrase family protein of *Sphingomonas wittichi* RW1	86	dbj|BAI98979
ORF32	29566-29961 C	*orf131*	Transcriptional factor (*copG*)	CopG/MetJ/Arc family protein	65	gb|ABQ71225
ORF33	30013-30390 C	*orf125*	Unknown	Hypotehtical protein	-	-
ORF34	30675-31184	*orf169*	Transcriptional factor (*copG*)	CopG/MetJ/Arc family protein	65	gb|ABQ71225
ORF35	31181-32027	*orf282*	Invertase/recombinase like protein (*int*)	Phage integrase family protein of *Sphingomonas wittichi* RW1	94	dbj|BAI98979
ORF36	32426-33014	*orf196*	Enzyme of glycolysis (*pgm*)	Phosphoglycerate mutase family protein	92	dbj|BAI99186
ORF37	33023-33305	*orf94*	Unknown	Conserved hypothetical protein of *Methylosinus trichosporium* OB3b	73	gb|EFH01001
ORF38	33296-33721	*orf141*	Signal transduction protein (*pilT*)	PilT domain containg protein of *Sphingomonas wittichi* RW1	59	gb|ABQ68989
ORF39	34228-34863	*orf211*	Unknown	hypothetical protein SJA_P1-00320 *Sphingobium japonicu* UT26	91	dbj|BAI99180
ORF40	34944-35579	*orf211*	Catalyzed conversion of cobyrinic acid to cobyrinic acid diamide	Cobyrinic acid ac-diamide synthase *Thauera* sp. MZ1T	58	gb|ACK55109
ORF41	35944-36854 C	*orf303*	Replication (*repB*)	RepB of *Gluconobacter diazotrophicus* PA15	83	gb|ACI53275
ORF42	37745-38847	*orf367*	Transposase (*y4qE*)	Putative transposase y4qE of *Roseibium* sp. TrichSKD4	59	gb|EFO28627

amino acids encoded by the *orf*. If any *orf* coded protein has significant homology to the functionally characterized proteins it is also given a functional name (Table 3. 3). While describing ORFs coded by plasmid pPDL2, they are divided into functional modules such as Replication and partition module, Mobilization module, Integration module, Degradation module and mobile genetic elements. Each functional module is independently described to facilitate easy description and understanding of plasmid pPDL2 sequence.

3.7. Replication and partition module

In general, replication module includes a well defined *oriV*, replication initiator protein, RepA, proteins involved in partitioning (Par) and plasmid maintenance (toxin-antitoxin modules).

3.7.1. Replicative origin (*oriV*)

The *oriV* generally contains sequence motifs (*cis*-elements) that interact with the replication initiator protein, RepA and other accessory proteins. In plasmids replicated through *theta* mode (Bramhill and Kornberg, 1988).They include (i) AT-rich region containing sequence repeats, often found to be located adjacent to RepA binding sites. At this AT-rich region of *oriV* the host initiation factors assemble immediately after its conversion into an open complex, and (ii) the next important *cis*-element found at origins of *theta*-replicating plasmids is DnaA binding box (Bramhill and Kornberg, 1988; Kornberg and Baker, 1992). Generally one or more *dnaA* boxes are found at the replicative origin of plasmids. Interaction of host encoded DnaA initiator protein is an essential event in initiation of plasmid replication process (Bramhill and Kornberg, 1988; Kornberg and

Baker, 1992). In pPDL2, two ORFs designated as *orf20* (18125-19168) and *orf21* (19995-20933) have shown very high homology (96%) to the RepA-like protein coding sequences of *Sphingobium japonicum* UT26S (Fig. 3.8). These two RepA sequences designated as

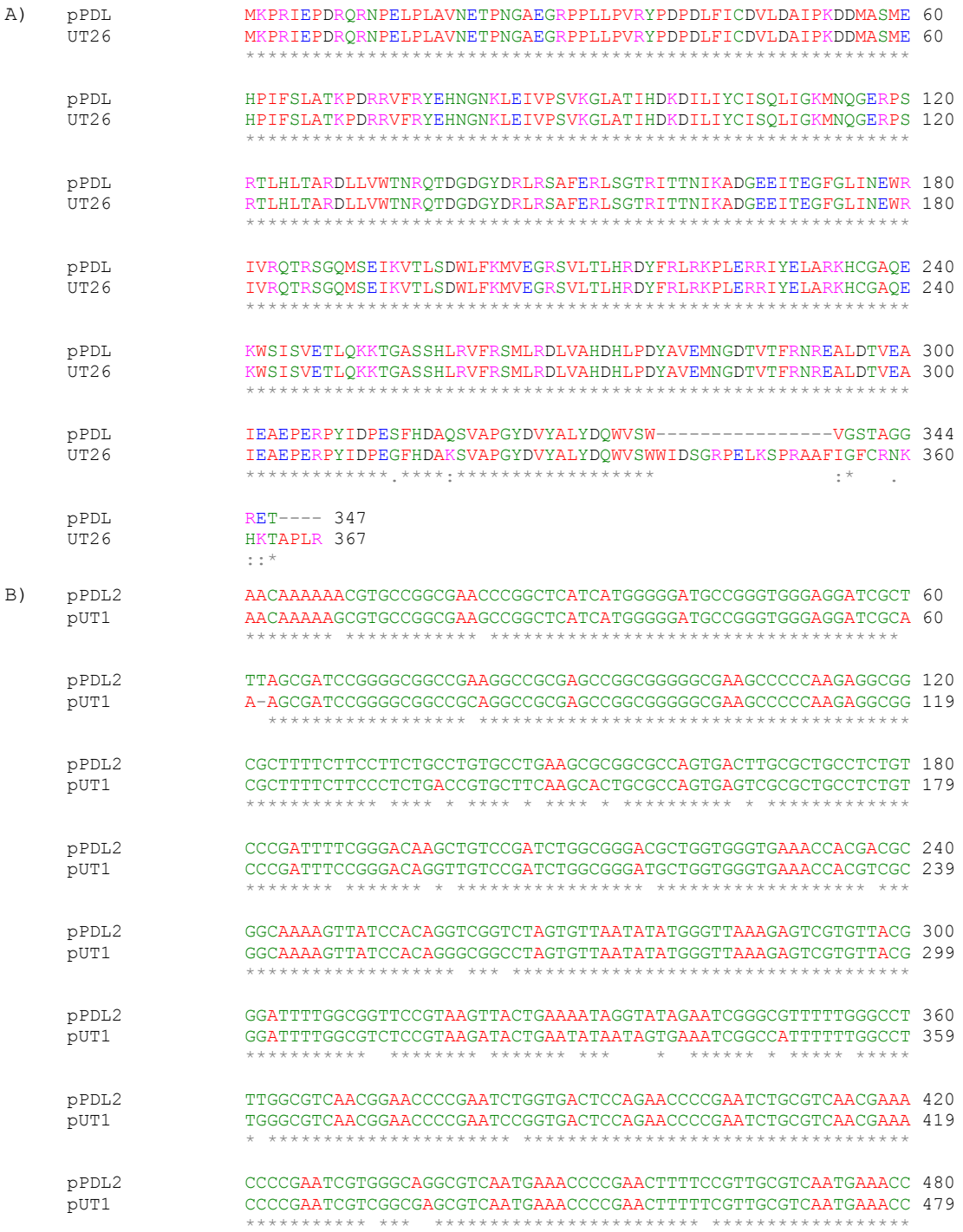


Fig. 3. 8. Pairwise alignment of A) RepAa of pPDL2 of *Flavobacterium* sp. ATCC 27551 with RepA of pUT1 of *Sphingobium japonicum* UT26S and B) alignment of *oriV* sequences of pPDL2 and pUT1 of *Sphingobium japonicum* UT26S. Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

RepAa and RepAb share 100% similarity. Though the reason for presence of two identical *repA* coding genes are unknown, considering its homology with well characterized RepAa protein of pUT1 of *Sphingobium japonicum* UT26S, the possible role of these two proteins in replication of plasmid pPDL2 is quite apparent. Since existence of two RepA proteins appeared unusual, a complete literature search was undertaken to gain more insights into this unusual phenomenon. As revealed by literature search, existence of more than one RepA is not uncommon. In an 184 kb indigenous catabolic plasmid pNL1 of *Sphingomonas aromaticivorans* F199 more than one *repA* genes were identified (Romine et al, 1999). However, no homology was seen between RepA of plasmid pPDL2 and RepA of pNL1 of *S. aromaticivorans.* When the upstream region of *repA* sequence was analyzed to indentify existence of promoter elements, a σ70 promoter was seen upstream of each *repA* gene, suggesting existence of two functional *repA* genes on plasmid pPDL2.

In order to assign incompatibility group to plasmid pPDL2, a two way approach was followed. Initially, the RepAa sequence was blasted to know its homologues from well characterized plasmids. The second approach was to identify similarity between the *oriV* sequence of pPDL2 and well characterized plasmid replicative origins available in the database. The RepAa sequence of pPDL2 has shown 96% identity to the RepA sequence of *Sphingobium japonicum* UT26S plasmid, pUT1 (BAI99177). Further, it has also shown about 85% identity to the plasmid, pAPA01-030 coded RepA of *Acetobacter pasteurians* IFO3283-01 and 62% identity with RepA of *Nitrospira multiformis* ATCC25196 (Fig. 3.9). As seen in phylogenetic tree constructed using blast output, no significant homology was found between RepA of pPDL2 to other bacterial RepA proteins (Fig. 3.9). However, the

plasmids of *Sphingobium* sp. to which *repA* of pPDL2 has shown strong identity are yet to be assigned with a distinct incompatibility group (AP010806).

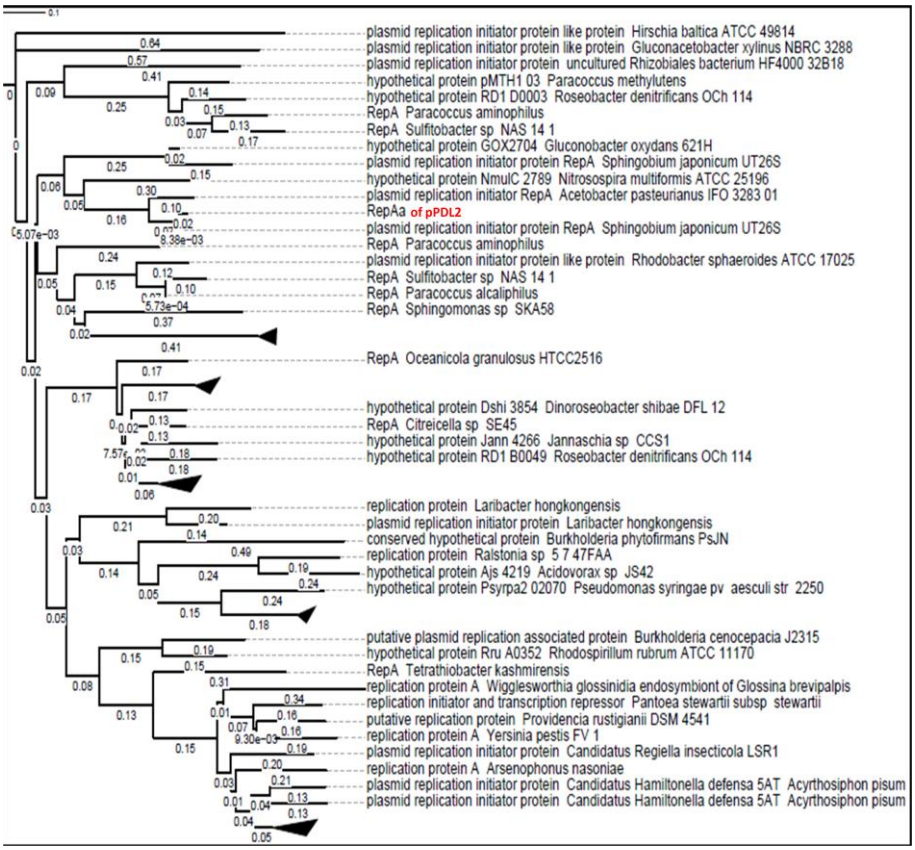


Fig. 3. 9. Phylogenetic analysis of RepAa of pPDL2 of *Flavobacterium* sp. ATCC 27551.

3.7.2. Origin of replication (*oriV*)

A 776 bp long putative *oriV* like sequence was seen immediately upstream of *repAa, spanning from sequence position* 17238 to 18114. In general the replicative origin is predicted based on the GC-skew analysis as described in materials and methods section (Grigoriev, 1998). As RepA protein present on plasmid pUT1 of *Spingobium japonicum* was 96 % identical with RepA protein of pPDL2, a thorough search was done to find out a sequence that has similarity with the replicative origin of plasmid pUT1 of *Sphingobium japonicum* UT26S. As shown in Fig. 3.8B considerable similarity was found between *oriV* sequences of pPDL2 and pUT1. Only very minor differences were noticed

between these two *oriV* sequences. The *oriV* predicted in pPDL2 has two imperfect tandem repeats. The first tandem repeat consists of 24bp sequence designated as A-24 and A'-27bp. There is a gap of 107 bp sequence between A-24 and A'27. The second tandem repeat is C-48 (48 bp) and C'-54 (54 bp) in size and are separated by a gap of 26 bp. In between these two tandem repeats a 17bp long palindromic sequence was identified (Fig. 3.10B). Interestingly, sequence found downstream of palindromic sequence is AT rich, while the upstream region is highly GC rich. In the upstream region of the palindromic sequence four typical DnaA binding boxes were identified with a consensus sequence of 5'-TTN4ACA-3' (Fig. 3.10C). The replicative origins of plasmids are shown to have configuration conservation (spatial arrangement of repeats) rather than showing strict sequence conservation (Gloria del Solar et al, 1988). When searched to find such configuration, homologues to *oriV* of plasmid pPDL2 the *oriV* of plasmid pSC01 isolated from *Pseudomonas* has revealed to have similar spatial arrangements of repeats and palindromic sequences (Fig. 3.10D). However, neither plasmid pUT1 nor pSC01 are assigned any compatibility group. Therefore, with the present data no clear incompatibility group can be assigned to the plasmid pPDL2. As shown in Fig. 3. 10A, the other important proteins that contribute for plasmid replication and maintenance are RepB, ParA and RelB. The *orf40* has shown 83% homology to replication initiator protein, RepB of *Gluconobacter diazotrophicus* PA15 (Fig. 3. 11A). RepB proteins are only seen in plasmids that replicate through rolling circle (RC) mode. The RC mode of DNA replication

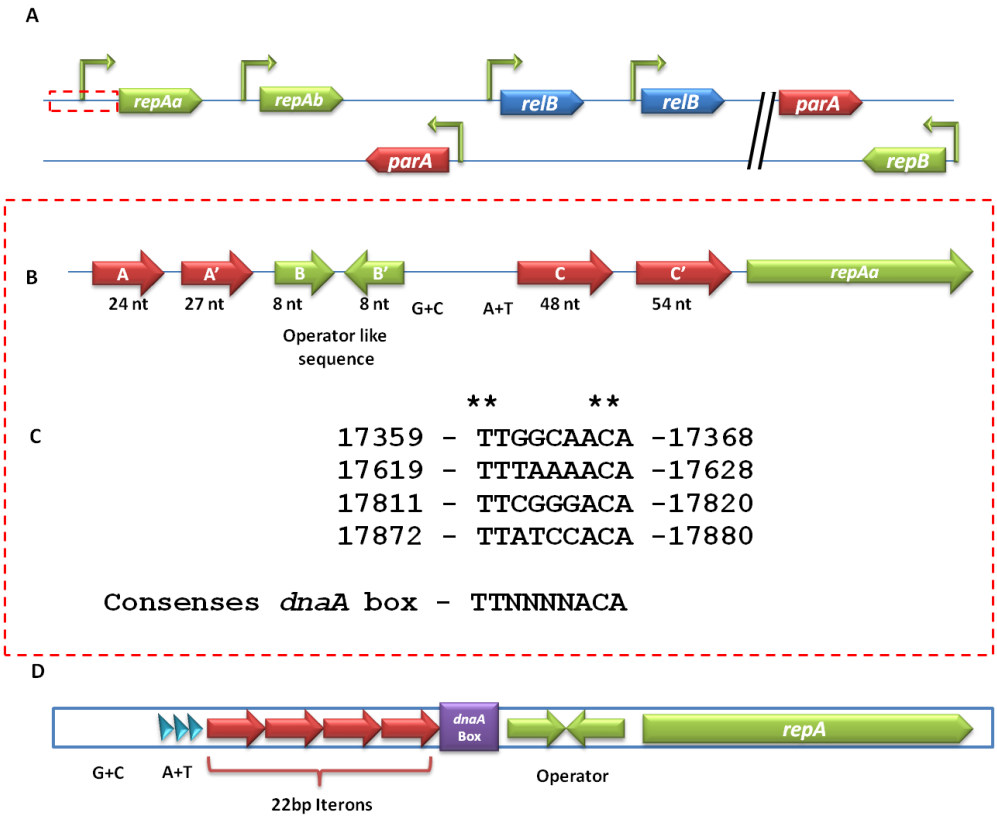


Fig. 3.10.Panel A shows physical map of plasmid pPDL2 DNA region showing the organization of ORFs coding for proteins involved in plasmid replication (*repA*, *repB*), partition (*parA*) and maintenance (*relB*). Panel B shows structural configuration of *oriV* found upstream of *repAa* in pPDL2 of *Flavobacterium* sp. ATCC 27551. In Panel B, the organization of tandem repeats A and A' found upstream of palindromic sequence (shown in green) also designated as operator sequence. The second tandem repeat sequence, C and C' found downstream of operator sequence is shown with maroon coloured arrows. The *repAa* is shown in green. Panel C. represents the alignment of predicted *dnaA* boxes found in the replicative origin *oriV* of plasmid pPDL2. The consensus *dnaA* box is shown separately. The configuration of *oriV* of pPS10 isolated from *Pseudomonas savastanoi* that show similarity to *oriV* of pPDL2 is shown in panel D

is generally seen in small plasmids (less than 10kb). The second place where such origin of replication is seen during conjugation process. In such plasmids, the replication initiators sequences are designated as the Mob class of initiators (Ilyina and Koonin, 1992).

A)

```
pPDL            MNHATSPVNGGKAKVALDGDTALTLAQKGRGNPFDPANYGEIVKPGELVDIVELSPLTLA 60
GPAl            ------------------------------------MNFG-------------------- 4
```

```
SPCl            MRVAAALQAKGGDEFAKPGSIVEVKFVKGQSLSLTASRLLALMILTAG------------ 48
                                                 .

pPDL            DRRIYNLLIANAWERIGEPVIHRIPKSALKGTHQGNERIESSLLRLMGTIAIVTIRKG-G 119
GPAl            ---------------FNEPIVS---------YAGGNERIEASLLRLMGTIAIVTIRKD-G 39
SPCl            -------------GDAWEDRPHKMRKADIRRGHKGNERISDMLEELHRTLFAVDDKSWRG 95
                            *                 *****.  * .*  *:  *  :.  *

pPDL            KSFKRRVQLLGPSDESLEKD-----GFLHYRIPEELIEILRNSEVYARLKTQVMYCFESK 174
GPAl            KSYKRRVQLLGPSDESLEKD-----GFLHYRIPEELIEILQNSEVYARLKTQVMYCFESK 94
SPCl            KKATLRFSLISSSREEAEDEEGADAGWIEWEFTPEARKLIQESETYAVLNRQAVLGFRST 155
                *. . *..*:..* *. *.:     *::.:.:. *  :::::**.** *: *.:  *.*.

pPDL            YALCLYEMIERRIGLEYKQSEEFTIAELRGLLNVPEGKLERFADFNKYCLKVAQEEINKL 234
GPAl            YTLCLYEMTERRIGLEYKQTEEFTIEELRGLLNVPDGKLERFADLNKYCLKVATEEINKL 154
SPCl            YALKLYEIGALRLHRRQ-SLWKGDMTALRALLGIAPDVYKDFAQLRRKVLEKAKAEIDQL 214
                *:* ***:   *:  .  .  :  :  **.**.:. .  : **::.:  *: *  **::*

pPDL            CPFWVEFTPIKKGRKVERVSMMWLPKTMSGRRDAQNLIDQHSIVRRAKLRGDIPEMPVLV 294
GPAl            CPFYVDFSPIKKGRKVERIAFHWFPKTSSGKRDAQILIDQHRIVRRAKLRGLAAELPLLL 214
SPCl            AHFRVEWREIRQGRTVTEIEFRFEPKDAPAQIATVDEIGRHSAGRKARREDEVETVAVEA 274
                . * *::  *::**.* .: : : **  ..:  :   *.:*   *:*: ..    :.:

pPDL            DFSAPAAQR---------------------------------------------------- 303
GPAl            DFGTEPPEK---------------------------------------------------- 223
SPCl            VTQAAVAALVSKDKAGAGEVTFPNGTIRFGSDTLAAIGRSAGGGWDIDLIADAYRAQMGE 334
                   :  .
```

B)

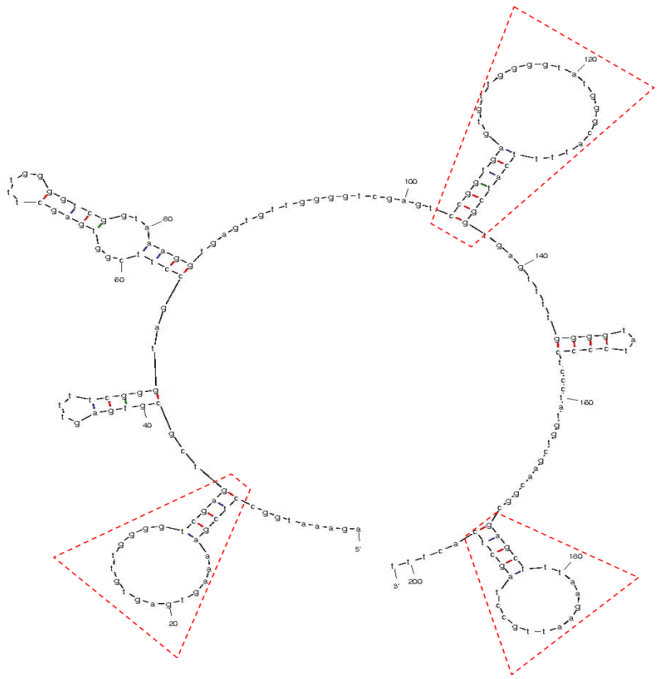


Fig. 3.11. Panel A shows multiple alignment of RepB of pPDL2 of *Flavobacterium* sp. ATCC 27551 with RepB of *Gluconacetobacter diazotrophicus* and *Sphingobium chlorophenolicum.* Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols. Panel B shows Secondary structure of putative rolling circle replication origin found in upstream of *repB* sequence of pPDL2. The hair-pin and stem loop structures found to be typical structural features of rolling circle replication origin are highlighted with red colour.

In general, the plasmids replicating through rolling circle replication usually have two origins, the double stranded origin of replication and a single stranded origin of replication with a potential to form a cruciform and a hair-pin like structures respectively (del Solar et al, 1998). In well characterized streptococcal rolling circle replication (RCR) plasmid pMV158, the RepB is a homo hexamer and usually binds to a region in the origin known as 'Bind region' which is characterized by presence of direct repeats (de la Campa et al, 1990; Ruiz-Maso´ et al, 2004). In pPDL2, a sequence that shows strong structural similarity to the consensus RC replication origin was identified immediately upstream of RepB. Like in typical RC replication origin this region contains 4 direct repeats each measuring a length of 22 bp with a consensus sequence of 5'-ACCCCAACACTCACCGGACTCG-3' and a sequence with a potential to form a secondary structure (Fig. 3.11B). Its typical structural features and strategic location upstream of RepB suggest a role in replication of plasmid pPDL2. RC mode of replication is only seen in small plasmids (del Solar et al, 1998). In the background of such reports, further experimentation is required to validate its involvement in replication pPDL2. RC mode of replication is also seen during horizontal mobility of plasmids through conjugation (Lanka and Wilkins, 1995). Therefore, its involvement along with RepB in generating a relaxosome required during conjugation process cannot be ruled out. Mobilizable nature of plasmid pPDL2 (described in chapter-2) adds strength to such proposal.

3.7.3. ParA locus

In addition to the replication initiator proteins, pPDL2 has an ORF designated as *orf*22 (21246-21893c) immediately downstream of *repAb* gene. The 215 amino acids long protein coded by *orf22* has shown 99% homology to the partitioning protein ParA of

Sphingobium japonicum UT26S (BAI99179) (Fig. 3.12). Therefore, *orf22* is henceforth designated as *parA* gene. In the light of such high degree homology with the well characterized ParA protein it appears to involve in partitioning of pPDL2 soon after its replication. Interestingly, there exists another *parA* homologue, *parAa* immediately upstream of *repB*. However, it has also shown 58% identity with cobyrinic acid ac-diamide synthase of *Thauera* sp. M21T (ACK55109) and 57% identity to ParA of *Laribacter hongkongensis* (ABC70161). With the existing information it is not possible to assign a confirmed role of this gene in the maintenance of pPDL2.

```
pPDL            MKVLAILSQKGGVGKRAPGTTLATCLAVAAEQAGKVAAIIDLDPQATASFWKDVRQLDTP 60
SJUT            MKVLAILSQKGGVGK----TTLATCLAVAAEQAGKVAAIIDLDPQATASFWKDVRQLDTP 56
                ***************    ****************************************

pPDL            AVASIQPVRLPAMLKACEDAGTDLVVIDGAAVARDVAYEAARQADFILIPTKTAVFDTMS 120
SJUT            AVASIQPVRLPAMLKACEDAGTDLVVIDGAAVARDVAYEAARQADFILIPTKTAVFDTMS 116
                ************************************************************

pPDL            MTHTLDVVRQLDRAFAVVLTFVPPQGQETGDAIQAVAELGATVCPVTIGNRKAFFRAQAA 180
SJUT            MTHTLDVVRQLDRAFAVVLTFVPPQGQETGDAIQAVAELGATVCPVTIGNRKAFFRAQAA 176
                ************************************************************

pPDL            GQAVQEFEPHGPAADEIHRLYEYTTIRLYNEAEAA 215
SJUT            GQAVQEFEPHGPAADEIHRLYEYTTIRLYNEAEAA 211
                ***********************************
```

Fig. 3.12 .Pairwise alignment of ParA of pPDL2 of *Flavobacterium* sp. ATCC 27551 (pPDL) with ParA of *Sphingomonas japonicum* UT26S (SJUT). Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

3.7.4. Toxin antitoxin module

Toxin-antitoxin systems (TA systems) increase the plasmid prevalence (number of plasmid containing cells/total number of cells) in growing bacteria populations by selectively eliminating daughter cells that did not inherit a plasmid copy during the process of cell division (Gerdes et al, 1986; Jaffe et al, 1985). This post-segregational killing mechanism relies on the differential stability of the toxin and antitoxin. Usually the toxin is stable which is rendered inactive by an unstable anti-toxin (Tsuchimoto et al,

1992; Van Melderen et al, 1994). In daughter bacteria devoid of a plasmid copy, because TA proteins are not replenished, the antitoxin pool rapidly decreases, freeing the stable toxin. These plasmid-free bacteria will eventually be killed by the deleterious activity of the toxin. Plasmid-encoded TA systems are also called addiction modules (Yarmolinsky, 1995) since this property renders the cell addicted to antitoxin production and therefore to the TA genes.

RelE and RelB is a toxin and antitoxin pair coded by the *relBE* toxin-antitoxin gene family which are discovered in *Escherichia coli* (Gotfredsen and Gerdes, 1998). The *relE* gene encodes a small (11-kDa) protein that is extremely toxic to bacterial cells, and the *relB* gene encodes an antitoxin of similar size that counteracts the cell killing activity of the RelE toxin (Gotfredsen and Gerdes, 1998; Grøndlund and Gerdes, 1999). Many *relBE* homologues have been identified in a broad range of both gram-negative and gram-positive bacteria and in archaea (Gotfredsen and Gerdes, 1998; Grøndlund and Gerdes, 1999).

3.7.5. RelB of plasmid pPDL2

Upstream of the *parA* gene, an ORF designated as *pilT* (33297-33722c) is identified. This ORF codes for a protein having 141 amino acid long PilT domain. The pilT domain of Orf38 shows 59% homology to PilT domain containing protein of *Sphingomonas wittichi* RW1. Monomers of PilT domain containing proteins have the ability to polymerise while forming Pilus fibre (Wall and Kaiser, 1999). Similarly, the PilT protein in its hexameric conformation is required for ATP-dependent retraction of the type IV pilus in gram-negative bacteria (Aukema et al, 2005). Retraction of type IV pili mediates intimate attachment and signalling to the host cells, surface motility, biofilm

formation (Chiang and Burrows, 2003; O'Toole and Kolter, 1998), natural transformation (Wolfgang et al, 1998; Whitchurch et al, 1994), and phage sensitivity. In the hexameric state, the ATPase activity of PilT, could actively promote dissociation of pilin monomers from the base of the pilus filament and thus has been shown to contribute to the pilin monomers pool observed within the cytoplasmic membrane (Morand et al, 2004). PilT could remove or inactivate a capping protein that prevents an energetically favourable retraction reaction. Alternatively, PilT could reverse the direction of the PilB motor, whose ATPase activity is required *in vivo* for the assembly of pilus filaments from pilin monomers (Turner et al, 1993). The exact function of the PilT domain in some proteins is unknown but this domain is present in some toxin proteins involved in bacterial plasmid stability such as the VapC (Francuski and Saenger, 2009; Robson et al, 2009). The exact nature of PilT domain containing protein of pPDL2 is not clear. However, in plasmid pPDL2, two ORFs, *orf20* (23492-23922) and *orf30* (26316-26582) encode for an antitoxin protein which shows 96% similarity to RelB, anti-toxin of *Sphingobium japonicum* UT26S. In the toxin - anti-toxin pair on pPDL2, absence of the toxin and the presence of pilT domain containing protein in plasmid pPDL2 suggests possible role of PilT domain containing protein as a toxin.

3.8. Mobilization module

3.8.1. Origin of Transfer

In Plasmid pPDL2, genes responsible for mating pair formation are not noticed. However on careful examination, genes responsible for initiation of its mobilization are identified on the plasmid. The initiation of transfer occurs through formation of relaxosome involving *oriT* and relaxase. The relaxase creates a nick at the *oriT*. In the

preceding sections existence of *repB* gene on plasmid pPDL2 is mentioned. As discussed before, the RepB creates a single stranded nick to initiate replication through RC mode of replication. In view of existence of RC replicative origin in its upstream region, it was implicated either in replication or in horizontal gene transfer (HGT) of pPDL2. RC mode of replication is seen in mobilizable plasmids and it is initiated from a sequence designated as *oriT* (Pansegrau and Lanka, 1991). The well characterized *oriTs* are characterized by a GC rich region with a stem loop structure (Lee and Grossman, 2007). Such GC rich region is present in the intergenic region of *repAa* and *repAb* (Fig. 3.13) This GC rich sequence (5'-AAAAAA**CCCCCCC**ACAAGAGA**GGGGGGG**GGGGG-3') has a potential to form a stem loop structure and possesses typical features of an *oriT* sequence (Fig. 3.13). Existence of RepB coding sequence along with putative *oriT* and RC replicative origin adds strength to the prediction of pPDL2 as a mobilizable plasmid. Such predictions have in fact formed basis for conducting experiments described in chapter II to show that plasmid pPDL2 is a mobilizable plasmid.

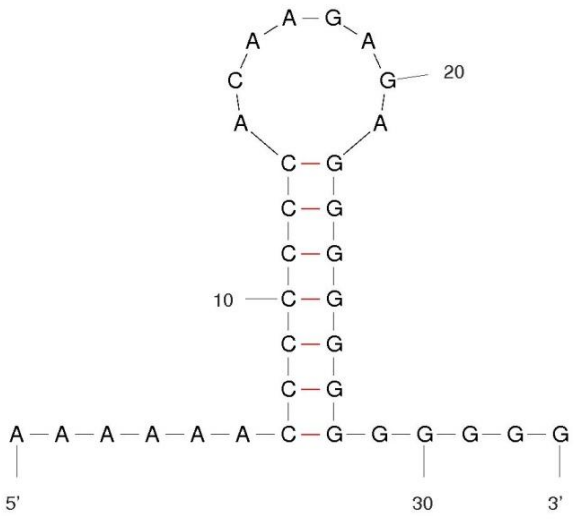


Fig. 3.13. Secondary structure of predicted *oriT* present in the intergeneic region of *repAa* and *repAb* in plasmid pPDL2.

3.9. Integrase module

An interesting observation in the entire sequence of pPDL2 is organization of integrase module. Integrase module contains two units, each consisting of *copG, int* and *pgm* genes. These units designated as CIP-I and CIP-II stretches from nucleotide position 26837 to 33015. Out of these three genes, both *copG* and *int* appears to be co-transcribed. The *copG* of CIP-I unit starts from 29962 nucleotide with a start codon ATG and ends with the presence of a stop codon present between nucleotide position 29564-29567. The stop codon of *copG* overlaps with the start codon of integrase. Integrase of CIP-I module is a 328 amino acid long protein and shows strong homology (86%) to phage integrase protein of *Sphingobium japonicum* UT26 and 70% homology to *Pseudomonas syringae* DC 3000 (Fig. 3.14).To gain further insights into the function of integrase, the protein coded by integrase

```
Int1            MNELAPLPPPPSSSALALPALVASADEAARLRFLEFFAVTIRNPHTRRAYMRAAGEFLAW 60
UT26            MNQLAPLPSP----ALVLPALIAAADERARLRFLEFFAVTIRNPHTRRAYARAAGDFLAW 56
SRW1            --------------------MIAAADDATRLRFLEFFAVTIRNPHTRRAYARAAGDFLAW 40
Bps             ------MNQIVRSSSATMPALVTAAGERAGVRFLEFFASAIRNPHTRRAYARAAEDFFAW 54
Psyr            ------------------------------MRFLEFFTANIRNPNTRRAYARATQEFLTW 30
                                               :******:  ****:***** **: :*::*

Tnt1            CEARGVASLAGVQPLHVAAWIEAQGGELAPPSVKQQLAGVRSLFDWLVMGQVVPANPAAS 120
UT26            CEARGVASLAGVQPLHVAAWVEALGRELAAPSVKQQLAGVRHLFDWLVTGHIVPVNPAGS 116
SRW1            CEARGVASLAGVQPLHVAAWVEALGRELAAPSVKQQLAGVRHLFDWLVTGHIVPVNPAGS 100
Bps             CANMGVTSIVAVQPLHVAAWVELQTQTLSAPTVKQRLAAIRHLFDWLVTGQVVPVNPAAS 114
Psyr            CQVVGVPSLTEVSPLHVATWIELQMQTLAAPSVKQRLAAIRHLFDWLVVGQVVPHNPAAS 90
                *   **.*:. *.*****:*:*     *:.*:***:**.:* ****** *::** ***.*

Int1            VRGPAYSQRRGKTPVLVPDEARHLLDTIDVATHAGLRDRALIGLMVYSFARIGAALAMRV 180
UT26            VRGPAHSQRRGKTPVLAPDEARRLLDSIDVITHAGLRDRALIGLMVYSFARIGAALAMRV 176
SRW1            VRGPAHSQRRGKTPVLAPDEARRLLDTIDVTTPAGLRDRALIGLMVYSFARIGAALAMRV 160
Bps             VRGPSHSSKVGKTPVLDATEARHLLDAIDVSTPAGLRDRALIALMVFSFARIGAALAMRV 174
Psyr            VRGPSHTSRTGKTPVLEPLEARQLLDSIDICTPAGLRDRALIALMVFSFARIGAALAMKV 150
                ****::::.: ****** . ***:***:**: * *********.***:***********:*

Int1            EDVFMQNRRLWVRLHEKGGKRHEMPCHHNLEDYLTAYIDGAACARIARGPLFRTIARGTG 240
UT26            EDVFMQNRRLWVRLHEKGGKRHEMPCHHNLEDYLSAYIDGCELREDRKGPLFRTIARGTK 236
SRW1            EDVFVQNRRLWVRLHEKGGKRHEMPCHHNLEHYLAEYLDGCELREDRKGPLFRTIARGTK 220
Bps             DDVYVQNRRLWVRLREKGGKRHEMPCHHTLEAYLHAYLDGTGLANESKGPLFRTIARGTG 234
Psyr            EDVYIQNQRLWVRLKEKGGKQHVMPCQHSLEAYLHAYLVETGIDNDPKGPLFRTIGRGTE 210
                :**::**:******:*****:* ***:*.** **  *:      .  :*******.***

Int1           QLSETPLPKPMLSRCAFAMVRRRAAAAGIGTAIGNHSFRATGITTYLKNGGTLETAATMA 300
```

```
UT26            RLSETPLPQAN----AFAMVRRRAGAAEIGTAIGNHSFRATGITTYLKNGGTLETAATMA 292
SRW1            RLSDTPLPQAN----AFAMVRRRAGAAEIGTAIGNHSFRATGITTYLKNGGTLETAATMA 276
Bps             QLSTTPLPQAN----AYAMVRRRAAAAGIATKIGNHTFRATGITAYLKNGGTIENAAAMA 290
Psyr            QLSVNALPQAN----AHAMVRRRALAAGIKTSIGNHTFRATGITAYLKNGGTLENAAAMA 266
                :** ..**:.     *.******* ** * * ****:*******:*******:*.**:**

Int1            NHSSTRTTQLYDRRPDDVTLDEVERVLIAPSQCPSRWCGGARRRPRSARRSAITVSARPG 360
UT26            NHSSTRTTQLYDRRPDDVTLDEVERVLI-------------------------------- 320
SRW1            NHSSTRTTQLYDRRPDDVTLDEVERVLI-------------------------------- 304
Bps             NHASTRTTQLYDRRRDDISLDEVERIRV-------------------------------- 318
Psyr            NHASTRTTQLYDRRRDEISLDEVERIRLDR------------------------------ 296
                **:*********** *:::******: :
```

Fig. 3.14. Comparison of integrase1 of CIP-I module with the integrases present in NCBI database. Integrase of pPDL2 (Int1) is compared with integrases of *Sphingobium japonicum* UT26S (UT26), *Sphingomonas wittichi* (SRW1), *Burkholderia pseudomallei* (Bps) and *Pseudomonas syrigyae* (Psyr). Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

gene of CIP-I module was aligned with similar sequences found in databases. Though the integrase has shown similarity to the entire stretch of tyrosine integrases, the C-terminus of the protein was found to be highly conserved than the central and N-terminal region (Fig. 3.15).

```
Prevotella                ISPHTLRHSFATELLKGGADLRAIQEMLGHESIGTTEIYTHIDISTLREEILNHHPRNIM 305
Alistipes                 ISPHTFRHSFATHLLEGGASIRQVQEMLGHESILTTEIYTHLEGDHLRDTVEKYLPL--- 298
Paenibacillus             ITPHTLRHSFAVHMLEGGADLRSVQEMLGHADLSTTQVYAQTARRNMKEVYEKHHPHGGN 307
Moorella                  ITPHTLRHSFATHLLENGADLRSVQELLGHADIGTTQIYTHLTRKKIREIYDHTHPRA-- 295
Desulfococcus             IKPHTLRHSFASHLLEGGADLRSVQIMLGHSDISTTQIYTHVTYRHLKDAHEKFHPR--- 297
Lawsonia                  ISPHTFRHTFATHLLEGGADLRSVQLLLGHVDMSATELYTHVQSDRLKYIHSMFHPRSNY 299
Chlorobium                ISPHTFRHTFATHLLEGGADLRAVQEMLGHSSISTTQIYTHIDRSFVKEVHKTFHPRG-- 304
Rhodopirellula            ISPHSLRHSFATHLLAGGADLRQVQEMLGHASIQTTQIYTHVEHSRLQRVHRDFHPRA-- 325
Bacillus                  ISPHVLRHTFATHLLDAGADLRAVQELLGHASLRSTQIYTHTTRERLLQVYLHAHPRA-- 304
Brevibacillus             VSPHTFRHTFATHMLNGGADLRTVQELLGHVNVSTTQVYTHVTKERLRHVYDTAHPRANP 304
Candidatus                VSPHTFRHSFATHLLDNGADLRAVQEFLGHSSLSTTQIYTHVTTERLKQVYDKTHPRA-- 298
Olsenella                 ITPHAMRHTYATELLSGGADLRSVQELLGHSSLSTTQIYTHLSVDRLKAAARQAHPRGE- 308
Atopobium                 LSPHAMRHTYATELLGGGADLRIVQELLGHESLSTTQVYTHLSVDRLKEAAKAAHPRSK- 305
Mycobacterium             IGPHGLRHSAATHLLEGGADLRIVQELLGHSTLATTQLYTHVTVARLRAVHDQAHPRA-- 300
Alkalilimnicola           VHPHMLRHSFASHLLESSGDLRAVQELLGHADIATTQVYTHLDFQHLARVYDQAHPRARK 298
Congregibacter            VHPHMLRHSFASHLLESSGDLRAVQELLGHSDISTTQIYTHLDFQHLAKVYDGSHPRARK 305
Conexibacter              VSPHALRHSFATHLLENGADLRSIQELLGHASISTTQVYTRVESARLRSAYANSHPRA-- 313
IntP of pPDL2             IGNHSFRATGITTYLKNGGTLETAATMANHSSTRTTQLYDRRPDDVTLDEVERVLI---- 328
Azoarcus                  IGNHSFRATGITEYLRNGGKLEIAQQMANHESARTTGLYDRRNDQLTLDEVERIVV---- 328
Polaromonas               MGPHALRATAATNALEHQADIAKVQEWLGHASISTTRVYDRRGSRPEDSPTFKVAY---- 331
Pseudomonas               LGVHGLRATAATNA---------------------------------------------- 279
Desulfobacterium          ITPHSARATFITQALENNCPIEAVQKTVGHAQIKTTQMYDKRTAKYRESASFAVRY---- 301
Frankia                   LSPHSLRATSVTLLLDAGASLRDAQDHARHADPRTTRAYDRARGSLDRAGTYQLVAYLDA 310
                          :  *  *  :
```

Fig. 3.15. Multiple alignment of C terminal region of integrase of pPDL2 with other integrases. The conserved motif H-X-X-R is boxed. Identical residues and gaps are

shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

Considering the similarity of integrase coded by pPDL2 and its near identity at the C-terminus especially in the region containing the H-X-X-R domain (Fig. 3.15), a proposal is made in this study to assign the status of integration mobilizable element (IME) to plasmid pPDL2. Existence of *oriT* like sequence and *repB*, involved in site-specific single stranded nick at *oriT* adds strength to this proposition.

The CopG/MetJ/Arc family regulatory protein (further referred as CopG) coded by pPDL2 shows highest similarity to the CopG regulatory protein coded by plasmid pUT1 of *Sphingobium japonicum* UT26S. The 131 amino acid long protein shows homology to a number of other proteins belonging to CopG/MetJ/Arc family of transcriptional regulators. The proteins of this family act both as transcriptional repressors and activators (del Solar and Espinosa, 1992; del Solar et al, 1995). In a well characterized streptococcal plasmid, pMV185, the CopG is shown to act as a repressor by regulating the expression of *repB* by binding to a pseudosymmetirc region present overlapping the -35 hexameric sequence of σ70 dependent promoter. Such binding is shown to prevent transcription from the *copG* promoter of the *copG –repB* genes which are co-transcribed in plasmid pMV158 (Farris et al, 2008).

In addition to these two transcriptionally coupled integrase and *copG* an additional *orf* that codes for a protein showing homology (98%) to *Sphingobium japonicum* UT26S phosphoglycerate mutase is identified in each unit of the integrase modules. The phosphoglycerate mutase gene in ICP-I module is found 1.1 kb away from the stop codon of integrase. This 216 amino acid long protein shows high homology (98%) to phosphoglycerate mutase present in *Sphingomonas japonicum* UT26S. When

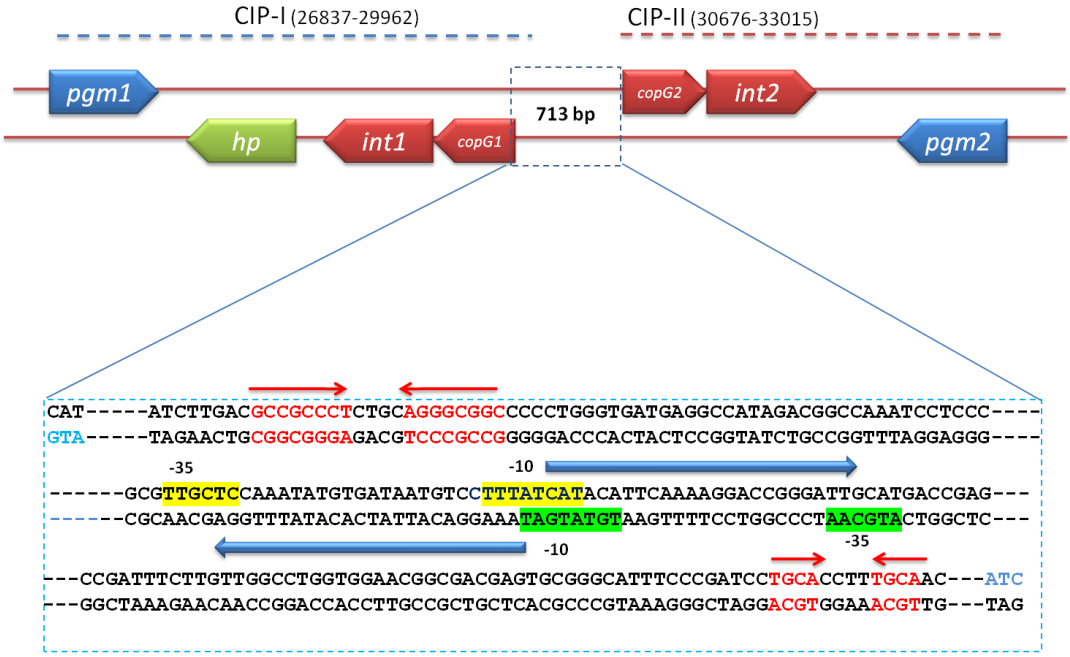
the sequence of phosphoglycerate mutase was analysed, it has shown existence of histidine phosphatase phosphoglycerate like domain (HP-PGM like domain) at the N-terminus of the protein. When examined the domain region has shown identity with similar a domain found in alpha-ribazole phosphatase (Fig. 3.16). Similarly, it has also shown similarity with the HP domain found in *Acinetobacter*.

```
PD1             -RGIRIDDAHAGPYSSATAKAPGNAGVPCHDLGAIELTELGQEQARE VAASWIEAPALIV 59
pUT1            MRAIFIRHGQS----------TGNAGVPCDDLGAIELTELGQEQARE VAASWTQAPALIV 50
                 *.* * ..::          .*******.********************** :******

PD1             TSPYTRTQQTAAPTIARFPGVPVEVWPIEEFTYLQPARWNGTRSAERMPHLERYWSAADP 119
pUT1            TSPYTRTRQTAAPTIARFPGVPVEVWPIEEFTYLQPARWNGTRSAERMPHLERYWSAADP 110
                *******:**************************************************

PD1             DYCDGEGAESFATLLRRCEAALARLAAMPVASPVYVFGHGQFIQAARAIVADAHMDDRAK 179
pUT1            DYCDGEGAESFATLLRRCEAALARLAAMPAASLVYVFGHGQFIQAARAIVADAHLDERAK 170
                ****************************.** **********************:*:***

PD1             MRAFWRKGEPPAIANAQRVGFHWEGGPLVLCAGAGGVMSAAPLIPTEDEIREHQAWDEFA 239
pUT1            MRAFWRKGEPPAIANAQRVGFHWEGG----------RWSCAP------------------ 202
                ************************           *.**

PD1             GGRPGAPRKGKIPLCDSPPPPQCGVVHLGAKDRWGAVGGKAE 281
pUT1            ------------------------------------ALAGTA- 208
                                                    *:.*.*
```

Fig. 3.16. Shows alignment of Phosphoglycerate mutase (PGM) coded by plasmid pPDL2 (PD1) and PGM (pUT1) coded by plasmid pUT1 of *Sphingobium japonicum* UT26S.

Phosphoglycerate mutases (PGMs) are very well characterized group of enzymes. Their existence is seen both in prokaryotes and eukaryotes. PGMs are basically transferases and are involved in transfer of phosphate group (Parkinson and Kofoid, 1992). Transfer of phosphates from 3rd position to 2nd position converting glyceraldehyde 3-phosphate to glyceraldehyde 2-phosphate is the classical biochemical reaction catalyzed by this group of enzymes. However, the PGMs containing HP-PGM like domain are also known to involve in signal transduction process (Matsubara and Mizuno, 2000). The histidine present in the catalytic site of this group of enzymes undergoes phosphorylation during the signal relay process (Parkinson and Kofoid, 1992). Strong link

between PGM and CopG and integrase genes suggests functional relevance among these proteins. Such organization is also seen in pCHQ1 (NC014007) and pUT1 (NC014005) of *Sphingobium japonicum* UT26S and pSWIT01 A)



B)

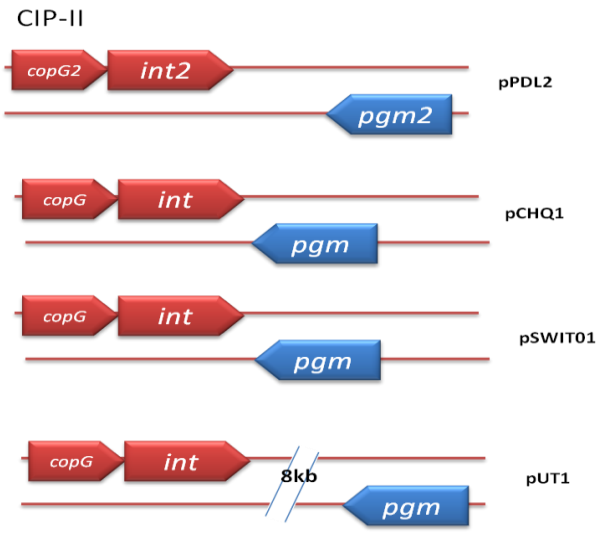


Fig. 3.17. Organization of Integrase modules in plasmid pPDL2. Panel A. The CIP units, CIPI and CIP-II are shown with dotted lines. The predicted regulatory region between CIP-I and CIP-II is boxed. The putative promoter elements identified upstream of *copGI* and *copGII* are highlighted with green and yellow colours respectively. The putative CopG

binding sites are shown with inverted arrows. Closed arrows indicate direction of transcription.
Panel B shows comparison of the organization of *cop*, *int* and *pgm* genes found in plasmids pPDL2, pCHQ1, pSWIT01 and pUTI.

of *Sphingomonas wittichi* RW1 (Fig. 3.17-B). Linkage of these three genes in plasmids isolated from different bacterial strains points towards having a functional relevance behind this conserved genetic organization. Further, studies are required to elicit the role of PGM in regulation of CopG-integrase operon expression.

As shown in Fig. 3.17-A, the integrase module contains a second copy of *copG*, *int* and *pgm*. We designate this region as CIP-II. This region also codes CopG, integrase and phosphoglycerate mutase as in CIP-I. As shown in unit-1, the *copG* and *int* are shown to be organized as one transcriptional unit. The stop codon of *copG* overlaps with the start codon of *int* gene. Such strong translational coupling suggests co-transcription of *copG* and *int* genes. CopG coded by CIP-II appears to have alternate start codon, ATC. Interestingly, the *copG* present in plasmid pCHQ1 of *Sphingobium japonicum* UT26S to which *copG* of pPDL2 shows high homology is also having an alternate start codon, ATC (AP010805). Like other proteins of plasmid pPDL2, CopG shows high homology to CopG coded by plasmids pCHQ1 and pUT1 of *Sphingobium japonicum* UT26S. CopGs coded by CIP-I and CIP-II are similar except that the CopG coded by CIP-II unit is using an alternate start codon, ATC (Fig. 3.18).

```
1.CopG1         ---------------------------------------------MHKDDDTAFADNYAER 16
2.CopG2         IWRRSRPPTPTHGAIATRCTTFGCGIAAILPRAFKRAGDPRPKGSPMTDDELPDDNPAER 60
                                                             **    ** ***

1.CopG1         DQARALREQARTGGLRFEAYLPGDMADWLLAQVEQGHFVDPSEAVFAIVKNFIEMEPHRD 76
2.CopG2         AQAKALREQARAGGLRFEAYLPGSMADWLLAQIERGRFADPSEAVFLIVQNFIEMEPHRD 120
                 **:*******:**********.********:*:*:*.******* **:**********

1.CopG1         LRDELLRRILDASVVRGLEDVKAGRVRPAEEVFDELRRKMAEPRPEPARWAKIAR 131
2.CopG2         LQDELLRRILQA---R-IDDPRPG--IPHDEACARIDRLLAEPRPDPARWEKIAR 169
                *:*********:*   * ::* :.*   * :*.  .: * :*****:**** ****
```

Fig. 3.18. Alignment of CopG1 and CopG2 sequences of integrase module of plasmid pPDL2. Pairwise alignment CopG1 and CopG2 coded by CIP-I and CIP-II units of integrase module are shown. Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

After analysing the ORFs of integrase module, the proteins coded by CIP-I were aligned to similar proteins coded by CIP-II. This comparison is done mainly to know if CIP-I is duplicated to give rise to CIP-II. If CIP-II has evolved through an event of gene duplication, the CopG, Int and PGM coded by CIP-I should have 100% sequence identity with their counterparts found in CIP-II. When CopG1 encoded by unit CIP-I is compared with CopG2 coded by CIP-II sequence, significant differences were noticed with respect to the primary sequence (Fig. 3.18). There exists only 67% identity between these two proteins. The N-terminal region of CopG2 encoded by CIP-II unit was found to be much longer (Fig. 3.18). Similarly, the C-terminus was not highly conserved. However, the central region of CopG1 and CopG2 proteins were found to be almost identical. Such diversity in the primary sequence indicates divergent origin of the CopG1 and CopG2 sequences. Similar situation was seen when *int1* and *pgm1* were aligned with *int2* and *pgm2*. Between *int1* and *int2* only 80% identity was seen (Fig. 3.19). However, the identity continued throughout the sequence. Interestingly,

```
1.Int-1_pPDL2          MNELAPLPPPPSSSALALPALVASADEAARLRFLEFFAVTIRNPHTRRAYMRAAGEFLAW
60
2.Int-2_pPDL2          MNQLAPLPSPASS--PALPALIAAADDDTRRRFLEFFAVTIRNPHTRRAYARAAGDFLAW
58
                     **:*****.*.**   *****:*:**: :* ****************** ****:****

1.Int-1_pPDL2          CEARGVASLAGVQPLHVAAWIEAQGGELAPPSVKQQLAGVRSLFDWLVMGQVVPANPAAS
120
2.Int-2_pPDL2          CEARGVASLAGVQPLHVAAWVEALGRELAAPSVKQQLAGVRHLFDWLVTGHIVPVNPAGS
118
                     ********************:** * ***.*********** ****** *::**.***.*

1.Int-1_pPDL2          VRGPAYSQRRGKTPVLVPDEARHLLDTIDVATHAGLRDRALIGLMVYSFARIGAALAMRV
180
2.Int-2_pPDL2          VRGPAHSQRRGKTPVLAPDEARRLLDSIDVTTHAGLRDRALIGLMVYSFARIGAALSMRV
178
                     *****:**********.*****:***:***:************************:***
```

```
1.Int-1_pPDL2        EDVFMQNRRLWVRLHEKGGKRHEMPCHHNLEDYLTAYIDGAACARIARGPLFRTIARGTG
240
2.Int-2_pPDL2        EDVFVQNRRLWVRLHEKGGKRHEMPCHHNLEHYLVEYIDGCGLREDRKGPLFRTIARGTK
238
                     ****:**************************.**. ****.. .  :**********

1.Int-1_pPDL2        QLSETPLPKPMLSRCAFAMVRRRAAAAGIGTAIGNHSFRATGITTYLKNGGTLETAATMA
300
2.Int-2_pPDL2        RLSDTPPAPSQCPSRWCGGARRRPRSA-----------RRSAITVSARPGSPPI------
281
                     :**:** . .  .    . .***. :*           * :.**.  : *..

1.Int-1_pPDL2        NHSSTRTTQLYDRRPDDVTLDEVERVLI 328
2.Int-2_pPDL2        ----------------------------
```

Fig. 3.19. Comparison of integrases coded by *int*-I and *int*-II genes of plasmid pPDL2. Pairwise alignment integrase-I of CIP-I and integrase-II of CIP-II modules is shown. Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

the percent identity between PGM1 and PGM2 is rather high. About 91% sequence identity was seen between these two proteins. The difference is only seen at C-terminus and N-terminus regions. About 10 amino acids gap was seen in PGM1 and PGM2 at both the N and C-terminal regions. PGM2 was found to be shorter by 21 amino acids due to existence of a 10 and 11 amino acids long gap at the N terminus and extreme C-terminus regions (Fig. 3.20). Considering the diversity in the primary sequence of the protein coded by CIP-I and CIP-II of the integrase module, the plasmid pPDL2 is proposed to have acquired the CIP-I and CIP-II integrase modules by independent sources possibly through unique recombination process.

```
1PGM1            -RGIRIDDAHAGPYSSATAKAPGNAGVPCHDLGAIELTELGQEQAREVAASWIEAPALIV 59
2PGM2            MRAIFIRHGES----------TGNAGVPCHDLATIELTERGQEQARAVAASWTEAPALIV 50
                  *.* * ...:           .**********.:***** ****** ***** *******

1PGM1                TSPYTRTQQTAAPTIARFPGVPVEVWPIEEFTYLQPARWNGTRSAERMPHLERYWSAADP
119
2PGM2                TSPYTRTRQTAAPTIARFPGVPVETWPIEEFTYLQPSRWNGTRSAERMPHLERYWSAADP
110
                     *******:****************.***********:**********************

1PGM1                DYCDGEGAESFATLLRRCEAALARLAAMPVASPVYVFGHGQFIQAARAIVADAHMDDRAK
179
2PGM2                DYCDGEGAESFGTLLRRCEAALARLAAMPADSLAYVFGHGQFIQAARAIVADAHMDDRAK
170
                     **********.****************. * .*************************

1PGM1            MRAFWRKGEPPAIANAQRVGFHWEGGRWSCAPALAA 215
```

```
2PGM2            MRAFWRKGEPPAIGNAQRVGFHWQGD---------- 196
                 ************.*********:*.
```

Fig. 3.20. Pairwise alignment phosphoglycerate mutase -I and phosphoglycerate mutases-II coded by of CIP-I and CIP-II units of integrase module. Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

The transcriptional organization of *copG* and *int* genes of CIP-I and CIP-II have revealed an interesting observation. As indicated before, in both CIP-I and CIP-II, these two genes are found to be co-transcriptional. In fact, in both the cases the translational stop codon overlaps of CopG with translational start codon of integrase. As shown in Fig. 3.17, *copG1* and *int1* of CIP-I and *copG2* and *int2* of CIP-II have opposite transcription orientation. Between these two transcriptional units there is a gap of 713 bp. When this gap region was analysed for identification of promoter and other *cis*-elements involved in regulation of gene expression, quite a few interesting features were noticed. As shown in Fig. 3.17 the gap region contained two independent putative promoter elements. The promoter element proposed to be involved in transcription of *copG1* and *Int1* operons has shown sufficient similarity to the consensus σ70 dependent promoter (Fig. 3.17 promoters are highlighted in green). Similarly, the promoter element predicted to be involved in transcription of *copG2* and *int2* operon was found to be similar to the consensus σ70 dependent promoter. There was an overlap of 5 bp between these putative *copG1* and *copG2* promoters. As indicated in earlier section CopG1 proteins are shown to be transcription repressors. As *copG* and *int* genes are co-trancriptional, the intergenic region found between the *copG1* and *copG2* operons were further analyzed to identify *cis*-elements that can serve as target sites for binding CopG proteins. In studies conducted by del Solar and his associates, CopG is shown to bind to a pseudosymmetric

cis-element present overlapping -35 region (del Solar et al, 2002). Inverted repeats are also shown to be putative binding sites of CopG in *Sulfolobus neozealandicus* (Greve et al, 2004). Interestingly, in CIP-I unit, a typical CopG binding site was predicted 276 bp upsteam of start codon of *copG1* (Fig. 3.17). Likewise in CIP-II an inverted repeat that can act as a potential CopG binding site was observed 45 bp upstream of translational start codon of *copG2* (Fig. 3.17). Existence of such CopG binding sites, if seen together with the CopG role as transcription repressor, the organization seen in integrase module of pPDL2 appears to be tailor made for regulation of integrase expression in *Flavobacterium* sp. ATCC 27551. Further work has to be done to gain better insights into the regulation of integrase expression.

3.9.1. The attachment (*attP*/*attB*) sites

Integrase is known to mediate integration of site-specific recombination between two conserved specific sites. In well studied bacteriophages λ, the phage integrase mediates site specific recombination between phage specific P site (*attP*) and bacterial chromosome specific B site (*attB*). These sites are called attachment sites. Conventionally, they are named as *attP* (P-phage) and *attB* (B-bacteria) sites. The site specific recombinase encoded by plasmid pPDL2 has high homology to tyrosine recombinases (Fig. 3.14 and 3.15). Tyrosine recombinases integrate target sequence at a specific site (*attB*), usually present at the 3' end of t-RNA genes (Williams, 2002). In the background of this information, the pPDL2 sequence was thoroughly analyzed to identify *attP* homologs. The bioinformatic searches have predicted existence of two such sites in plasmid pPDL2 and were designated as *attP*-I and *attP*-II. The *attP*-I is located between nucleotide position 15197 to 15186 and *attP*-II is found between nucleotide positions

37889 to 37878. The predicted *att* sites are of 12 bp long. When these two sites were aligned there was absolute identity between last 6bp of *attP*-I and first 6 bp of *attP*-II (Fig. 3.21). After identifying putative *attP* sites an attempt was made to identity *attB* sites. In order to identify *attB* in *Flavobacterium* sp. ATCC 27551 from where pPDL2 was isolated it is required to know complete genome sequence of the host organism. However, such information is not available for *Flavobacterium* sp. ATCC 27551. *Flavobacterium* sp. ATCC 27551 has recently been reclassified as *Sphingobium fuliginis* (Kawahara et al, 2010). A number of *Sphingobium* total genome sequences are available in the public domain as reported in the earlier sections. Plasmid pPDL2 has been shown to have considerable homology to plasmid pUT1 isolated from *Sphingobium japonicum* UT26S. The total genome sequence of *Sphingobium japonicum* UT26S is available in the public domain (www.**ncbi**.nlm.nih.gov/). The total genome sequence was used as input to identify putative *attB* sites that have homology to *attP* sites predicted in plasmid pPDL2.

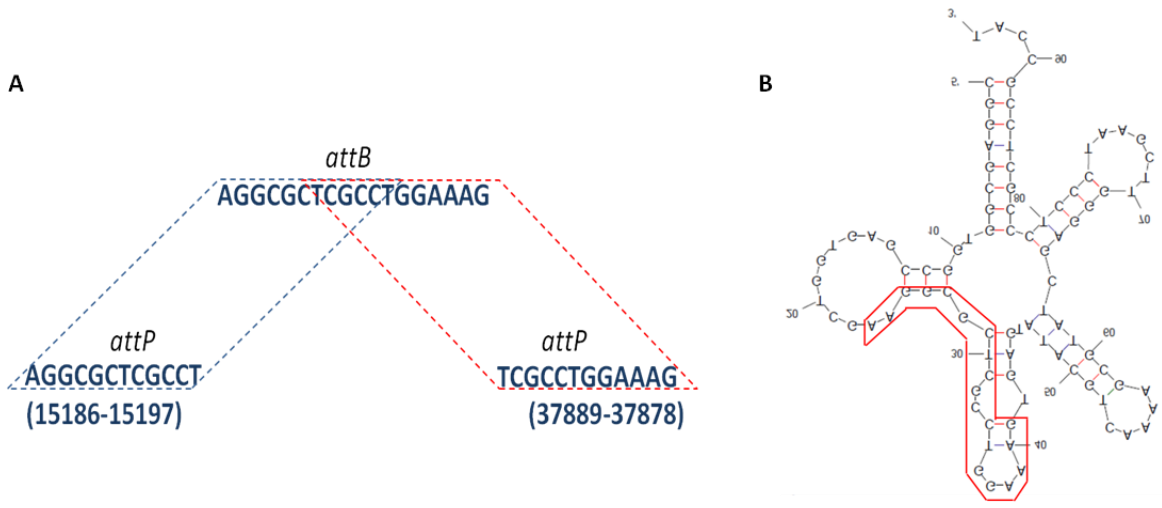


Fig. 3.21. Organization of *attP* sites on pPDL2 of *Flavobacterium* sp. ATCC 27551. Panel A shows the positions of *attP* sites on plasmid pPDL2. In Panel B the putative *attB* site on the tRNA-serine gene is outlined.

Interestingly, identical sequences were identified in the t-RNA gene sequences of *Sphingobium japonicum* UT26S (Fig. 3.21). If existence of *attP* site is taken together with the *attB* sequence in the *Sphingobium* genome, there exists an ample scope for plasmid pPDL2 to integrate into the genome.

3.10. Degradative module

As mentioned in aforementioned section, plasmid pPDL2 is associated with degradation of organophosphorus compounds. A well conserved parathion hydrolase coding *opd* gene has been shown as part of transposon-like element (Siddavattam et al, 2003). The sequence determined in this study agreed in total with sequence information reported from our laboratory. As indicated before the *opd* sequence is given a generic name as *orf11* (Table 3.3). A fine diagrammatic representation indicating the gene involved in degradation of organophosphorus compounds is made (Fig. 3.22). The degradation module includes parathion hydrolase or organophosphorus hydrolase coded by the *opd* gene, *meta* fission product hydrolase (*mfhA*), protocatechuate hydrolase (*pcaH*) and β-ketoadipate enol lactone hydrolase. Ample literature is available on catalytic properties, substrate range and membrane targeting of organophosphorus hydrolase (Karpouzas and Singh, 2006, Gorla et al, 2009).

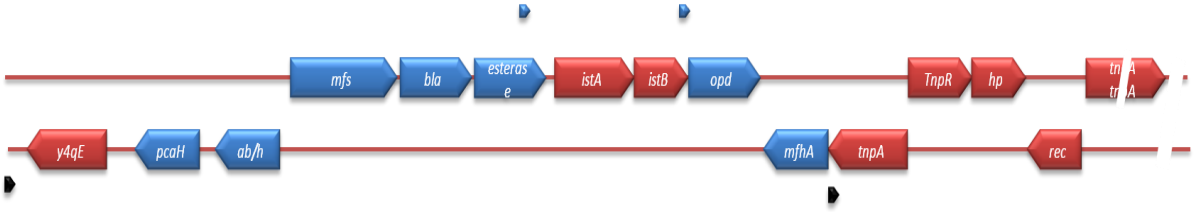


Fig. 3.22. Organization of degradative module in pPDL2. Transposases and resolvases are shown in red colour. Degradative genes are shown in blue coloured arrows. The

repeats of Tn*3* and IS*21* element are shown as black arrows and blue arrows respectively. Degradative reactions catalyzed by OPH, MfhA, P45O, b-keto adipate lactonase are shown using methyl parathion as model compound.

In our previous studies, we have we have reported existence of a *meta*-fission product hydrolase gene, *mfhA*, immediately downstream of *opd* gene (Khajamohiddin et al, 2006). The *mfhA* gene is also shown to be transcriptionally linked to the *tnpA* coding gene of transposon Tn*3* (Siddavattam et al, 2003). In this study, two more important genes that contribute to mineralization of organophosphorus compounds have been identified. Previous studies conducted by our lab has identified IS element belonging to IS*21* class upstream of *opd* gene (Siddavattam et al, 2003). When the complete sequence of pPDL2 was analyzed upstream of the IS element IS*Flsp1*, two more ORFs designated as *orf2* and *orf3* have been identified. These two ORFs were shown to code proteins that have significant homology to β- subunit of Protocatechuate 4, 5 dioxygenase and β-ketoadipate enol lactonase. Therefore these two ORFs were designated as *pca1* and *pca2*.

Organophosphorus hydrolase has been shown to hydrolase a variety of well known insecticides that are found to have aromatic moieties are linked to central phosphoric acid residue through an ester-linkage (Karpouzas and Singh, 2006). Upon OPH mediated hydrolytic cleavage most of OP-insecticides release aromatic compounds like 4-nitrophenol, 3-methyl 4-nitrophenol. These aromatic compounds are found to be much more toxic to the microbes than the parent OP compounds (Crbella et al, 2001). As shown in figure. 3. 22, existence of protocatechuate 4, 5 dioxygenase, β-ketoadipate enol lactonase in the upstream region of *opd* gene, flanked by transposase y4qE and transposase of Tn*3* suggests evolution of a well conserved degradation module to

mineralize OP insecticides. The dimethyl thiophosphoric acid generated from OP insecticides like methyl parathion, paraxon and parathion can be quickly utilized as source of carbon and sulphur. However, the aromatic compounds thus generated through OPH activity, require dioxygenase, β-ketoadipate enol lactonase and *meta* fission product hydrolase, to convert them as an intermediates of TCA cycle. Existence of β-ketoadipate enol lactonase and *meta*-fission product hydrolase and a dioxygenase coding region might contribute for such conversion. Though experimental evidence has been shown on the functions of *opd* gene and *mfhA* (Mulbry and Karns, 1989; Khajamohiddin et al, 2006), the role of these two genes *pca1* and *pca2* in mineralization of aromatic compounds need to be established by further experimentation. However, if degradation module is carefully examined, evolution of degradation traits in the form of a mobilizable element is very apparent.

3.10.1. **Protocatechuate 4, 5 dioxygenase (P45O)**

The *pca* designated with a generic name as *orf27* is present between the nucleotide positions 1185 to2067. The 277 amino acids long protein coded by *pca* has a domain belonging to extradiol dioxygenase 3B-like super-family and the complete sequence shows 79% identitiy to the Protocatechuate 4,5 dioxygenase identified in *Xanthomonas campestris* pv. *campestris.* Therefore, the *pca1* encoded protein is designated with a functional name as Protocatechuate 4,5 dioxygenase (P45O). The Protocatechuate 4,5 dioxygenase is the key enzyme in the benzoate degradation pathway (hydroxylation pathway) and 2,4 dichlorobenzoate degradation (Adriaens et al, 1989;). In addition to these two compounds the P45O has been shown to cleave many aromatic compounds (Fig. 3.23).

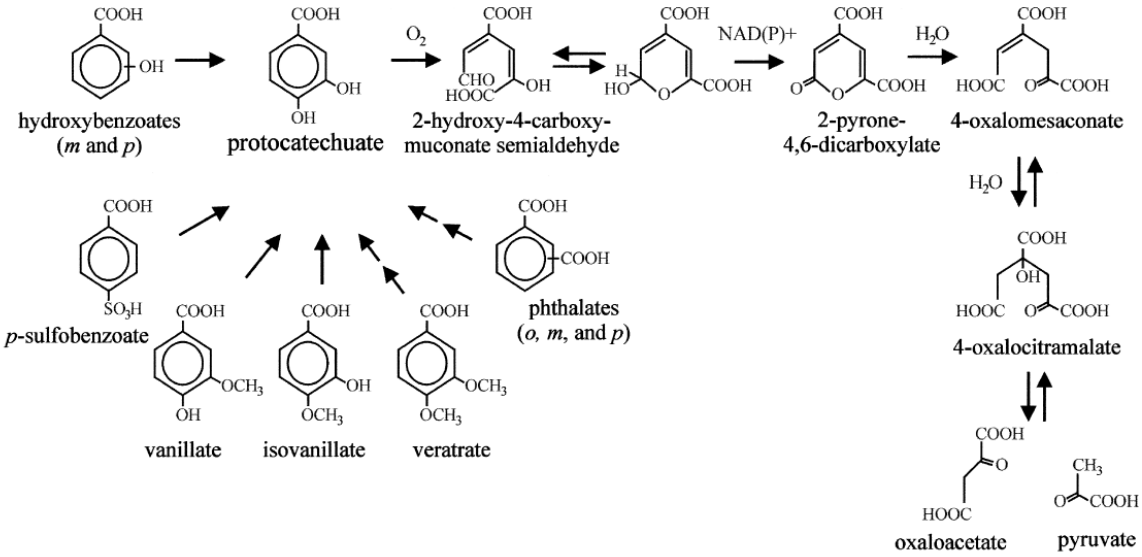


Fig. 3.23. Channelling of aromatic compounds into the TCA cycle through the Protocatechuate degradation pathway (Providenti et al, 2001).

Existence of such ring cleavage enzyme might contribute for conversion of 4-nitrophenol into intermediates of TCA cycle through a well known pathway known as protocatechuate degradation pathway. Degradation of aromatic compounds *via* protocatechuate pathway generates ring cleavage products with lactone ring. These lactones are further channelled into the TCA cycle through β-ketoadipate pathway (Khajamohiddin et al, 2008).The β-ketoadipate enol lactone hydrolase and Mfha coding *meta* fission product hydrolase contribute for such channelization. In degradation module along with OPH and P45O these two ORFs are present. The *orf311* coding protein shows 43% homology to the 3-oxo-adipate enol lactonase of *Xanthomonas campestris*. Its existence in the degradation module is yet another evidence to claim plasmid pPDL2 has evolved to mineralize OP compounds used as insecticides.

3.10.2. Major facilitator super-family protein

In the degradation module the next prominent ORF is *orf5* (3433-4887), designated as *orf484* encodes for a protein of 484 amino acids. The sequence of Orf484 shows homology to many of the aromatic acid transporters (Fig. 3.24). It shows 42% similarity with vanillate transporter of *Xanthomonas campesteris* pv. *campesteris.* Vanillate is shown to metabolize via protocatechuate degradation pathway (Fig. 3.23). As shown in Fig. 3.24 the transporter found immediately downstream of protocatechuate dioxygenase, is highly similar to number of aromatic acid transporters. Presence of similar transporters in close association of other genes that contribute for mineralization of organophosphates provides *prima facie* evidence to show existence of complete information on plasmid pPDL2 for mineralization of OP insecticides.

In Pseudomonads and in many Gram negative strains, genes are usually organized into operons and are co-ordinately regulated. Such organization is also seen in genes coding for enzymatic machinery involved in degradation of a number of aromatic compounds (Harayama and Rekik, 1990; Harwood and Parales, 1996; Yen and Serdar, 1988).

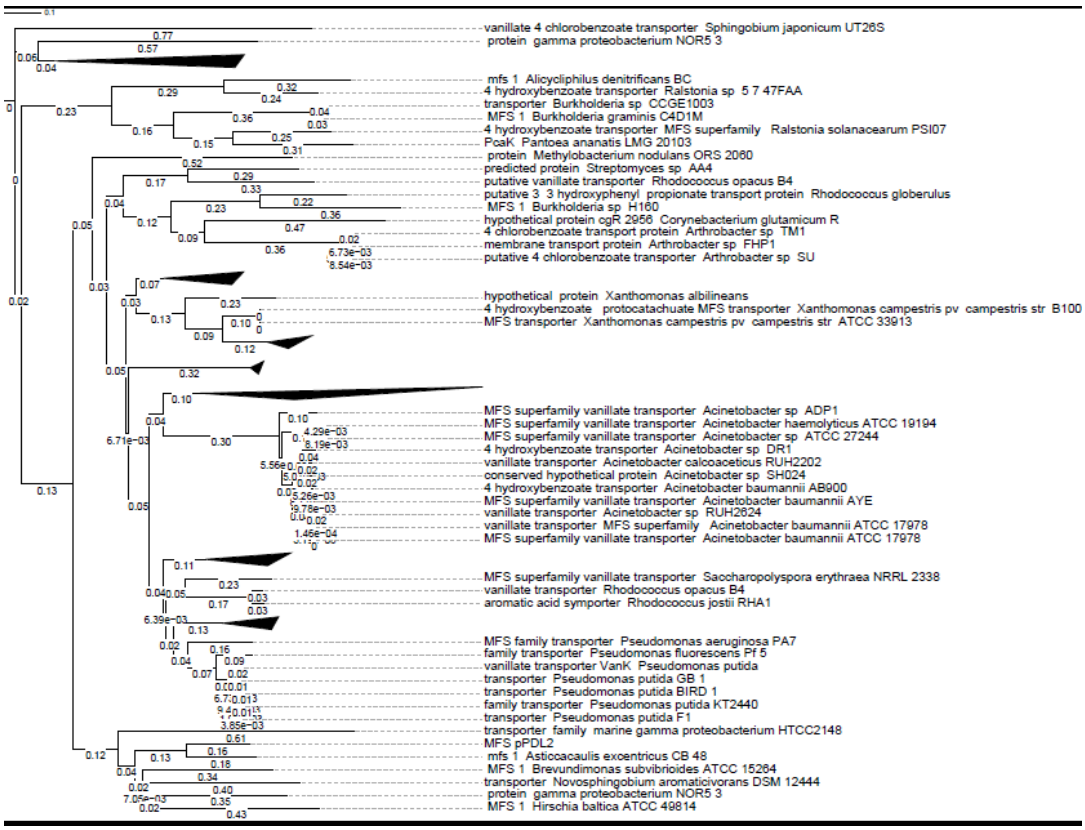


Fig. 3.24. Phylogenetic tree of Major Facilitator Superfamily (MFS) proteins. MFS of pPDL2 is seen in a clade that shows similarity with vanillate and other aromatic compounds

In the degradation module found in pPDL2 is compared with similar modules found in TOL and NAH plasmids the degradative genes on pPDL2 are dispersed. This kind of organization might be due to presence of transposons and recombinases causing rearrangement of genes. *Flavobacterium* sp. ATCC 27551 was reclassified as *Sphingobium fuliginis* based on chemotaxonomic and phylogenetic evidences (Kawahara, 2010). Dispersed organization of otherwise clustered genes is a common feature in Sphingomonads. Especially the genes coding for catabolic pathway enzymes in *Sphingomonas* strains are often found localized away from one another creating an unusual organization, where having a co-ordinately regulated operons are seen as a rare phenomenon (Basta, 2004). This has been described for the genes involved in the degradation of γ-hexachlorocyclohexane (lindane) by *S. paucimobilis* UT26 (Miyauchi et

al, 1998; Nagata, 1999), pentachlorophenol by *S. chlorophenolica* (Cai and Xun, 2002), protocatechuate by *S. paucimobilis* SYK-6 (Masai, 1999), naphthalene, biphenyl, and toluene by *S. yanoikuyae* B1 and *S. aromaticivorans* F199 (Romine et al, 1999; Zylstra and Kim, 1997), and dibenzo-*p*-dioxin by *S. wittichii* RW1 (Armengaud et al, 1998). Such dispersed organization, as seen in *opd* element of pPDL2 (Fig. 3.22.), may be a typical characteristic feature in Sphingomonads in which *Flavobacterium* sp. ATCC 27551 is placed according to the new classification (Kawahara et al, 2010).

3.11. Mobile genetic elements

In total 1 IS element and 3 transposons are present in the sequence of pPDL2. Existence of an IS element that shows homology to IS*21* class of Insertion elements is reported in our previous studies (Siddavattam et al, 2003) and this IS element, designated as ISFlsp*1* and was deposited in the IS database (http://www-is.biotoul.fr). This IS element is present in the upstream region of *opd* gene, spanning the nucleotide positions from 6941 to 9300. This IS element consists of two ORFs designated as *istA* and *istB* which encode for proteins of 507 and 279 amino acids, respectively. The stop codon of *istA* overlaps with start codon of *istB* which imply that the two genes, *istA* and *istB* are translationally coupled. The 507 amino acid protein designated as IstA, shows 59% identity with transposase ISMdi7 (IS*21* family) of *Methylobacterium extorquens* DM4 (YP_003065654) and 55% identity with IS*21* family transposase of *Agrobacterium tumefaciens* str. C58 (NP_355800) (Fig. 3. 26).

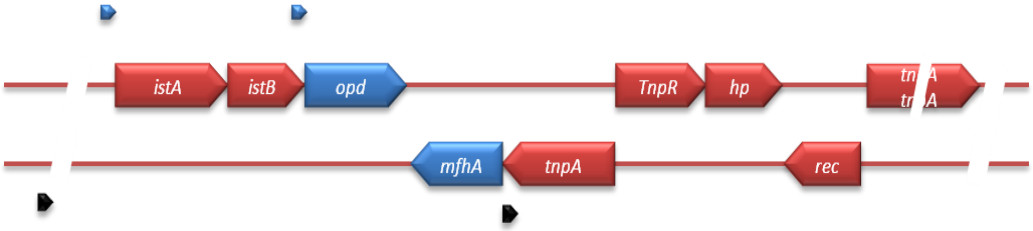


Fig. 3.25. Organization of mobile genetic elements on pPDL2. Transposases and resolvases are shown in red colour. Degradative genes are shown in blue coloured arrows. The repeats of Tn*3* and IS*21* element are shown with black and blue arrows respectively.

```
PD              MKSVEIYAKVRRAVLVEGMTRREAARYFGVHRNTITKMLQYAEPPGYRRAVPRVSEKLAP 60
Auma            MFAVEVYAAVRHFVLIERNSQREAARVFGLSRETVSKMCRFSLPPGYTRVKPVARPKLGA 60
Meex            MFVVEVYAAVRQFVFIEGQSRREAARVFGLSRETIAKMCRFSLPPGYTRSKPVEKPKLGP 60
                *  **:** **: *::*  ::***** **: *:*::** ::: **** *  *    **..

PD              FETLIDEILRSDKGAPPKQRHTCKRIYERLRTEHGYTGGLTILSDYVRSQRLRSREVFIP 120
Auma            LLPVIDWILEADGTAPVKQQHTAKRIFERLRDEHGYGGGLTVVKDYVRIARGRLRETFVP 120
Meex            LLPVIAAILEADRTAPLKQRHTAKRIFERLRDEHGYAGGYTVVKDHVRICRARGQETFVP 120
                : .:*  **.:*  ** **:**.***:**** **** ** *::.*:**  * * :*.*:*

PD              LSHRPGHAQVDFGEADAIIAGKRVRLHYFCMDLPQSDGCFVKAYPAEVAEAFCDGHVSAF 180
Auma            LAHSPGHAQVDFGEAIGVIGGVRQKIHFFCMDVPQSDAPFVKAYPRETTEAFLDGHVSAF 180
Meex            LAHPPGHAQVDFGEAVATIAGVRRKIHFFCMDLPHSDACFVKAYPRETTEAFLDGHVAAF 180
                *:* *********** . *.* * ::*:****:*:**. ****** *.:*** ****:**

PD              AFFGGVPTRILYDNTRLAVARILGDGRRERSRMFAGLQSHYLFDDRFGRPGKGNDKGKVE 240
Auma            DFFGKVPLSILYDNTTIAVARICGDGRRERTRAFTELQSHYLFADRFGRPGKGNDKGKVE 240
Meex            AFFGGVPLSILYDNTKIAVAKICGDGQRERTRAFTELVSHCLFRDRFGRPGRGNDKGKVE 240
                 *** **  ****** :***:* ***:***:* *: * ** ** *******:********

PD              GLVGYVRRNFMVPIPAAASIEELNARFADQCRRRGAAVLRGQSQSITARMEADSAAFMPL 300
Auma            GLVKYARSNFMTPIPQAASFDDLNAMLAERCRQRQGEVAGRHSETIGERLVADLEAFKDL 300
Meex            GLVKFARSHFMTPAPEAASFEALNADLERRCRARQNECAGRHPESIGTRLMADRVVLRAL 300
                *** :.* :**.* * ***:: *** :  :** *       :.::*  *: **  .:  *

PD              PEVAFDPCHIDSGCASSMALVRYRTNDYSVPTAFAHQQVVIKGYVDRVDIVCRGTCIASH 360
Auma            PATPLEPCEKRAARVSSTALVRYRCNDYSVPTSFGFRDVLVKGFVDEVVILCAGEEIARH 360
Meex            PAVPLEPCEKRAGRVSSTALVRYRGNDYSVPTTYGFRDVLVKGFVEEVVILCAGVEIARH 360
                * ..::**.  :. .** ****** *******::..::*::**:*:.* *:* *  ** *

PD              VRRYEREDFIANPLHYLALLEHKPGALDQAAPLDGWHLSEPVHRLRRLMEARSGKEGRRA 420
Auma            RRSYATGTFVFDPLHYLMLLEMKPNALDQAAPLQGWDLPETFQHLRHLLEARMGNRGKRE 420
Meex            PRSYGSGVFVAEPLHYLALIETKPNALDQAAALQGWDLPEAFQHLRHLLEARMGNRGKRE 420
                 * *    *: :***** *:* **.******.*:**.*.*..::**:*:*** *:.*:*

PD              FIQVLRLCEHYEQSLVEWAVARALELGAISFDAVKMILLARLEHRPARLDMSLYPYLPRA 480
Auma            FIQVLRLMEAMPMGIVAAAVTEAIRLGAIGFDAVKLIALSRIERRPLRLDLSRYPHLPKM 480
Meex            FIQVLRLMEAMPKDLVAWAVTEAIRLGAIGFDAVKLIALARLERRPPRLDLSAYPHLPRP 480
                ******* *    .:*  **:.*:.****.*****:* *:*:*:** ***:* **:**:

PD              NVGVTDTRAYLGLIPDAHRVTMKGASA 507
Auma            DVRTTAAADYAVLVPGKAA-------- 499
Meex            AVRATMAADYTVLVPEVAA-------- 499
                 * .* :  *  *:*
```

Fig. 3.26. Multiple alignment of IstA of pPDL2 of *Flavobacterium* sp. ATCC 27551 (PD) with similar proteins present in *Aurantimonas manganoxydans SI85-9A1*(Auma) and *Methylobacterium extorquens* DM4 (Meex). Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

Similarly IstB, shows 71% identity with resolvase IstB of *Mesorhizobium* sp. F28 (ABY59054) and *Sphingomonas wittichii* RW1 (YP_001260016). The IS*21* family transposons have a length of 2 kb to 2.5 kb and are found to be among the largest bacterial IS elements. They carry related terminal IRs whose lengths may vary between 11 to 50 bp generally terminating with a dinucleotide 5'-CA-3' (Mahillon and Chandler, 1998). The *istA* and *istB* of pPDL2 are flanked IR sequences of 15 bp present 100 bp upstream of *istA* (IRL) and 150 bp downstream of *istB* (IRR).

3.11.1. Tn*3* transposon

The sequence of pPDL2 found immediately downstream of *mfhA* gene showed existence of two copies of Tn*3* elements separated by an open reading frame coding for a hypothetical protein. These two copies of Tn*3* elements are designated as Tn*3*-I and Tn*3*-II (Fig. 3.25). These two Tn*3* like elements span from nucleotide position 11424 to 16077. The transposon Tn*3 is a* well characterized mobile genetic element found in a number of taxonomic groups. The Tn*3* transposon contains two genes, *tnpA* and *tnpR* and a 38-bp terminal inverted repeat at the left (IRL) and right terminus (IRR) (Heffron et al, 1979). The *tnpA* and *tnpR* code for transposase and resolvase. Transposition event takes place in two steps (Heffron et al, 1979; Shapiro, 1979; Grindley, 1983). The first step is formation

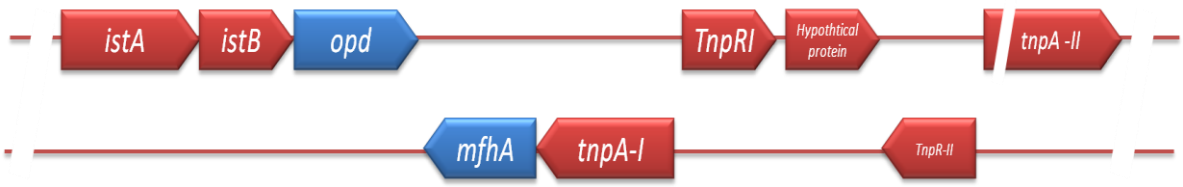


Fig. 3.27. Organization of transposon Tn*3* copies in plasmid pPDL2. The transpoase, resolvase coded by Tn*3*-I and Tn*3*-II are designated as *tnpa*-I *tnpA*-II, *tnpR*-I and *tnpR*-II respectively.

of cointegrate by the transposase, TnpA and the second step is resolution of the cointegrate into two separate replicons (Heffron et al, 1979; Gill et al, 1979; McCormick, et al, 1981). The resolvase catalyzes a site-specific recombination at the internal resolution site (IRS or *res*).

During the resolution process the cointegrate molecule is resolved to give recipient and donor replicons (McCormick, et al, 1981; Reed and Griendly, 1981). The TnpR protein is also a repressor that inhibits synthesis of both transposase and itself at the level of transcription (Gill et al, 1979; Chou et al, 1979).

3.11.1.1. Transposon Tn*3-I*

The Tn3-I spanning from nucleotide position 11424 to 138885 contains two oppositely transcribed open reading frames showing high homology to Tn3 family of transpoase (TnpA) and resolvase (TnpR) Therefore these two *orfs* are designated as *tnpA1* and *tnpR1,* which code for 583 amino acid long protein, TnpA1 and 189 amino acid long protein, TnpR1 respectively. TnpA1 of pPDL2 shows high similarity with TnpAs present in *Sphingomonas* and *Sphingobium,* especially to the TnpA sequence of (BAF03245) of *Sphingomonas* sp. KA1 (Fig. 3.27). Usually, full-length TnpAs coded by transposon Tn*3* are 985 amino acids. On comparison of TnpA1 of pPDL2 with that of the full-length transposases, a deletion of more than 400 amino acids is seen in the central region. The portion from amino acid 162 to 644 are found missing in TnpA-1 of pPDL2 (Fig. 3.27). The Tn*3* family transposases have a conserved domain which are typically seen in Transposase 7 super-family and is conserved in transposases coded by Tn*21*, Tn*1721*, Tn*2501*, Tn*3926* transposons. The domain contains an invariant triad, Asp689, Asp765, Glu895 (numbering as in Tn3) also reffered as D-D-35-E motif implicated in the catalysis of

transposition reactions by numerous transposases. Though there is deletion of more than 400 amino acids, the catalytic motif (D-D-35-E) is undisturbed as it is present at the C-terminus of Tn*3* transposase coded by *tnpA*-I of pPDL2 (Fig. 3.27). As substantial portion of TnpA is deleted, its functional status, despite of existing catalytic domain remains to be examined. In chapter-II, where transposition event is described an attempt is made to establish the functionality of TnpA.

```
TnpA1_PD      MTKRKHQLLTESERDQILAIPTDRDHLARLYTFEPSDIEIIGARRERRNQLGVALQLALL
TnpA_KA1      ------------MLAEHFDPSLDEREIARHFTLTRDDLELIASRRGDATRLGYAMLLLYL
TnpA_UT       --MARRRLVSLEIWAGHYDAPLDEREIARHYTLTSDDLEIVGRRRGDATRLGFAMLLLTM
TnpA_LB1      --MARRRLVSAEIWAGHYGAPLDEREIARHYTLTGDDLEIVGRRRGDATRLGYAMLLLYM
                                . *. .:** :*:  .*:*::. **   .:** *: *  :

TnpA1_PD      RHPGITLAQLIQDRGAIPHDLAAFVAEQLGLHVTELANYAARDQTMTDHVRELAARLGLR
TnpA_KA1      RWP----GRVLEAGEAPPMPILAFVARQLNVSPAAWRDYARRDETRRTHLADLSRRFGHG
TnpA_UT       RWP----GRALEAGEVPPAPVLGYVARQLGVAPDAFADYARRDQTRREHLVEIRRSHGFR
TnpA_LB1      RWP----GRALEAGEVPPAPVLAYVAQQLGVAPEAFADYAHRDQTRREHLVEIRRSHGFR
              * *    .: ::   . *  : .:**.**.:      :** **:*   *: ::    *

TnpA1_PD      GPTRADIPFMVEAAARTAWATDKGMTIAMGVVTALREARILLPS----------------
TnpA_KA1      AFGRADFHTLVAFAMPIAQTVTQPSRLAGIIMDEMRRRRLLLPPVTIIEAIVRRARQQAG
TnpA_UT       IFDRDAFREVVAFSIPIAQTIIHPGQMAGVIVDELRRRQILLPSSSILEAVLRRARQQAE
TnpA_LB1      IFDRKAFHEVVAFSIPIAQTIVHPGQMAGVIVDELRRRQILLPSSSVLEAVLRRARQQAE
                 *  :  :*  :   * :  :   :*  ::  :*. ::***.

TnpA1_PD      ------------------------------------------------------------
TnpA_KA1      DMIHDVLAGDLGEPERTRLDALLSRRDDKSATWLSWLRNPPLSPAPRNILRLIERLDHVR
TnpA_UT       QLTYEVLTNGLRPDTLQDLDDLLARRTGQAATWLSWLRNASQSPAARNILRLIERLAYVR
TnpA_LB1      QLTYEVLTNGLRPDTLQGLDDLLARRTGQAATWLSWLRNAPQSPAARNILRLIERLTHIR


TnpA1_PD      -IGIEPP--------------------------------------------------------
TnpA_KA1      TLGIAASRAATIPQAAFDRIADEAARITPQHLAELPDKRRHAILAAAGIRLEESLTDAVL
TnpA_UT       ALGLDRGRADMIPASTFDRLADEGSRITPQHLGELNALRRHATLAATGIRLEEDLTDATL
TnpA_LB1      ALDLDRARADMIPALTFDRLADEGSRITPQHLGELNALRRHATLAAQGIRLEESLTDATL
               :.:

TnpA1_PD      ------------------------------------------------------------
TnpA_KA1      TMMDKFLGSMMRRAENRTKEKAIGTIRSLQAQLRLITGSCRTLLDARARGVDSLAAIGSI
TnpA_UT       TMFDKLLGSMVRRAENRTRDKALKTVRELQGHLRTLTGSCRILIDARTNGVDSLAQIEAL
TnpA_LB1      TMFDKLLGSMSRRAENRTRDKALKTVRELQGHLRTLTGSCRILIEARTNGVDSLAQIEAL


TnpA1_PD      ------------------------------------------------------------
TnpA_KA1      DWERLGTAVVNAELLIAPETIDRTAELIERQRSLRSVIGPFLNAFEFRGAGAVQGLLDAA
TnpA_UT       DWQRFAVSVEQAEVLSRPETVDRTAELIERHRTVKLFAGAFLNTFEFRGAGAVQGLLSAL
TnpA_LB1      DWQRFAVAVARAEVLGRPETVDRTAELIERHRTVKLFAGAFLNTFEFRGAGAVQGLLSAL


TnpA1_PD      -----------------LPLAARPSIFR---------------------------------
TnpA_KA1      RLVADIYRTGRRRFPDKPPLRFVPPSWRPFVLRDGEVVRAAYELCVLTQLRDRLRGGDIW
TnpA_UT       AIIAELYRTGKRRLPDRVPLRFVPSAWRPFILRDGIVDRAAYELCALSQLRERLRAGDIW
TnpA_LB1      TIIAELYRTGKRRLPDRVPLRFVPSAWRPFVLRDGIVDRAAYELCALSQLRERLRAGDIW
                          **   *. :*

TnpA1_PD      ------------------------------------------------------------
TnpA_KA1      VAESRQYRAFDSYLLPPATFEAMRARGPLPLAIETDFDKFIAGRRASLDTALERVTILAR
TnpA_UT       VAGSRQFRDFDSYLIPPATYAALREKGPLPLAIETDFERHIEERRTRLDTAIEQVTVLAR
TnpA_LB1      VSGSRQFRDFDSYLIPPATFDALREKGPLPLAIETDFDRHIEERRARLDTAIEQVTVLAR


TnpA1_PD      -------------------RTAAPP-------------------------------------
TnpA_KA1      QGELPQVRLDGNGLVISPLKAITPPDAEDMRRVAYDRLPRVKITDLLLEVDSWTGFSECF
TnpA_UT       QGELPQVRLDENGLIISPLKAATPPATEIARRAAYDRLPRVKITDLLLEVDAWTGFSECF
```

Fig. 3.27. Multiple alignment TnpA1 of pDPL2 of *Flavobacterium* with TnpAs reported from *Sphingomonas* (TnpA KA1, TnpA LB1) and *Sphingobium japonicum* UT26S (TnpA UT1). Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols. The D-D-35-E motif is highlighted in yellow.

3.11.1.2. TnpR-I

Unlike TnpA-I the TnpR-I encoded by first copy of Tn3 (Tn3-1) shows high homology (99%) to resolvase of *Sphingomonas japonicum* UT26S (Fig. 3.29). Both these resolvases of Tn*3* show conserved features of the serine recombinase family such as the C-terminal DNA binding HTH motif and a conserved N-terminal catalytic domain having a serine residue in its active site (Fig. 3.30). As mentioned

before, resolvases are involved in resolution of co-integrates formed during transposition. However, in TnpA-I substantial portion is deleted formation of such co-integrate is questionable. In such a scenario the role of TnpR in transposition of *opd* cluster remains to be established.

3.11.2. Transposon Tn*3-II*

3.11.2.1. TnpA-II

As shown in figure the second copy to Tn*3* transposon starts at nucleotide position 15639 and ends at 16973nt. In between these two transposons, Tn*3*-1 and Tn*3*-2 a sequence that codes for an open reading frame of 315 amino acids is identified. The sequence of Orf315 has shown no homology to any other protein found in database. The TnpA coded by transposon copy II is designated as TnpA-II which has an opposite transcription orientation when compared to TnpR-II. Translated sequence of *tnpAII* when compared with full-length Tn*3* transposase has shown large deletions in the central region. The conserved catalytic motif having catalytic triad (D-D-E) is also not seen in TnpA-II. Therefore, the second transposase of Tn*3* family, TnpAII, present in pPDL2 is assumed as an inactive transposase (Fig. 3.28).

TnpRII (*orf16*) present in the second copy of Tn*3* spanning nucleotide postions14744 to 15325 is designated as *tnpR*-2 as it shows high homology (99%) to resolvase of *Sphingomonas japonicum* UT26S (Fig. 3.29). The resolvase designated as TnpR-II of Tn*3* show conserved features of the serine recombinase family such as the C-terminal DNA binding HTH motif and a conserved N-terminal catalytic domain having a serine residue in its active site (Fig. 3.30). Further, the sequence of TnpR-I and TnpR-II have shown high similarity to TnpR sequence found in transposon Tn3 elements

identified in *Sphingobium japonicum* UT26S. Considering the extensive homology found throughout the protein and

```
Tn3S_KA1        MLAEHFDPSLDEREIARHFTLTRDDLELIASRRGDATRLGYAMLLLYLRWPGRVLEAGEA 60
Tn3PD2          -------------------------------------------MLLLYMRWPGRALEAGEV 18
                                                           *****:*****.*****.

Tn3S_KA1        PPMPILAFVARQLNVSPAAWRDYARRDETRRTHLADLSRRFGHGAFGRADFHTLVAFAMP 120
Tn3PD2          PPAPVLAYVAQQLGVAPEAFADYAHRDQTRREHLVEIRRSHGFRIFDRKAFHEVVAFSIP 78
                ** *:**:**:**.*:* *: ***:**:*** **.:: * .*.  *.*  ** :***::*

Tn3S_KA1        IAQTVTQPSRLAGIIMDEMRRRRLLLPPVTIIEAIVRRARQQAGDMIHDVLAGDLGEPER 180
Tn3PD2          IAQTIVHPGQMAGVIVDELRRRQILLPSSSVLEAVLRRARQQAEQLTYEVLTNGLR---- 134
                ****:.:*.::**:*:**:***::***. :::**::******* :: ::**:..*

Tn3S_KA1        TRLDALLSRRDDKSATWLSWLRNPPLSPAPRNILRLIERLDHVRTLGIAASRAATIPQAA 240
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        FDRIADEAARITPQHLAELPDKRRHAILAAAGIRLEESLTDAVLTMMDKFLGSMMRRAEN 300
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        RTKEKAIGTIRSLQAQLRLITGSCRTLLDARARGVDSLAAIGSIDWERLGTAVVNAELLI 360
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        APETIDRTAELIERQRSLRSVIGPFLNAFEFRGAGAVQGLLDAARLVADIYRTGRRRFPD 420
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        KPPLRFVPPSWRPFVLRDGEVVRAAYELCVLTQLRDRLRGGDIWVAESRQYRAFDSYLLP 480
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        PATFEAMRARGPLPLAIETDFDKFIAGRRASLDTALERVTILARQGELPQVRLDGNGLVI 540
Tn3PD2          PDTLQACKLQG---------------------------------------------------- 145
                * *::* : :*

Tn3S_KA1        SPLKAITPPDAEDMRRVAYDRLPRVKITDLLLEVDSWTGFSECFTHRRSGRVADDRNALL 600
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        TVILADGINLGLTRMAETCQGATLRQLAHLHDWHISEAAYGEALGRLIDVHRTVPLSALW 660
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        GDGTTSSSDGQLFHSGGRGASIGDINARNGNEPGVSFYTHVSDQYDPFASRVIAATAGEA 720
Tn3PD2          ------------------------------------------------------------

Tn3S_KA1        PYVLDGLLYHATGLSIEEHYTDTGGASDHVFGLMPFFGYRFAPRLRDLKDRRLHLLPGQE 780
Tn3PD2          ----------------------------------MPFFGYRFAPRLRDIKQRRLHLLPGQE 172
                                                 **************:*:**********

Tn3S_KA1        AGPLLAGMTGDAVAIGHVADHWDELLRLTTSIRSGTTTASAILRRLSAYPRQNGLALALR 840
Tn3PD2          AGPLLAGMTAEPIALGHVAAHWDELLRFATSIRTGTATASAMLRRLSAYPRQNGLALAMR 232
                *********.:.:*:**** *******:*****:**:*****:****************:*

Tn3S_KA1        EVGRVERSIFMLDWLRDLDLRRRTQAGLNKGEARNALARALFFNQLGELRDRRFENQAYR 900
Tn3PD2          ELGRLERSIFMLDWLRDIDLRRRTQAGLNKGEARNALARALFFNQLGELRDRRFENQTYR 292
                *:**:************:****************************************:**

Tn3S_KA1        ASGLNLLVAAIILWNTRYLEQAVGTLSIPGNIARHIAPLGWEHISLTGDYRWNVESRPDP 960
Tn3PD2          ASGLNLLVAAIILWNTRYLEVALADIGTPDEIARHVAPLGWEHISLTGDYSWNVEDRPDP 352
                ******************** *:. :. *.:****:*************** ****.****

Tn3S_KA1        GK-----------LRPLRTPSSLLAA----- 975
Tn3PD2          GCPAATARHQFVARRVTFTIRSRLACGCVTF 383
                *            *   *  * **.
```

Fig. 3.28. Pairwise alignment of Transposase TnpA-2of pPDL2 of *Flavobacterium* sp. ATCC 27551 is aligned with similar ssequences coded by *Sphingomonas* sp. _KA1. Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

important catalytic site residues (Fig 3.29 and 3.30), the TnpR sequences coded by these two copies of Tn*3* elements can certainly be considered functional. However, in the absence of functional transposases, their presence in plasmid pPDL2 makes no functional significance. However, if transposition event occurs due to existence of a functional TnpA coding Tn*3* element located either in the chromosome or another indigenous plasmid, the resolvases found on plasmid pPDL2 can act to resolve the generated cointegrate.

```
TnpR2           --MWRSAMRASRRQTKKALTPARRAERGPGATRIFDDHASGAKADRPGLAEALAYLRSGD 58
TnpR_Sj_UT      -MALIGYARVSTADQKLSLQLD--ALNAAGCDRIFDDHASGAKADRPGLAEALAYLRSGD 57
TnpR1           MGGILGYARVSTGDQDVAGQTMR--LENAGAIKVFTDVISGKSMERPGLAELIAYARKGD 58
                      .  *.*  : . :         . .*. ::* *  ** . :****** :** *.**

TnpR2           TLVVWKLDRLGRSMSHLIEKVGELATRGIGFRSLTENIDTTTSGGMLVFNIFGSLAQFER 118
TnpR_Sj_UT      TLVVWKLDRLGRSMSHLIEKVGELATRGIGFRSLTENIDTTTSGGMLVFNIFGSLAQFER 117
TnpR1           TLAVVRLDRLGRSLTELLATVETLRSQGIALLSLEEKIDTSSAAGELIFHVFGAIAHFER 118
                **.* :*******::.*: .*  * ::**.: ** *:***::::.* *:*::**::*:***

TnpR2           DLIRERTHAGLKAARERGRPGGRRPVVTPDKLRKAREHIASGLTVREAAARLKIGKTALY 178
TnpR_Sj_UT      DLIRERTHAGLKAARERGRPGGRRPVVTPDKLRKAREHIASGLTVREAAARLKIGKTALY 177
TnpR1           RLISERTRDGIAAARAKGKQPGRQPLDMS-KVDAAIKLVEARISPTEAARQLGIGRSTIY 177
                 ** ***: *: *** :*:  **:*:  . *:  * : : : ::  *** :* **::::*

TnpR2           KALEATEKNTKSQRSRSVRSRADK 202
TnpR_Sj_UT      KALEATEKNTKSQRSRSVRS---- 197
TnpR1           REMRRLGVERPA------------ 189
```

Fig. 3.29. Multiple alignments of resolvases of Tn*3* transposon. The resolvase of TnpRI and TnpRII of plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551 are compared with TnpR found in *Sphingobium japonicum* UT26S (TnpR Sj UT). Identical residues and gaps are shown with an (*) and (-) respectively. Similar residues are shown with (:) and (.) symbols.

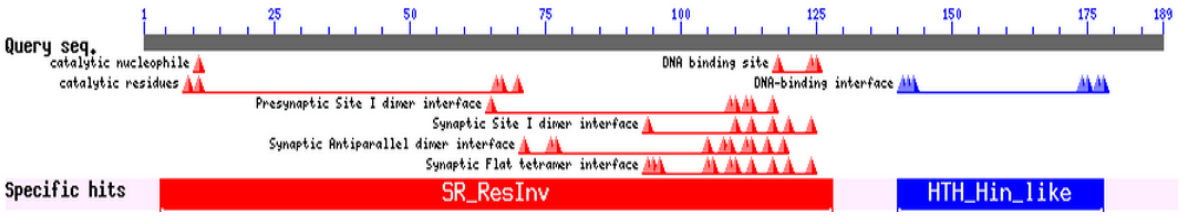


Fig. 3.30. Resolvases of pPDL2 showing N-terminal catalytic domain and C-terminal DNA binding HTH motif.

3.11.3. Transposon Tn3 specific terminal repeats

In pPDL2, two direct repeats containing characteristic features of Tn*3* family of transposons are present at positions 11354-1138 and 37419-37453. As mentioned earlier, transposon Tn*3* is present in many strains of the *Sphingomonas* sps. Analysis of their terminal repeats has showed minor variation in their length (Fig. 3.31). In plasmid pSY3

```
pPDL2RT         GGGGTCACTACACGAAAGTGCATTTTACGTACGCT---------------
35
pPDL2LT         GGGGTCACTACACGAAAGTGCATTTTACGTACGCT---------------
35
pCAR3           GGGGTCACTACACGAAAGTGCATTTTACGTACGCT---------------
35
PCAR3.2         ---------ACACGAAAGTGCATTTTACGTACGCT---------------
26
pSY3            --------------GGGGT-CACTACACGAAAGTGCATTTTACGTACGCT
35
Ibu-2           --------------GGGGT-CACTACACGAAAGTGCATTTTACGTACG--
33
UT26s           --------------GGGGT-CACTACACGAAAGTG---------------
20
pCHQ1           --------------GGGGT-CACTACACGAAAGTG---------------
20
                                    ** ** *  *** * *
```

Fig. 3.31. Terminal repeats of Tn*3* transposon found on plasmid pPDL2 (pPDL2RT and pPDL2LT) of *Flavobacterium* sp. ATCC 27551, plasmid of *Sphingomonas* sp KA1 (pCAR3), *Sphingomonas chungbukense* strain DJ77 (pSY3), *Sphingomonas* sp. Ibu-2 and *Sphingobium japonicum* chromosome UT26S (UT26S) and plasmid (pCHQ1).

isolated form *Sphingomonas chengbukense, a* 35bp terminal repeat was seen with 4 mismatches. The Tn*3* is also present both on the chromosome and plasmid pLB1 of *Sphingobium japonicum* UT26S. These are only 20 bp long. Terminal repeats in the Tn*3* element found in *Sphingomonas* sp., and *Sphingobium japonicum* UT26Ss suggesting that the first twenty base pairs of the repeats are enough for successful transposition of transposon Tn*3*.

3.11.4. The y4qE element

The ORF42 present at the region spanning the nucleotide positions 37745 to nucleotide position 38847 is designated as *y4qE* as it codes for a 374 amino acid long protein that shows 59% identity with the putative transposase y4qE (EFO28627) of *Roseibium* sp. *Trich*SKD4 and 57% identity with IS*116*/IS*110*/IS*902* family protein (ABS70259) of *Xanthobacter autotrophicus*. The IS*110* family elements are usually flanked by 51 bp terminal repeats (García-Trigueros et al, 2007). However, no repeats of 51 bp are noticed in the flanking region of y4qE transposase. In the absence of such terminal repeats the functional status of y4qE element is questionable. If the mobile elements found on pPDL2 are considered, Tn*3* alone appears to have TnpA with proper catalytic domain. The terminal repeats that are specific to Tn*3* are only seen flanking the degradative module. Experimentlly the transposition event needs to be determined to assess its status in transposition.

3.12. Discussion

The sequence of plasmid pPDL2 contains both features of a typical plasmid and Integration Mobilizable Element (IME). Existence of a double stranded replication origin (*oriV*) and a well defined *par* locus that is involved in maintenance of a plasmid together with a toxin anti-toxin domain perfectly justify assigning pPDL2 the status of a well defined plasmid. Historically pPDL2 was isolated as part of investigation aimed at understanding the molecular basis for degradation of organophosphate insecticides used in modern agricultural practices. Subsequent studies conducted on this plasmid have shown organophosphate degrading (*opd*) gene as part of a complex transposon-like element (Siddavattam et al, 2003). The present study is designed to understand horizontal mobility of *opd* gene, which is considered to be all most certain due to

existence of identical *opd* genes in bacterial strains isolated from bacterial strains belonging to diverse taxonomic groups. The complete sequence has indeed added an interesting dimension to genetics and biology of organophosphate degradation. In addition to the well defined replication and maintenance modules, the plasmid pPDL2 contained a well defined degradative module with a typical features of a complex transposon. The degradative module included organophosphate degrading (*opd*) gene, dioxygenase gene, and a *meta* fission product hydrolase gene and a β-ketoadipate hydrolase coding sequence. As shown in figure 3.25 transposon Tn*3* element and y4qE element are found flanking this degradative module. The perfect IR sequences that serve as target sites to the Tn*3* coded transposase have identified at the extreme ends. If the arrangement is seen it clearly indicates that *opd* gene is part of a well structured mobile element designed to contribute for the lateral transfer of *opd* information among soil bacteria. If the plasmid maintenance and degradative modules are alone taken into consideration it clearly assigns pPDL2 the status of a typical bacterial plasmid. However, it also contains a well defined integration module, which includes an integrase, CopG and phosphoglycerate mutase. The transcriptional arrangement of these genes and presence of *attP* site that show similarity to the chromosomally located *attB* site found at the 3' end of serine tRNA gene perfectly justifies the capability of pPDL2 to integrate into and excise from the chromosome. If these features are seen together with existence of *oriT*, the pPDL2 should be given a status of an integrative mobilizable element (IME). Before presenting the structural status of pPDL2 a brief discussion is given on the Integrative Conjugative Elements (ICEs) so that a structural comparison can be made between ICEs and plasmid pPDL2.

The conjugation systems encoded by chromosome-borne mobile genetic elements (MGEs) were recently identified. Such elements are often referred to as integrative and conjugative elements (ICEs)(Burrus et al, 2002; Wozniak and Waldor, 2010). ICEs are self-transmissible MGEs that encode for conjugation machinery as well as intricate regulatory systems to control their excision from the chromosome and their conjugative transfer (Salyers et al, 1995; Osborn and Boltner, 2002; Burrus and Waldor, 2004). The ICEs encompass all self-transmissible integrative and conjugative mobile elements regardless of their mechanisms of integration or conjugation (Wozniak and Waldor, 2010). These include elements that are commonly characterized as conjugative transposons, which often integrate into the host chromosome with minimal sequence specificity and, consequently, are capable of both intracellular and intercellular transfer (Burrus and Waldor, 2004). The Tn*916* of *Enterococcus faecalis* (Velikonja et al, 1994) and CTnDOT in *Bacteriodes thetaiotaomicron* are the first known MGEs with ICE-like properties (Shoemaker et al, 2001). Certain chromosomal elements which are previously classified as genomic islands also have properties of ICEs. Xenobiotic island, ICE*clc*^{B13} of *Pseudomonas knackmussii* sp. strain B13 and symbiotic island (Ravatn et al, 1998), ICE*MlSym*^{R7A} of *Mesorhizobium loti* are examples of genomic islands which have ICE properties (Ramsay et al, 2006).

3.12.1. Structure and Function of ICEs

Integrative and conjugative elements (ICEs) typically have modular structures in which genes with related functions are clustered together (Mohd-Zain et al, 2004; Juhas et al, 2007; Roberts and Mullany, 2009; Wozniak and Waldor, 2009). All ICEs have three

simple, distinct functional modules designated as i) maintenance, ii) dissemination and iii) regulation modules.

i) Maintenance modules

All ICEs encode an integrase (*Int*) that enables their integration into the host chromosome. The process of integration requires integrase and no additional factors are required. But for excision to occur, additional factors, known as recombination directional factors are required along with integrase. Integrases determine the site and frequency of ICE excision. Moreover, regulation of *int* expression is one of the key means of controlling ICE transmission. Most of the known ICE integrases are members of the tyrosine recombinase family (Argos et al, 1986). The best studied recombinase of this family is the Int encoded by phage λ. The phage Integrase uses a topoisomerase I-like mechanism to promote site-specific recombination between identical or near identical sequences in the host chromosome (referred to as *attB* sites) and the phage chromosome (the *attP* site). The strand exchange reactions catalysed by Integrase do not require a high-energy cofactor such as ATP, and no sequence duplication or deletion results from recombination. Integration of ICEs usually occurs into specific sites known as primary sites of integration (Table 3. 3). However, integration may also occur at secondary sites in absence of the primary sites (Burrus and Waldor, 2003; Lee et al, 2007).

Table 3. 4. ICEs of various sources and their sites of insertion

ICE	Host	Size (kb)	Site of insertion	Phenotype	Reference
SXT	*Vibrio cholera*	99.5	*prfC*	Cm^R, Sm^R, SXT^R	Beaber et al, 2002

R391	*Providencia rettgeri*	89	*prfC*	Hg^R, Kn^R	Boltner et al, 2002
ICE*Bs1*	*Bacillus subtilis*	20	$tRNA^{Leu}$ gene	None Known	Burrus et al, 2002
PAPI-1	*Pseudomonas aeruginosa*	108	$tRNA^{Lys}$ gene	Virulence factors and regulation of biofilm formation	
ICE*clc*^{B13}	*Pseudomonas knackmussii*	105	$tRNA^{Gly}$ gene	3-chlorobenzoic acid degradation	Ravatn et al, 1998
ICE*MISym*^{R7A}	*Mesorizobium loti*	502	$tRNA^{Phe}$ gene	Symbiosis with Lotus corniculatus involving nodulation and nitrogen fixation	Sullivan et al,2002
ICE*Hin1056*	*Haemophilus influenza*	49.4	$tRNA^{Leu}$ gene	Tet^R, Cm^R, Amp^R	Mohd-Zain et al, 2004
pSAM2	*Streptomyces ambofaciens*	10.9	$tRNA^{Pro}$ gene	None Known	Pernodet et al,1984
Tn*916*	*Enterococcus faecalis*	18	AT-rich regions	Tet^R	Clewell et al, 1995
CtnDoT	*Bacteroides spp.*	65	GTANNTTTTGC	Tet^R, Erm^R	Cheng et al, 2000
TnGBS2	*Streptococcus agalactiae*	33.5	Intergenic regions upstream of RNA polymerase sigma A promoters	None known	Brochet et al, 2009

Tn*4371*	*Ralstonia sp. A5*	55	TTTTTCAT	Biphenyl degradation	Toussaint et al, 2003
Tn*5397*	*Clostridium difficile*	21	Single site	Tc^R	Wang et al, 2000
Tn5252	*Lactococcus lactis*	70	TTTTTG	Sucrose utilization and nisin synthesis	Vijayakumar et al, 1993
pRS01/sex factor	*Lactococcus lactis*	48.4	Single site	Tellurium resistance	Gasson et al, 1995
Tn5801	*Staphlyococcus aureus*	25.8	3' end of a gene encoding GMP synthase	Tc^R	Kuroda et al, 2001

ii) Dissemination modules

Like conjugative plasmids, ICEs disseminate *via* conjugation. ICEs contain the genes that specify the synthesis of the 'mating machinery' that enables intimate contact between donor and recipient cells and that delivers DNA to the recipient cell. The dissemination modules of ICEs are also diverse. In most described cases, ICEs are thought to transfer as single-stranded DNA with few exceptions (pSAM2 from *Streptomyces mbofaciens)* (Grohmann et al, 2003). In some cases, such as in the *V. cholerae*-derived ICE SXT, the transfer genes bear similarity to those found in well-characterized conjugative plasmids such as the F plasmid (Beaber et al, 2003). In other cases, such as in Tn*1549* from *Enterococcus* spp., ICE transfer genes appear to be distantly related to those found in Gram-positive conjugative plasmids (Garnier et al, 2000). Finally, the genes required for the transfer of some ICEs, such as the Bacteroides-derived CTnDOT, are for the most part unrelated to previously characterized transfer genes (Bonheyo, 2001).

iii) Regulation modules

The genes and the mechanisms that regulate ICE transfer are just beginning to be defined; however, it is already apparent that ICE regulation modules are very diverse (Wozniak and Waldor, 2010). Transfer of both Tn*916* and CTnDOT is induced by sub-inhibitory concentrations of tetracycline (Slayers et al, 1995). The expression of the *int* gene of the *clc* element from *Pseudomonas* sp. was recently found to be stimulated by growing the bacterium on 3-chlorobenzoate-containing medium, but not by high cell density, heat shock, osmotic shock, UV irradiation or ethanol stress (Sentchilo et al, 2003). Interestingly, aromatic chlorinated compounds, such as 3-chlorobenzoate, are substrates for the degradation pathway encoded by *clc*. It was hypothesized that a metabolite of 3-chlorobenzoate modulates interactions of a putative activator and repressor with the integrase-encoding gene promoter (Sentchilo et al, 2003). Thus for three ICEs, Tn*916*, CTnDOT and *clc*, specific compounds in the environment seem to trigger their dissemination thereby conferring upon new hosts the ability to resist to or metabolize these compound. *V. cholerae* derived ICE, SXT encodes SetR, an orthologue of the phage lambda repressor CI and transfer of this ICE is regulated by the SOS response.

Integrative mobilizable elements (IMEs) are non-self-transmissible elements which can excise from chromosome, and transferred to a new host with the help of conjugative functions provided in *trans* by other elements, through the formation of a covalent closed circular molecule and integrate into the chromosome of the host. IMEs are reported from *Salmonella* and *Bacteroides* species (Doublet et al, 2005; Shoemaker et al, 1996). In *Salmonella*, the genomic island, SGI1 is not self-transmissible but can be transferred from a donor strain of *Salmonella* enteric to non-SGI1 containing S. *enteric*

and *Escherichia coli* recipient strains (Doublet et al, 2005). The functions necessary for transfer are provided in *trans* by the donor strains having the conjugative *IncC* plasmid R55 (Doublet et al, 2005). *Bacteroides* species have mobilizable insertion elements known as non-replicative Bacteroides units (NBUs). Four members of the NBU element family (NBU1, NBU2, NBU3, and Tn4555) are shown to be mobilizable *via* a covalently closed circle intermediate (Li et al, 1993, Li et al, 1995). This circle does not replicate but can be transferred by conjugation, starting from an internal transfer origin (*oriT*). NBUs possess *mob* gene but are not self-transmissible, nor can they excise on their own. Both excision and mobilization of NBUs require transacting functions provided by a Bacteroides conjugative transposon (CT) (Shoemaker and Slayers, 1988; Steven et al, 1990; Stevens et al, 1992).

As mentioned in earlier sections all modules required for maintenance and distribution of pPDL2 as an ICE are present in its primary sequence. However, the genes responsible for the formation of mating pair are not present. Instead, genes responsible for its mobilization and integration are present along with the *attP* site required to mediate site specific recombination. The replication initiator, RepB found on pPDL2 can initiate the replication process in rolling circle mode. In addition to *repB* two potential *oriT* sites needed to generate relaxed pPDL2 required to mediate mobilization process are identified on plasmid pPDL2.

The plasmid, pPDL2 of *Flavobacterium* sp. ATCC 27551 has an Integrase coding sequence which can perform its integration and excision into and out of the chromosome. As typically seen in ICE/IMEs,a a putative regulator *copG* is found translationally linked to integrase coding sequence. Further, the Integration Host Factor

(IHF) binding sites have been identified in the upstream region of *copG*, the protein shown to play a major role in excision of the integrative elements (Nash and Robertson, 1981; Bushman et al, 1985). Presence of such *cis*-element in fact strengthens the proposal of existence of an integration and excision event between chromosome and pPDL2. Further, *copG, int* and *pgm* genes are found clustered in pPDL2. Existence of linkage between *copG*, *int* and *pgm* genes, which are named in the present study as CIP unit suggests functional relevance (Fig. 3.17). In pPDL2 there exists, two sets of CIP units, without having an absolute identity. The PGM is shown to have a HP-PEM domain involved in dephosphorylation of histidine phosphates. Presence of such domain indicated regulation of integration /excision event through means of signal transduction. However, further studies are needed to prove existence of such regulation. It is not yet known if growth of *Flavobacterium* sp. ATCC 27551 in presence of OP compounds or its intermediates stimulates the transfer of pPDL2.

What is pPDL2?

After examining the complete sequence of plasmid pPDL2 and looking at the various modules present in the sequence an attempt is made to give a correct name to pPDL2. Historically, it's identified as a plasmid containing organophosphate degrading (*opd*) gene. The replication origin and well conserved replication initiation proteins RepA, maintenance module containing toxin and anti-toxin protein coding sequences justify calling it as a typical bacterial plasmid. However, it also contains a number of features that justify grouping it under Integrative Mobilizable Element. Existence of an integration module (Fig. 3.17 and 3.21), attachment site *attP* and a module that contribute for the mobilization of pPDL2 are typical features of an IME (Juhas et at, 2009). In the light of

these facts how should pPDL2 be called? A plasmid or an IME? It reminds mythological situation, a Chimera and Ganesha….. though considered powerful Gods, for a common man, they neither have complete human nor animal features. Most of the simplest things are hard to explain.

3.13. Conclusions

1. *Flavobacterium* sp. ATCC 27551 contains four indigenous plasmids designated as pPDL1, pPDL2, pPDL3 and pPDL4.

2. The indigenous pPDL2 alone contains *opd* gene and it is rescue cloned into *E.coli pir*-116 cells.

3. Plasmid library was constructed by cloning *Eco*RI and *PstI* fragments of pPDL2 in pBluescript KS II.

4. Complete sequence was determined for the 39.75 kb pPDL2.

5. The GC content of pPDL2 (61.76%) is found to be very close to the GC content of plasmids pSWIT02 and pUT2 of *Sphingomonas wittichi* RW1 and *Sphingobium japonicum* UT26S respectively.

6. Nucleotide sequence of plasmid pPDL2 of *Flavobacterium* sp. ATCC 27551 showed maximum homology to the indigenous plasmids pUT1, pCHQ1 of *Sphingobium japonicum* UT26S and pSWIT01of *Sphingomonas* sp.

7. Sequence analysis of pPDL2 has revealed presence of 42 ORFs out of which only 18 are hypothetical and the remaining 24 ORFs have shown strong homology to the well characterized proteins found in NCBI database

8. Open reading frames present on pPDL2 are organized into modules, such as (i) the replication and partition module, (ii) mobilization module, (iii) integrase module, (iv) degradation module and (v) mobile genetic elements.

The complete sequence information of plasmid pPDL2 is presented in chapter-I. In the sequence of pPDL2, *oriT* and *repB* that contribute for horizontal mobility of plasmid pPDL2 were identified. Further, the literature available on the structure and organization of *opd* genes coding organophosphate hydrolase clearly suggest existence of horizontal transfer of *opd* genes among soil bacteria (Siddavattam et al, 2003; Horne et al, 2003; Wei et al, 2009). However no experimental evidence is available to support the hypothesis. This chapter describes horizontal mobility of *opd* genes through experimental design described in materials and methods section. The well characterized *opd* plasmids pPDL2 of *Flavobacterium* sp. ATCC 27551 and pCMS1 of *Brevundimonas diminuta* were taken as experimental tools while gaining the experimental evidence on horizontal gene transfer (HGT) of *opd* gene.

4.1. Horizontal mobility of pPDL2 of *Flavobacterium* sp. ATCC 27551

Horizontal transfer of genes among bacterial strains is a well defined subject. In order to accomplish HGT among bacterial strains plasmids should have genes that code for mating-pair formation (MPF) between the donor and the recipient cells. In addition to the mating pair formation module, an *oriT* sequence and a *mob* gene coding for a relaxase is required to achieve successful transfer of plasmid DNA (Holmes and Jobling, 1996; Russi et al, 2008). In plasmid pPDL2 no MPF machinery is present. However, existence of *oriT* like structures and a replicase, RepB, the rolling circle replication initiator suggests mobilizable nature of pPDL2. In order to know the mobilizable nature of pPDL2, biparental and triparental mating experiments were performed as described in materials and methods section. While performing biparental mating experiments, the pPDL2 derivative, pPDL2::Tn5<R6Kγori-Kan-2> containing *E. coli pir*-116 was used as donor. Usage of pPDL2 derivative generated by inserting mini-transposon, EZ-Tn5<R6Kγori-Kan-

2>, confirms kanamycin resistance on pPDL2:: Tn5<R6Kγori-Kan-2>. Usage of an *E. coli* cell having pPDL2::Tn5<R6Kγori-Kan-2> as donor facilitates easy monitoring of pPDL2 mobility by selecting on a kanamycin antibiotic resistance plate. In an experiment conducted using *E. coli pir*-116 (pPDL2::Tn5<R6Kγori-Kan-2>) as donor and *Acinetobacter* sp. DS002 as recipient no exconjugants of *Acinetobacter* sp. DS002 were identified on a selective plate having kanamycin and benzoate (Fig. 4. 1). Logically on a minimal agar plates supplemented with 5mM benzoate and kanamycin colonies should appear if pPDL2 is a self transmissible plasmid. *Acinetobacter* sp. DS002 strains are kanamycin sensitive and grow on benzoate using as sole source of carbon. If pPDL2::Tn5<R6Kγori-Kan-2> is mobilized into recipient strain colonies of *Acinetobacter* sp. DS002 exconjugants would have grown on benzoate + kanamycin plates. However, no exconjugants were observed on the selection plates, which are also in agreement of the sequence information of pPDL2 (Fig. 4. 1).

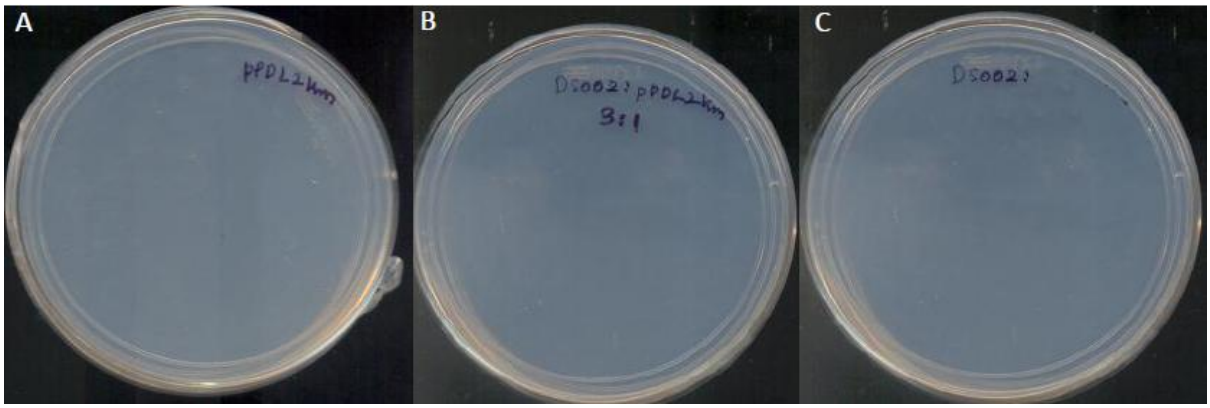


Fig. 4.1. Biparental mating using *E. coli pir*-116 (pPDL2::<R6Kγori-Kan-2>) and *Acinetobacter* sp. DS002 as donor and recipients. Panel A represents selection plate spread with *E. coli pir*-116 (pPDL2<R6Kγori-Kan>), B represents selection plate spread with conjugation mixture having *E. coli pir*-116 (pPDL2<R6Kγori-Kan>) (donor) and *Acinetobacter* sp. DS002 (recipient). Panel C represents selection plate with *Acinetobacter* sp. DS002 (recipient). No exconjugants were seen in selection plates having mating mixture.

4.1.1. Triparental mating

The above experiment clearly rules out the possibility of pPDL2 as a self-transmissible plasmid. However, if sequence information taken as basis to assess lateral transfer of pPDL2 it clearly suggests that it can only be mobilized if donor strain can supplement genetic machinery to code for a MPF. Therefore, such situation is created by performing triparental mating experiments, where the helper *E. coli* HB101 harbouring plasmid pRK2013 (Figurski and Helinski, 1979), *E.coli pir*-116 harbouring pPDL2::Tn5<R6Kγori-Kan-2> and *Acinetobacter* sp. DS002 strains served as helper, donor

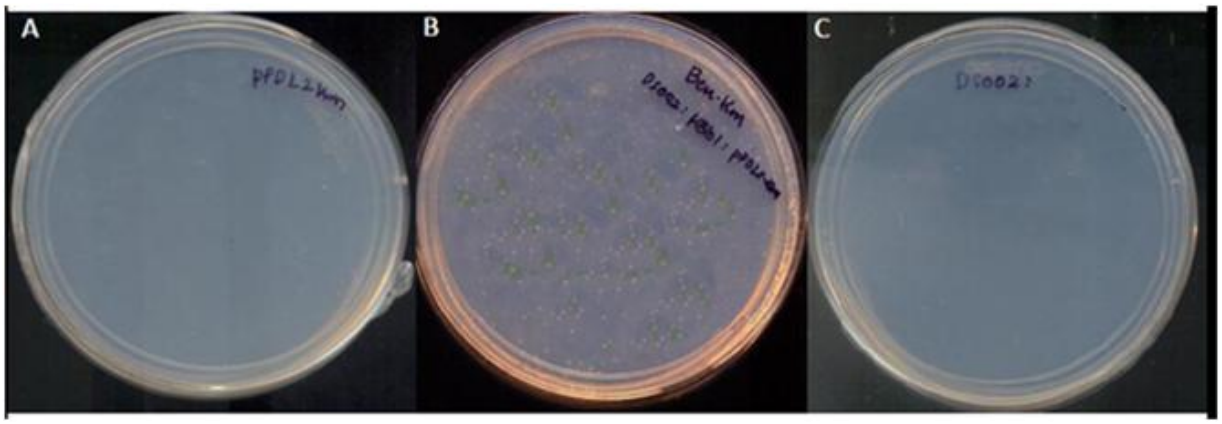


Fig. 4. 2. Triparental mating using *E. coli pir*-116 (pPDL2<R6Kγori-Kan>), *E. coli* HB101 (pRK2013) and *Acinetobacter* sp. DS002. Panel A represents selection plate spread with *E. coli pir*-116 (pPDL2<R6Kγori-Kan>), B represents selection plate spread with mating mixture having *E. coli pir*-116 (pPDL2<R6Kγori-Kan>) (donor), *E. coli* HB101 (pRK2013) (helper) and *Acinetobacter* sp. DS002 (recipient). Panel C represents selection plate with *Acinetobacter* sp. DS002. Appearance of exconjugants is seen in panel B.

and recipients respectively. The helper plasmid pRK2013 provides the genetic machinery for formation of mating pair formation and mobilization (Figurski and Helinski, 1979). Exconjugants appeared on the selection plates indicating the mobilization of plasmid pPDL2::Tn5<R6Kγori-Kan-2> into *Acinetobacter* sp. DS002 (Fig. 4.2). In the previous chapter, based on sequence information, a well conserved *oriT* was identified along with

RC replication origin and *repB* gene. The experiment described above has shown existence of functional mobilization module plasmid pPDL2.

4.1.2. Characterization of exconjugants

4.1.2.1. Detection of *opd* gene

After mobilization of plasmid pPDL2::Tn<R6Kγori-Kan-2> from *E.coli pir*-116 into *Acinetobacter* sp. DS002 through the helper strain *E.coli* HB101 (pRK2013), the presence of plasmid in the recipient strain was analyzed by performing PCR using the plasmid borne *opd* specific primers. When PCR was performed using *opd* domain specific primers, a specific amplicon of 500 bp was seen only in PCR mix having a colony of exconjugant and donor. No such amplicon was seen in recipient cells (Fig. 4.3).

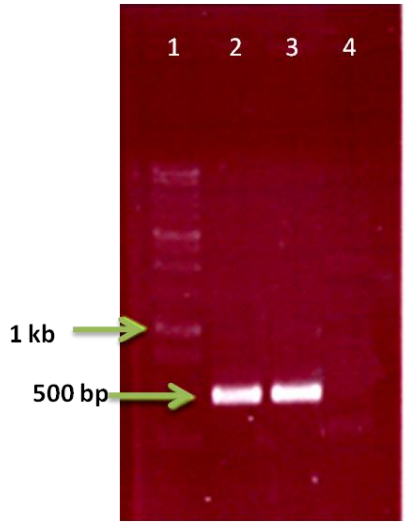


Fig. 4. 3. Confirmation of pPDL2 mobilization into *Acinetobacter* sp. DS002 by colony PCR using *opd* specific primers. Lane 1 represents 1 kb DNA ladder. Lanes 2-3 represent amplicons obtained from exconjugant and donor colony containing PCR mix. The PCR mix of recipient is shown in lane 4.

4.1.1.2. OPH assay

After establishing stable maintenance of plasmid pPDL2::Tn5<R6Kγori/Kan-2> in *Acinetobacter* sp. DS002, experiments were conducted to test their ability to degrade OP compound, paraoxon. Our lab has recently shown presence of OPH in the inner

membrane of *B. diminuta* and its dependence on Twin Arginine Transport (Tat) pathway for membrane targeting (Gorla et al, 2009). Therefore, the *Acinetobacter* cells harboring plasmid pPDL2::Tn*5*<R6Kγori/KAN-2>were fractionated in to cytoplasmic and membrane fractions and were assayed for OPH activity as described in methods. Most of the OPH activity was found in membrane fraction and very little activity was seen in cytoplasmic fraction (Fig. 4. 4). This result clearly suggests expression of active OPH from plasmid pPDL2::Tn*5*<R6Kγori/KAN-2>.

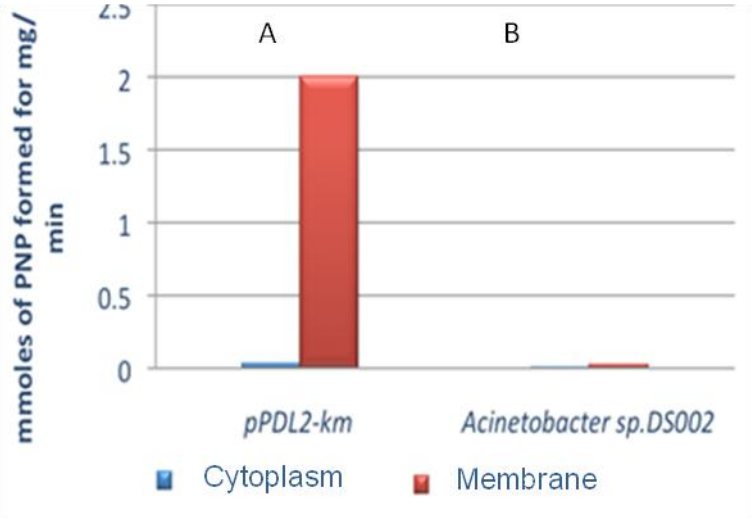


Fig. 4.4. Assay of organophosphorus hydrolase activity in cytoplasmic and membrane fractions of A) exconjugant (*Acinetobacter* sp. DS002 having plasmid pPDL2::Tn5<R6Kγ/Kan-2>) and B) wild type *Acinetobacter* sp. DS002.

4.2. *In vivo* transposition assay

As shown in Fig. 3.22 and described in chapter-1 the *opd* gene cluster of pPDL2 has shown a transposon-like module. Transposon Tn*3* and y4qE flank all the degradative traits that code for enzymatic machinery with a possible involvement of mineralizing OP insecticides like methyl parathion and fenitrothion. The organophosphorus hydrolase is a well characterized triesterase involved in hydrolysis of tri-ester linkages present in diverse group of OP compounds (Benning et al, 1994; Cho et al, 2004). Similarly, the

mfhA gene present immediately downstream of *opd* is a *meta* fission product hydrolase. Downstream of *mfhA* and *opd,* transposon Tn*3* consisting of *tnpA* and *tnpR* which codes for transposase

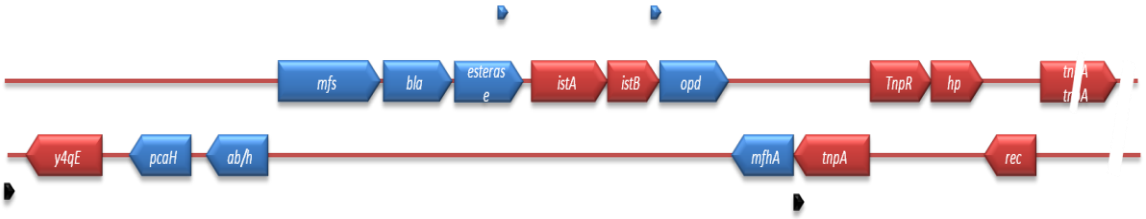


Fig. 3.22 of Chapter 1. Organization of degradative genes and mobile elements on pPDL2. Transposases and resolvases are shown in red colour. Degradative genes are shown in blue coloured arrows. The repeats of Tn*3* and IS*21* element are shown as black arrows and blue arrows respectively.

and resolvase was identified. In the upstream region of *opd* gene an IS element designated as IS*Flsp1* was identified (Siddavattam et al, 2003). In the sequence information generated in the present study two more open reading frames designated as ORF2 and ORF3 are identified upstream of IS*Flsp1*. One of them codes for β-sub-unit of protocatechuate 4,5 dioxygenase and the second ORF has homology to β-ketoadipate lactonase. Existence of another transposable element y4qE downstream of these genes indicates transposition of *opd* cluster. Existence of Tn*3* transposase specific inverted repeats flanking downstream of y4qE and upstream of transposase Tn*3* supports possible transposition event of *opd* element found in plasmid pPDL2.

In order to gain experimental evidence for such transposition event an *in vivo* transposition assay was performed by following procedures described in materials and methods section. The *in vivo* transposition assay was performed using three independent compatible plasmids. The plasmid pTras::*tet* contains the entire pPDL2 cluster cloned in

pUC18 vector. Its construction is described in one of the earlier reports published from our laboratory (Siddavattam et al, 2003). This plasmid serves as donor of *opd* cluster. The second plasmid, pJQ210SK (Quandt and Hynes, 1993) contains *sacB* gene. Presence of *sacB* gene is lethal for gram negative bacteria in presence of sucrose. The third plasmid is pMMBTnpA, derivative of pMMB206 generated by cloning *tnpA* under the control of inducible promoter. All the three plasmids were transformed into *E. coli* (pTrans::tet + pJQ210 + pMMB-TnpA) and were induced for expression of transposase by adding low concentrations of IPTG. After induction the cells were then plated on LB plates containing sucrose + gentamycin +tetracycline. A schematic representation is given in Fig. 4.5 to explain the functioning of the *in vivo* transposition assay.

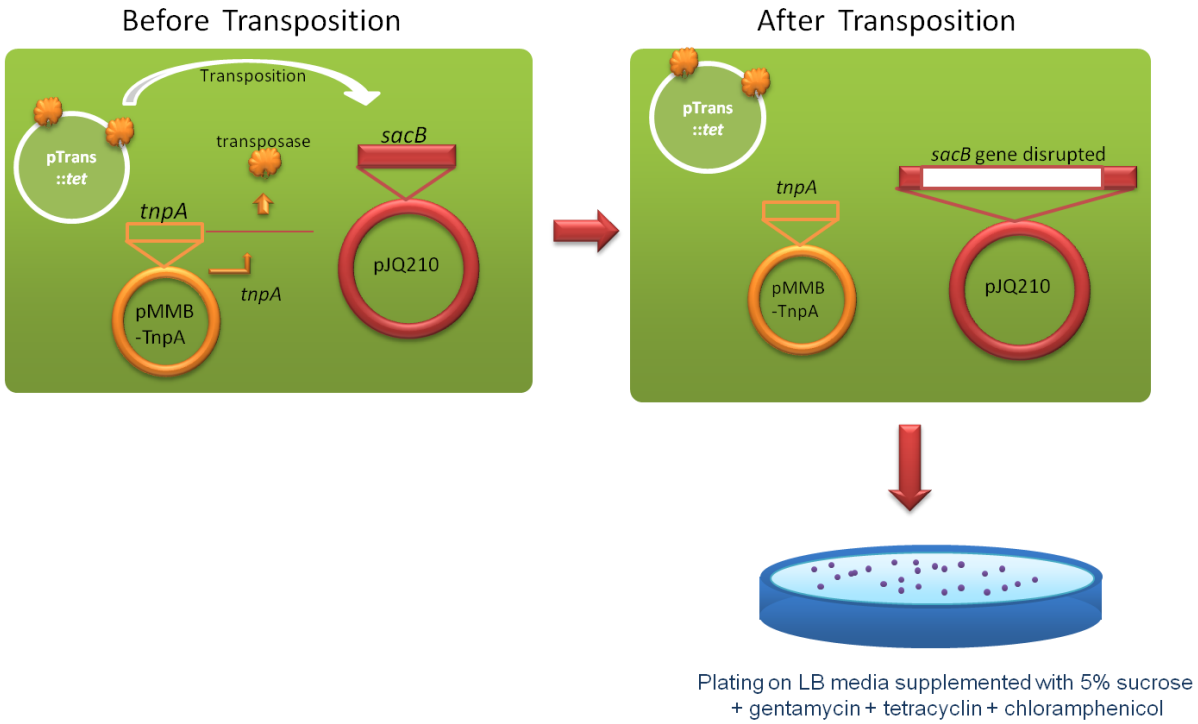


Fig. 4.5. Schematic representation of *In vivo* transposition assay

E.coli cells having either pTrans::*tet* (donor of *opd* cluster) or reporter plasmid (pJQ210SK) served as negative control. As shown in Fig. 4.6 a number of colonies are

found in selection plates containing sucrose + gentamycin + tetracycline + chloramphenicol. Such appearance of colonies on sucrose plates is only possible in the event of disruption of *sacB* gene due to transposition.

A) B)

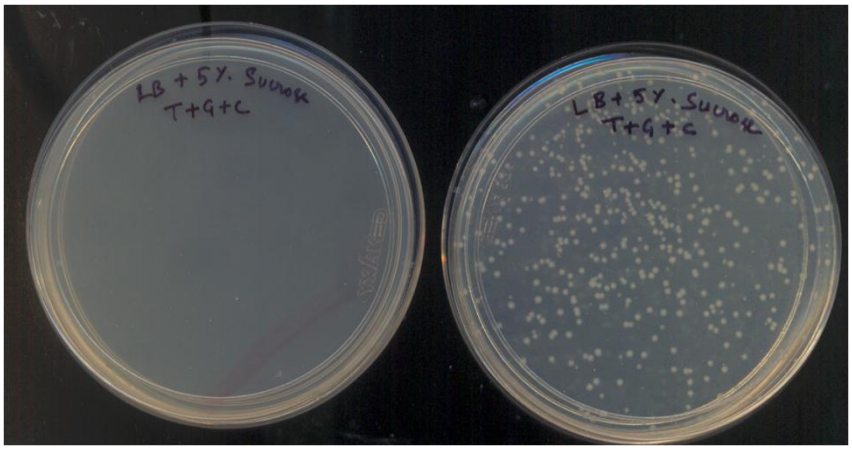


Fig. 4.6. *In vivo* transposition assay for demonstrating transposition of *opd* gene cluster.Panel A and B represent LB + 5% sucrose (Kanamycin + gentamycin + tetracycline) plates spread with *E.coli* having (pTrans::*tet* + pJQ210 + pMMB206) and *E.coli* having (pTrans::*tet* + pJQ210 + pMMB-TnpA). Sucrose resistant colonies are only seen in cells expressing TnpA.

4.2.1. Analysis of sucrose resistant colonies

In order to gain further insights into the nature of DNA fragment causing disruption of *sacB* gene, the pJQ210 derivatives were isolated from sucrose resistant colonies. The plasmids prepared from the resistant colonies were used as templates for performing PCR using *sacB* specific primers. PCR amplification using these primers gave 1.2 kb *sacB* amplicon, if it is not disrupted. However, if *sacB* is disrupted due to insertion of *opd* element found on the donor plasmid pTrasn::*tet*, there will be increase in size of the amplicon. Figure 4.7 shows that amplicons obtained from a number of colonies have

a size greater than the *sacB* gene. Colony PCR performed using *opd* specific primers from these colonies gave

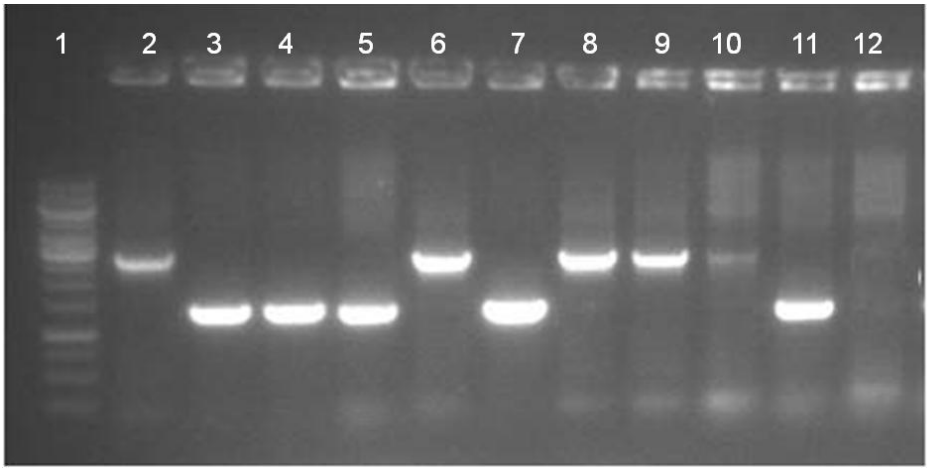


Fig. 4.7. Screening of plasmids isolated from sucrose resistant colonies using *sacB* specific primers. Increase in size indicates disruption of sacB gene due to transposition. Lane 1 represents 1kb DNA Ladder. Lanes 2 -10 represent amplicons from sucrose resistant colonies. Lane 11 and 12 represent positive control (pJQ210SK) and negative control (pTrans::*tet*, donor plasmid).

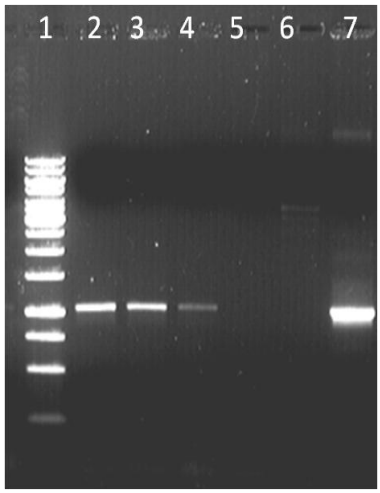


Fig. 4.8. Screening of sucrose resistant colonies using *opd* specific primers. Lane 1 represents 1kb DNA Ladder. Lanes 2 -6 represent amplicons of *opd* from sucrose resistant colonies. Lane 7represent positive control (pTrans::*tet*).

amplicons corresponding to the size of *opd* gene (Fig. 4.8). These two experiments provide a clear evidence to claim the transposable nature of the *opd* element found on plasmid pPDL2. However, these pJQ210 derivatives containing *opd* element insertion in

sacB gene needed to be sequenced to identify boarder sequences of the *opd* element. Due to paucity of time such experiments were not conducted in this study. However, our laboratory is undertaking such studies to identify precise junctions of *sacB* and *opd* element to determine *opd* specific sequences.

The studies described in chapter-2 document horizontal nature of *opd* genes found on pPDL2 of *Flavobacterium* sp. ATCC 27551 and clarify to large extent for the reasons behind the existence of identical *opd* elements on dissimilar plasmids and chromosomes of soil bacteria isolated from diverse geographical regions.

4.3. Horizontal transfer of plasmid pCMS1

Brevundimonas diminuta is one of the first microorganisms reported to have the capability to degrade organophosphorus pesticides (Serdar et al, 1982). It was further demonstrated that the OP compound degradation capability of *Brevundimonas diminuta* was due to existence of a large indigenous plasmid, pCMS1 (Mulbry et al, 1986; McDaniel and Wild, 1988). The comparison of restriction profile of pCMS1 and pPDL2 revealed existence of sequence similarity only in 5.1 kb DNA region containing *opd* gene. As described before plasmid pPDL2 is clearly a mobilizable plasmid. It can constantly contribute for HGT of *opd* gene under certain conditions, especially when it gains access to mating pair formation coded by another plasmids/ Integrative conjugative element.

4.3.1. Random sequencing of pCMS1

Unlike in pPDL2, no sequence information is available for pCMS1. In order to test the HGT of pCMS1 two things need to be generated. Initially the plasmid pCMS1 need to be tagged with an antibiotic marker so that its mobility can be monitored by performing a typical conjugation experiment. Secondly the plasmid pCMS1 need to be sequenced to

identify genetic modules coding for mating pair formation (MPF) and *oriT* and *mob* genes. In order to obtain quick information on the sequence of pCMS1 two fosmid clones having entire pCMS1 DNA were taken and *in vitro* transposon tagging experiments were performed to insert mini-transposon EZ-Tn5<R6Kγori-Kan-2> randomly into fosmid clones of pCMSA and pCMSB. After *in vitro* transposon tagging experiments, the kanamycin resistant *E. coli pir*-116 cells were taken and the fosmids having mini-transposon insertions were independently sequenced using transposon specific primers. Such sequence stretches were then analyzed to identify presence of *tra* genes that contribute for mating pair formation (MPF). The sequence information has shown the presence of *tra* genes on plasmid pCMS1 providing *prima face* evidence for horizontal mobility of plasmid pCMS1 (Table 4.1). As

Table 4.1. List of *tra* genes identified in plasmid pCMS1.

S. No.	Name of the recombinant fosmid.	Name of the Tra protein deduced from nucleotide sequence & % homology	Role in horizontal gene transfer (HGT)	*Reference strain/* GeneBank Accession No.
1	pCMSB314	TraI ; 54	DNA relaxase	*Comamonas sp.* CNB-1 / ABM06255
2	pCMSB32R	TraE ; 72	DNA topoisomerase	*Achromobacter xyloxidans sub-sp. denitrificans* / AAS49467
3	pCMSB32F	TraE ; 64	DNA topoisomerase	*Achromobacter xyloxidans sub-sp. denitrificans* / AAS49467
4	pCMSB31R	TraE ; 70	DNA topoisomerase	*Achromobacter xyloxidans sub-sp. denitrificans* / AAS49467
5	pCMSB31F	TraM ; 56	Conjugal transfer protein	*Achromobacter xyloxidans sub-sp. denitrificans* /AAS49476

shown in Table 4. 1, *traI*, *traE* and *traM* genes were identified on plasmid pCMS1. All the *tra* sequences share considerable similarity (54-72%) with *tra* genes present on plasmid

pEST4011 of *Achromobacter xyloxidans* sub-sps. *denitrificans*. The TraI is a relaxase which performs strand scission at the transfer origin (*oriT*) and TraM acts a topoisomerase which enhances TraI activity. TraE is one of the proteins present in the pilus assembly (Karl et al, 2003). As existence of the *tra* genes was apparent on pCMS1, further experiments were done to obtain experimental evidence on its horizontal mobility.

As no markers are available to track mobility of pCMS1, the *B.diminuta* strains having its pCMS1: *tet* generated previously by replacing the *opd* with *opd::tet* (Gorla et al, 2009) was used as donor strain. In a typical conjugation experiment, performed by using *B. diminuta* (pCMS1::*tet*) and *P. putida* KT2440 as donor and recipient, respectively, plasmid pCMS1::*tet* was successfully transferred with a frequency of 0.72 x 10^{-6} . This result clearly suggests that plasmid pCMS1 is a self transmissible plasmid and supports the sequence information that gave clear indication about the presence of *tra* genes.

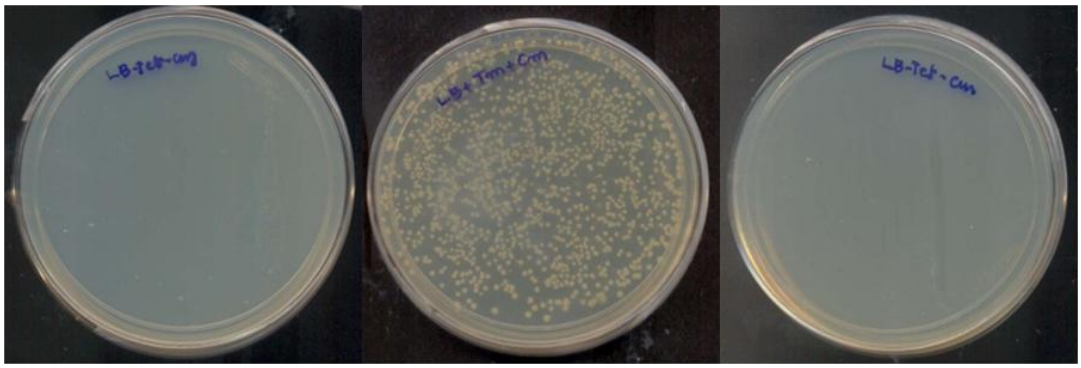


Fig. 4.9. Panel A, B and C indicate LB (Km+Cm) plates spread with donor (*B.diminuta* with pCMS1::*tet*), mating mixture and recipient strains (*Pseudomonas putida* KT2240). Colonies were only seen on plates spread with mating mixture.

4.3.2. Analysis of exconjugants

The exconjugants generated in this study were analyzed to see presence of *pCMS1*::*tet*. Initially, the presence of the plasmid was detected in exconjugant and donor by isolating plasmids following protocol described in materials and methods section

(Currier and Nester, 1976). As shown in figure 4.10-A a clear plasmid band was seen in exconjugant and donor. No such plasmid band appeared in recipient cells. The plasmid preparations were then analyzed for presence e of *opd*::*tet* by performing PCR using *opd* specific primers. Agarose gel analysis of PCR samples has shown amplification of *opd::tet* from an exconjugant. The *tet* gene has an internal *BamHI* site. The amplicon when digested with *BamHI* generated two fragments of 1.1 kb and 0.8 kb respectively, confirming existence of *opd*::*tet* in plasmid preparations made from exconjugant (Fig. 4.10).

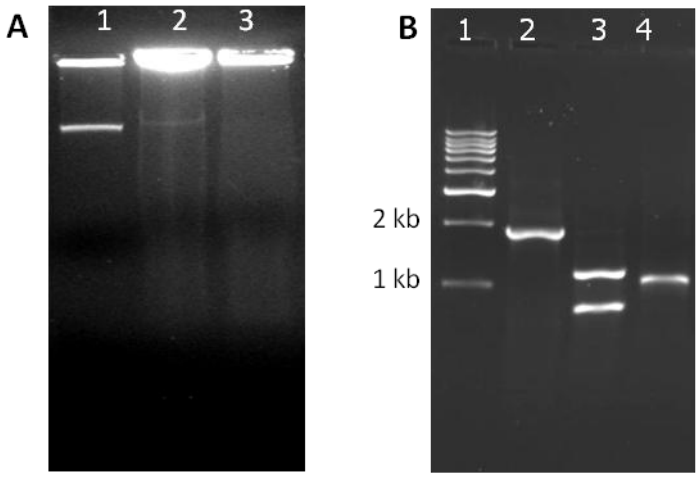


Fig. 4.10 Panel A. Agarose gel electrophoresis showing existence of pCMS1::*tet* in donor (lane1) and exconjugant cells (lane 2). No plasmid was seen in *Pseudomonas putida* KT4220 used as recipient (lane 3). Panel B. Agarose gel showing PCR amplification of *opd::tet* from exconjugants of *B. diminuta.* Lane 1, kb ladder, lane 2 PCR amplification of *opd::tet* from exconjugants. Lane 3 shows digestion of *opd::tet* with *Bam*HI, an unique site found in *tet* gene. Lane 4 represents amplification of 1.2 kb *opd* gene from wild type *B. diminuta* cells.

4.4. Discussion

The work presented in this chapter clearly suggests possible Horizontal Gene Transfer (HGT) of the *opd* gene among soil bacteria, in the light of these results a

thorough literature search was done to have an understanding of the origin, evolution and degradation of organophosphorus degrading traits. The organophosphorus hydrolyzing enzymes were originally named as phosphotriesterases due to their ability to hydrolyze tri-ester linkage found in organophosphates and nerve agents. The tri-esterases found in prokaryotes can be divided into three independent groups. There exists no sequence identity among these three groups. They are organophosphorus hydrolases encoded by *opd* gene identified in *Flavobacterium* sp. ATCC 27551 and *Brevundimonas diminuta*. The second group belongs to methyl parathion hydrolases (MPH) group. These are mainly isolated from Chinese agricultural soils. The third group belongs to organophosphorus acid anhydrolase (OPAA) group. A The OPAAs were later identified as prolidases, the dipeptidases found in variety of bacterial strains. The structural similarity between natural substrate, dipeptide and nerve agent strain is shown to be responsible for the triesterase activity (Cheng and DeFrank, 2000). Therefore, in the subsequent sections, a brief description is given only on the structure and function of *opd* and *mpd* genes, which are considered to be evolved from lactonases and β-lactams respectively. In the preceding sections a detailed description is given on organization and HGT of *opd* elements found on plasmid pPDL2 and pCMS1. Therefore further description on *opd* elements are avoided and a brief mention is made in the following sections on structural organization of other phosphotriesterase sequences found in taxonomically diverse group of organisms.

4.4.1. The Tn*opdA* element

The *opd* homologue of *Agrobacterium radiobacter* P230 is *opdA* (Horne et al, 2002). The chromosomally located *opdA* gene was later shown to be part of a

transposable element, which contained three further ORFs in addition to *opdA* along with inverted repeats typically seen in transposon Tn*610* of *Mycobacterium fortuitum* (Horne et al, 2003). The transposase TnpA is identical with the TnpA sequence of Tn*610*. The other two ORFs found between *tnpA* and *opdA* were predicted to code for a truncated transposase (*orfA*) and an ATP binding protein (*orfB*). Transposition was successfully shown in *E. coli* confirming the horizontal mobility of *opdA* sequences among soil bacteria (Horne et al, 2003).

4.4.2. The *mpd* elements

All methyl parathion-degrading (*mpd*) genes reported to date have been isolated from Chinese agricultural soils or from the activated sludge collected from a Chinese pesticide manufacturing unit. The first *mpd* gene was cloned from a *Plesiomonas* sp. strain M6 (Zhongli et al, 2001). This chromosomally located *mpd* gene surprisingly has shown no homology to any of the known *pte* genes. Following this discovery, a number of *mpd* sequences were cloned from Chinese agricultural soils (Liu et al, 2005; Zhang et al, 2006). However, the horizontal mobility of *mpd* genes among soil microbes gained acceptance only with the discovery of plasmid-borne *mpd* gene in *Pseudomonas* sp. strain WBC-3. In this soil isolate an indigenous plasmid of 70 kb designated pZWL0 contained both a *mpd* gene and genes responsible for degradation of *p*-nitrophenol. This strain uses methyl parathion and its degradation product *p*-nitrophenol as sole source of carbon, nitrogen and energy (Liu et al, 2005). Further investigations into the genetics of methyl parathion degradation revealed that the organization of the *mpd* gene in this strain was like that of a functional transposon (Fig. 4.11).

4.4.2.1. The Tn*mpd* element is a typical class I Transposon

Sequence analysis of the *mpd* region of pZWLO of *Pseudomonas* sp. strain WBC-3 revealed the existence of a functional *mpd* element. When the 6.5 kb *Kpn*I-*BamH*I fragment was sequenced, it revealed existence of the IS*6100* class of IS elements that flanking the *mpd* gene (Fig. 4.11). Further studies conducted by Wei et al (2009) have elegantly demonstrated the transposition event of the *mpd* gene in *Pseudomonas* sp. strain WBC-3 (Wei et al, 2009). This is the first functional transposable element with a *mpd* gene and the second one in the entire *pte* family of genes that codes for a phosphotriesterase.

4.4.2.2. Distribution of *mpd* elements

After establishing the existence of a functional *mpd* element, Tn*mph,* in *Pseudomonas* sp. strain WBC-3, Zhang et al, have isolated seven bacterial strains capable of degrading methyl parathion from different locations of Chinese soils with a history of using methyl parathion (Zhang et al, 2006). In all of them a *mpd* gene containing a 4.7 Kb region is highly conserved. In this conserved DNA region a total of five ORFs were identified. One of them shows similarity to the TnpA-coding sequence of an IS element, IS*6100*, and contains a perfect 14 base pair inverted repeat in its flanking sequences. The second ORF, designated as *orf463*, found immediately downstream of the IS element codes for a protein that shows considerable homology to a house-keeping sigma factor. Significance of its existence as part of *mpd* element is unclear. However, presence of other ORFs as part of the *mpd* cluster appears to have a strong functional relevance. The ORFs found upstream of *mpd* sequences, *orf232* and *orf259*, code for a permease of an ABC-transport system and an ExeB found to be the part of general protein secretion pathway. Although there is no experimental evidence to support, due to the presence of

these two ORFs in association with the *mpd* gene, which codes for a MPH precursor with a 35 amino acid-long signal peptide, it is proposed that these two proteins are involved transport and maturation (Zhang et al, 2006). Of the seven *mph* sequences known, two of them are identical, whereas and the rest of the five sequences code for MPH proteins with amino acid substitutions at 9 positions. Nevertheless, these variations have been shown to have positive effect on the catalytic properties of the MPH (Dong et al, 2005). The existence of such highly conserved *mpd* clusters in seven different bacterial strains that show a weak taxonomic relationship strongly supports horizontal mobility among *mpd* genes in soil microbes. The presence of IS*6100* in all these clusters adds strength to this proposal.

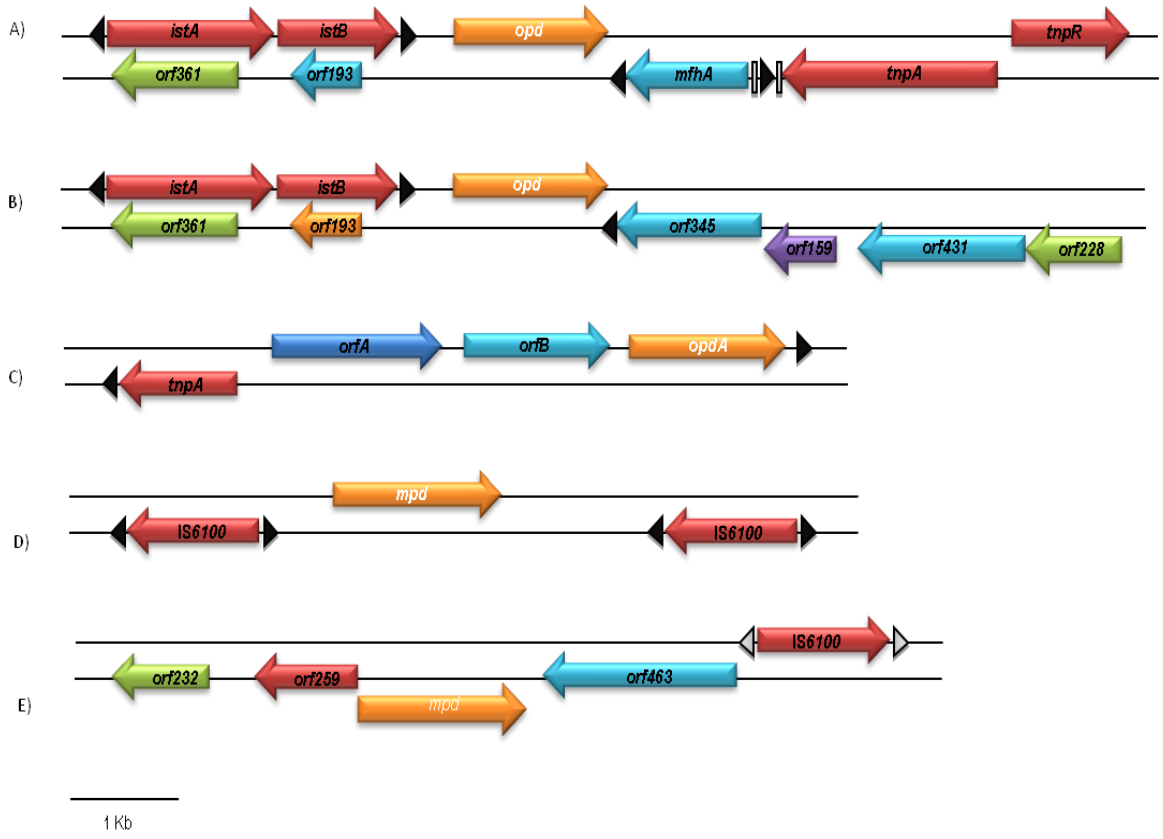


Fig. 4.11. Physical maps showing organization of the *opd* clusters of pPDL2 from *Flavobacterium* sp. ATCC 27551 (A) and pCMS1 from *Brevundimonas diminuta* (B), *opdA* of *Agrobacterium radiobacter* P230 (C), *mpd* of *Pseudomonas* sp. WBC-3 (D) and *Plesiomonas* (W). Arrows indicate the direction of transcription.

4.4.3. The *opaA* genes

After purification of organophosphate acid anhydrolase (OPAA) enzymes from various halophilic and *Alteromonas* species (DeFrank et al, 1993) and demonstration of their ability to degrade G-class nerve agents (Cheng et al, 1999), the *opaA* gene was cloned from the *Alteromonas* sp. strain JD6.5. The *opaA* gene product has more than 50% amino acid similarity to *E. coli* PepQ. Further investigations into the physiological role of OpaA have established that it has prolidase activity. There are no indications that the gene is organized as a mobile genetic element or of its presence on plasmid. The *opaA* gene does not appear to have evolved to code for a phosphotriesterase. The activity of its product, prolidase, on G-class nerve agents, is seen as an ancillary activity of these enzymes due to structural similarity of their substrates (Merone et al, 2005).

4.5. Evolutionary link between phosphotriesterases and lactonases

Promiscuous activities play a key role in the evolution of enzymes. They actually serve as starting point for acquiring a new function through gene duplication (Kolalowski et al, 1997; Lai et al, 1995; Rastogi et al, 1997; Benning et al 1994; Harper et al, 1988). In fact, these promiscuous activities are considered to be the vestiges of the function of their ancestral protein (Kolalowski et al, 1997; Lai et al, 1995). The phophotriesterases have been shown to have promiscuous phosphodiesterase, carboxyl esterase, and lactonase activities (McDaniel et al, 1988; DeFrank et al, 1993; Cheng et al, 1997). In general, the family members that have presumably diverged from a common ancestor often share promiscuous activities (Poelarends et al, 2005, Roodveltdt et al, 2005, Yew et al, 2005, Elias et al, 2008). Afriat et al have elegantly shown the existence of reciprocal promiscuities between lactonases and triesterases (Afriat et al, 2006). Based on the

structural differences, especially in the loops 1, 7 and 8 that comprise substrate binding sites, they have classified OPH homologues into three groups (Fig. 4.12). In the first group, designated as phosphotriesterases, with more than 86% identity to *bd*-OPH, they have shown existence of promiscuous lactonase. The *ec*-OPH is kept in the second group of OPH homologues, as it contained relatively shorter substrate binding loops. The third group of enzymes have only loop 7 (Fig. 4.12). These proteins annotated in the database as putative parathion hydrolases including AhlA from *R. erythropolis*, PPH from *Mycobacterium tuberculosis* and SsoPox of *Sulfolobus solfataricus* and are all re-classified as phophotriesterase like lactonases (PLLs). All of them proficiently hydrolyzed lactone with distictively low Km (10 -230uM) values and a very weak phosphotristerase activity (102 to 106 fold). The arylesterase activity was shown only by SsoxPox but not by any of the other PLLs. In principle, the promiscuous activity shown by a family member is not seen with other members of the family. If activity is shown by all the members, it is considered to be indicative of native function (Khersonsky et al, 2006). If this analogy is taken into consideration the PLLs are primarily lactonases with promiscuous phophotriesterase activity and probably the phophotriesterases are evolved from PLLs in the recent past. The substrate-binding loops contribute the main structural difference between PTEs and PLLS.

Indeed insertions, deletions and loop-swapping are believed to be a primary mechanism for creating enzyme diversity (Twafik, 2006, Park et al, 2006, Soskine and Tawfik, 2010). A number of studies have used PLLs as templates for directed evolution and succeeded in either enhancing the substrate range, catalytic efficiency (Chow et al, 2009) or converting PLLs to catalyze altogether new reactions (Mandrich and Manco,

2009). Among the proposed PLLs the archaeal triesterase alone is shown to have arylesterase activity (Afriat et al, 2009). As *ec*-OPH has also shown arylestrase activity the PLL, SsoxPox, due to the existence of similar activity and structural similarities, is proposed to be a "generalist" molecule that served as a template for evolution of phosphotriesterases found in mesophilic organisms (Afriat et al, 2006, Merone et al, 2008). In fact, the recent discovery of phosphotriesterase-like carboxyesterase, (*Mlo*PLC) from *Mesorhizobium loti* and its transformation into a diesterase through *in vitro* evolution supports the proposed hypothesis by Afriat et al, 2006 (Mandrich and Manco, 2009).

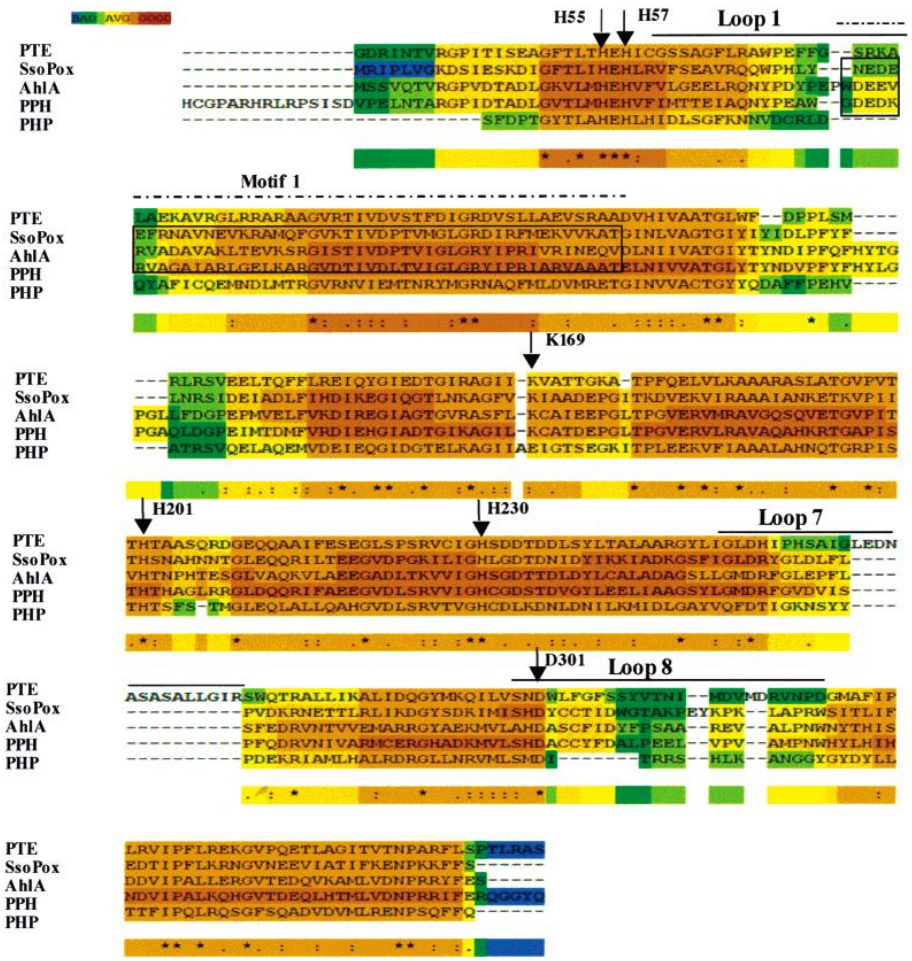


Fig. 4.12. Multiple Alignment of Phosphotriesterases of *B.diminuta* (PTE), *Sulfolobus sulfatricus* (SsoPox), *R. erythropolis* (AhlA), *Mycobacterium tuberculosis* (PPH) are shown using T-Coffee program.

4.6. MPH Scenario

The scenario with the evolution of methyl parathion hydrolases appears to be in no way different from the evolution of the OPHs. They appear to have evolved from ß-lactamases with which they share considerable structural homology. The N-acyl-L-homoserine lactone (AHL) lactonases are members of the metallo-ß-lactamase superfamily and contain two zinc ions in their catalytic center (Aravind, 1999; Daiyasu et al, 2001; Crowder et al, 2006). The recently solved crystal structure of *Bacillus thuringiensis* AHL lactonase (Liu et al, 2008) has shown striking similarity with the crystal structure of MPH (Dong et al, 2005) (Fig. 4.13). The MPH is also shown to have promiscuous lactonase activity (Afriat et al, 2006). If these findings are seen with the aforementioned experimental evidence gathered to show the structural relationship between lactonases and phosphotriesterases, the proposal that metallo ß-lactamases were the progenitors of MPH is worthy of consideration.

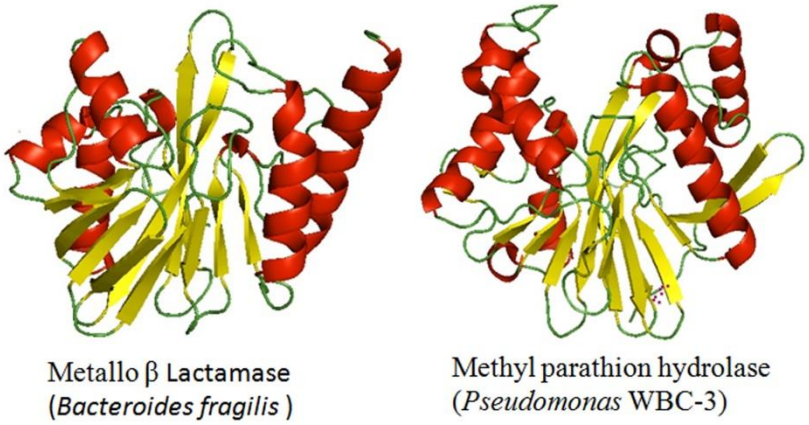


Fig. 4.13. Ribbon diagram of Metallo-β Lactamase and methyl parathion hydrolase showing similarities in their structure

Evolution of such new traits coding for biodegradation of recalcitrant xenobiotics and recalcitrant aromatic compounds is not uncommon and has been reported frequently in

Organophosphates, as stated in the introduction chapter, are widely used as insecticides to control various insect pests that affect economically important crops. The organophosphorus hydrolase (OPH) encoded by plasmid pPDL2 borne organophosphate (*opd*) gene hydrolyses triester linkage found in variety of OP insecticides (Mulbry and Karns, 1989; Dumas et al, 1989). This hydrolytic step inactivates OP compounds and thus reduces their toxicity towards mammals and other non-target organisms having well developed nervous system. However, this hydrolytic cleavage generates a number of aromatic compounds, especially nitrophenols (Fig. 5.1), which are highly toxic to soil microflora and are shown to adversely influence soil ecosystem (Camper et al, 1991). Complete mineralization of OP compounds means not only OPH mediated hydrolysis, the products

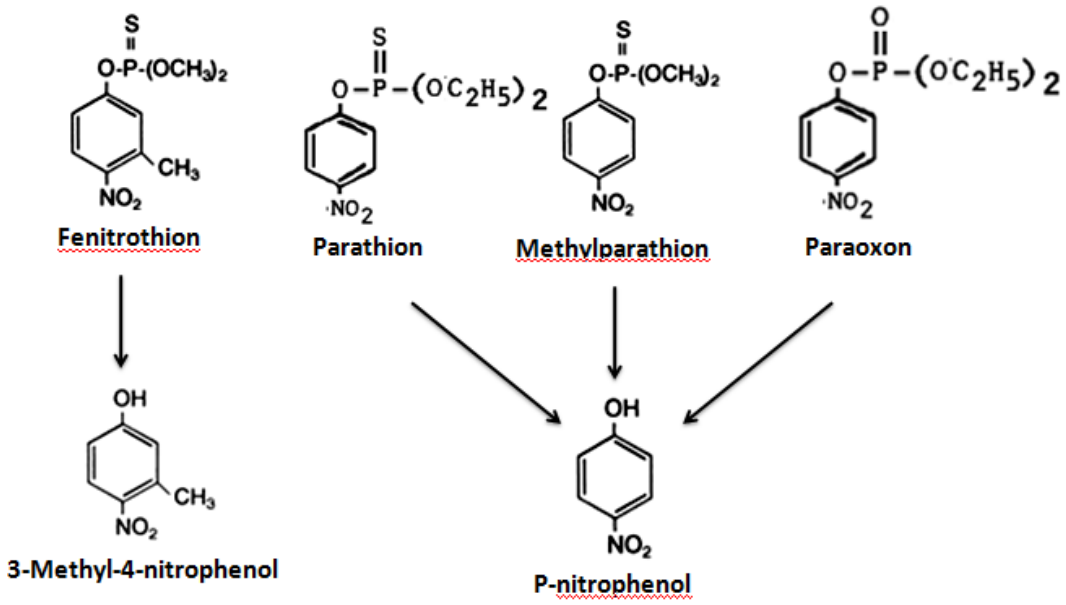


Fig. 5.1. Generation of nitrophenols through OPH mediated hydrolytic cleavage of certain OP_compounds.

generated through hydrolytic cleavage must be mineralized. In a very few cases the microbes are shown to possess genetic capability to mineralize OP compounds, due to existence of *opd* gene along with genes coding for enzymatic machinery needed for

degradation of aromatic compounds (Ou and Sharma, 1989; Rani and Lalithakumari, 1994; Keprasertsup et al, 2001). The alternative strategy is to mobilize *opd* containing native plasmids into native isolates having innate ability to mineralized aromatics and nirophenols. In our attempt to search for such native soil bacterial strain our laboratory has isolated an *Acinetobacter* sp. DS002 from a methyl parathion contaminated soil. When tested no *opd* gene was found in *Acinetobacter* sp. DS002. However, it has grown on a variety of aromatic compounds. Biodegradation of aromatic compounds is a well-studied aspect of biocatalysis. The degradation pathways operational in both gram-positive and gram-negative bacteria are well known (Harwood and Parales, 1996). Understanding of aromatic compound degradation pathway operational in *Acinetobacter* sp. DS002 is expected to generate basic information required for its manipulation towards achieving the mineralization of phenolic and nitrophenolic compounds. On careful examination, the isolate was shown to use a number of aromatic compounds and dicarboxylic acids as sources of carbon. Through comparative growth studies, the benzoate has been shown to serve as a better carbon source for *Acinetobacter* sp. DS002. Therefore aromatic degradation pathway found in *Acinetobacter* sp. DS002 was elucidated by growing it using benzoate as sole source of carbon. A combinatorial approach involving both metabolomic and proteomic tools were followed while elucidating the degradation pathway. Before actually analyzing the degradation pathways, the growth conditions were optimized and the cells grown in a physiological condition where the pathway enzymes were maximally induced.

5.1 Growth behavior of *Acinetobacter* sp. DS002 in benzoate

Initially, the optimal benzoate concentration for optimal growth of *Acinetobacter* sp. DS002 was determined. Growth was observed in all concentrations ranging from

5mM to 100mM benzoate. However, in high benzoate concentrations growth was seen only after 20 hours. A typical growth curve with lag, log and stationary phases were observed when grown in 5mM benzoate (Fig. 5.2). At higher concentrations of benzoate (50 mM) a typical diauxic growth curve with two exponential phases was observed. About 97% decrease in concentration of sodium benzoate was seen by the end of the logarithmic phase (Fig. 5.3).

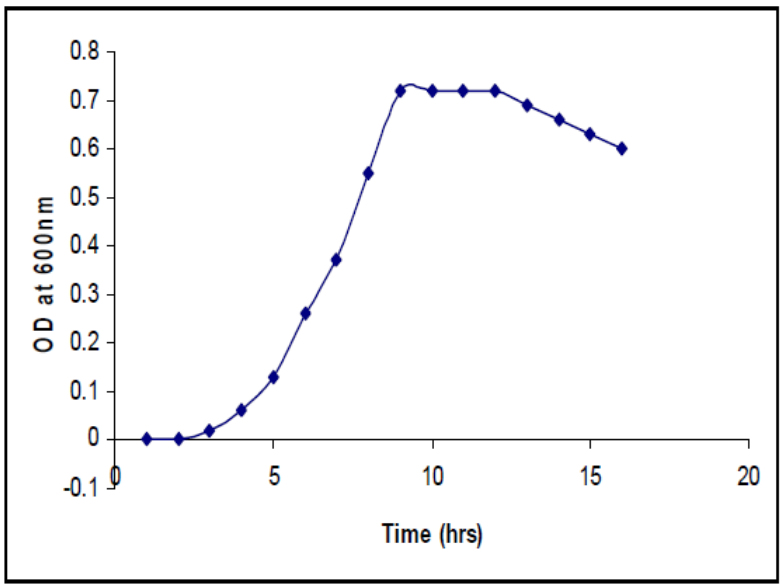


Fig. 5.2. Growth curve of *Acinetobacter* sp. DS002 in benzoate

When spent medium collected after eight hours of growth was analyzed on HPLC there was drastic reduction in benzoate peak with a concomitant appearance of additional peaks with retention times less than 4.4 min. The metabolites associated with these two new peaks were identified using LC/MS.

5.2 LC/MS analysis of catabolites

In order to identify the catabolic intermediates of benzoate, the spent medium collected from *Acinetobacter* sp. DS002 culture was extracted at different time intervals

following the procedures described in materials and methods section. When these extracts were separated on HPLC, three major peaks with retention times of (Fig. 5. 4) 1, 2.1 min (Fig. 5.5) and 3.9 min1.6 min (Fig. 5. 6) were observed. Though there were considerable differences in the individual peak intensities, they were constantly found from the extracts

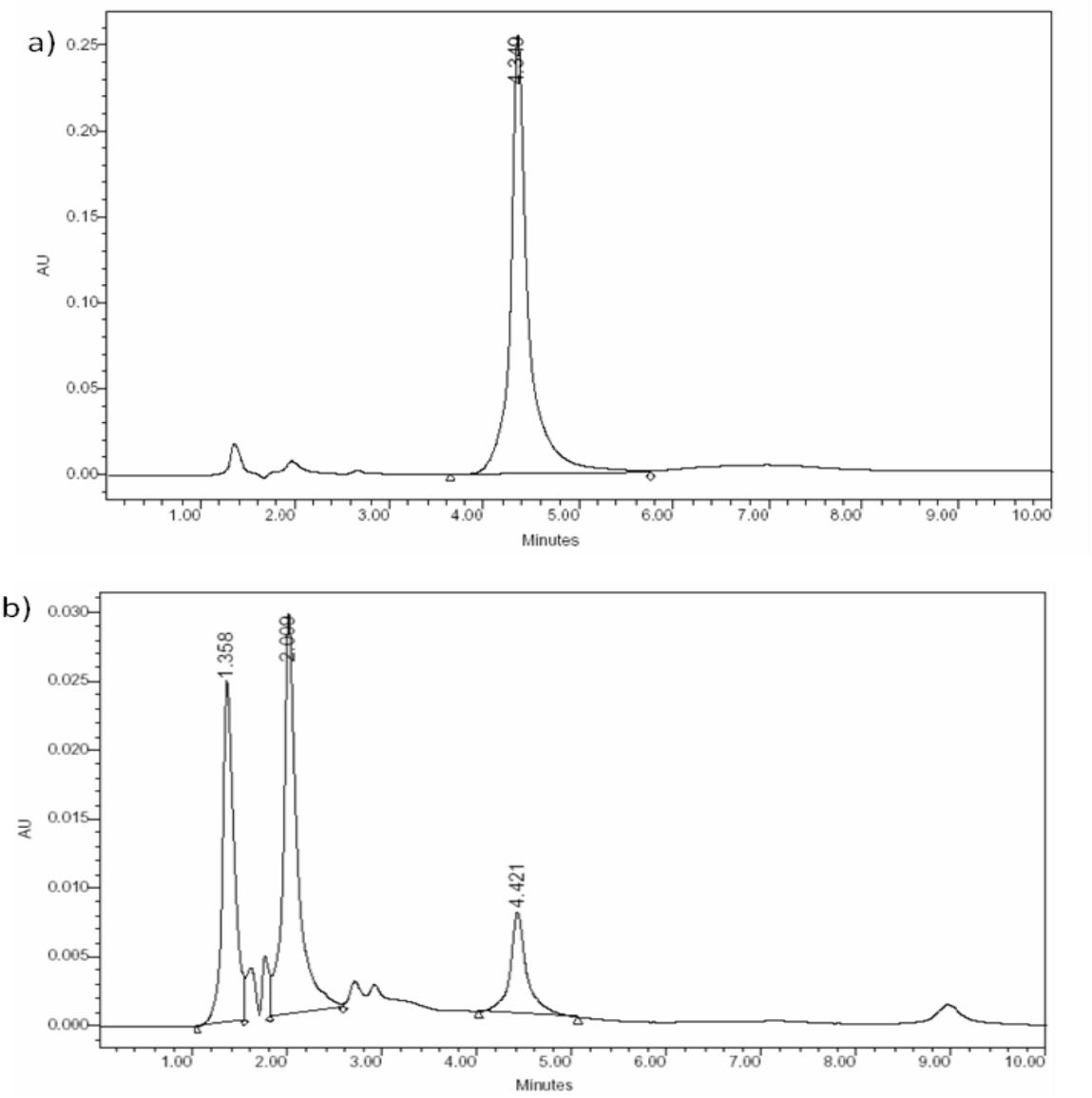


Fig. 5.3. HPLC analysis of catabolites at a) 0 hrs and b) 8 hrs. Peak with a retention time of 4.4 min represents benzoate.

prepared both at 8 h and 16 h time periods. These base peaks were then used for doing MS/MS in the negative mode to obtain molecular ion [M-] m/z values. The retention

times and molecular ions [M-] having the m/z values 141, 109 and 121 matched with m/z values of *cis, cis* -muconate, catechol and benzoate. As evidenced in the peak intensities, the concentration of the catabolic intermediates in the spent medium varied with time due to utilization of these compounds by *Acinetobacter* sp. DS002. The catabolite *cis,cis* muconate is the product formed from catechol by the action of catechol 1,2 dioxygenase, the first enzyme in the *ortho* cleavage pathway of catechol (Stanier et al, 1970).

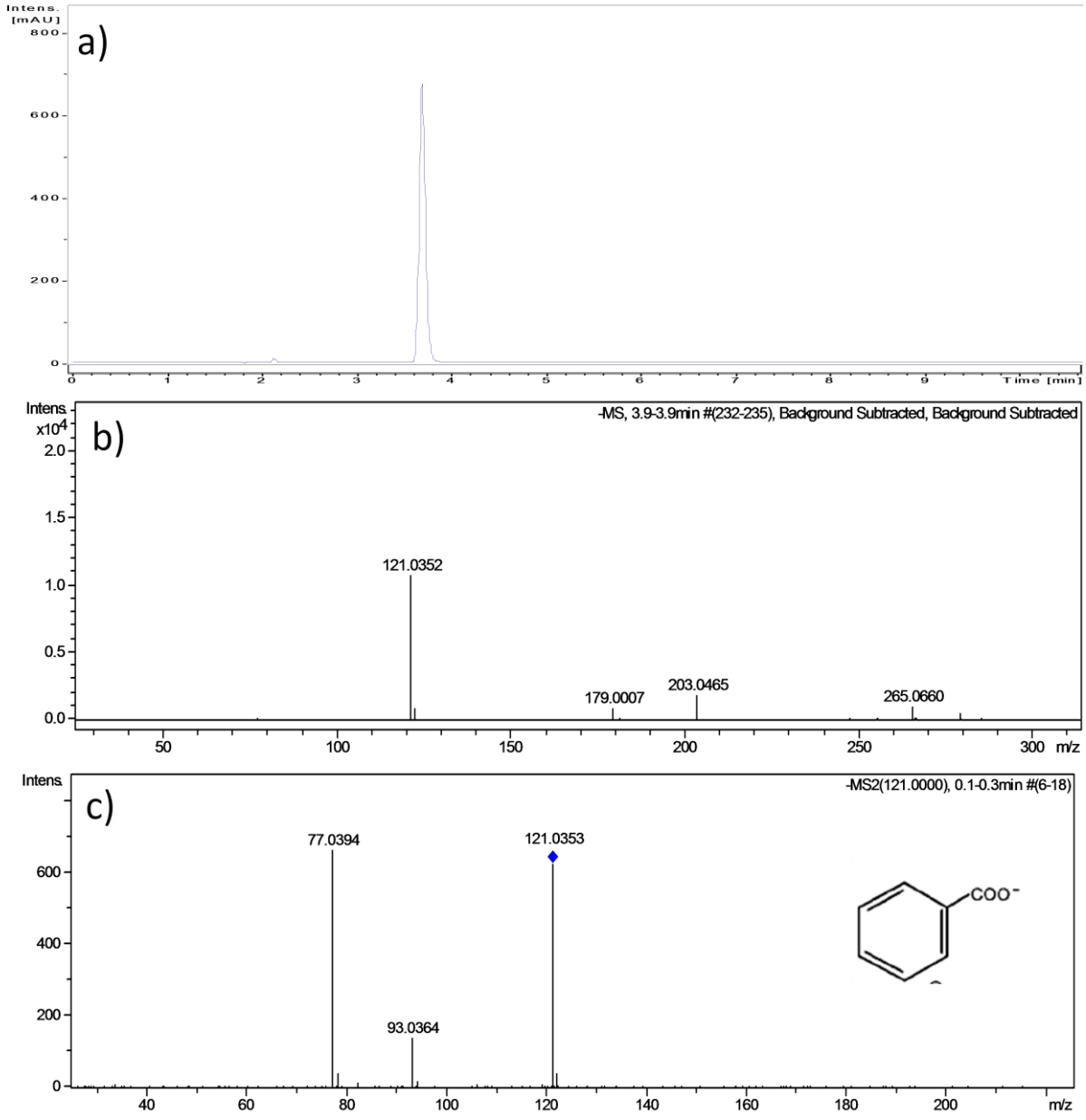


Fig. 5.4. LC-MS analysis of metabolites collected at 0 hrs time. Panels a, b and c represent LC profile, MS pattern and MS/MS pattern of a compound with retention time 3.9 min.

The initial step in the aerobic biodegradation of benzoate is incorporation of molecular oxygen into the aromatic nucleus by the enzyme benzoate 1, 2 dioxygenase to form a non-aromatic *cis-diol,* 2-hydro-1, 2-dihydroxybenzoate (DHB) which is further converted to catechol by the action of DHB dehydrogenase (Fig. 5)(Reiner, 1972; Reiner and

Hegeman, 1971). Catechol is cleaved either through the *ortho* pathway or *meta* pathway by the

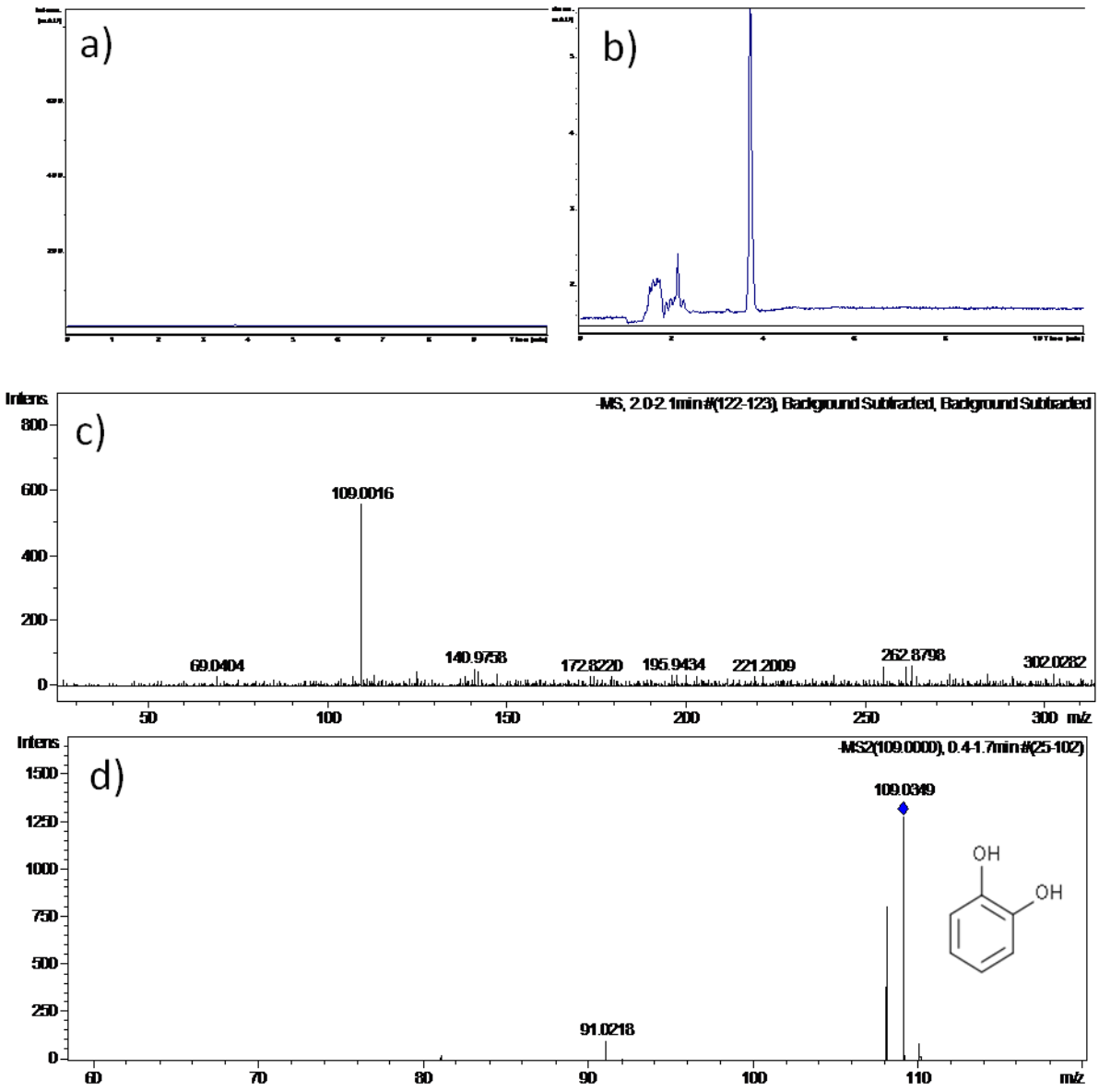


Fig. 5.5. LC-MS analysis of metabolites collected at 8 hrs time. Panels a, b, c and d represent LC profile, enlarged LC profile, MS pattern and MS/MS pattern of a compound with retention time 2.0 min.

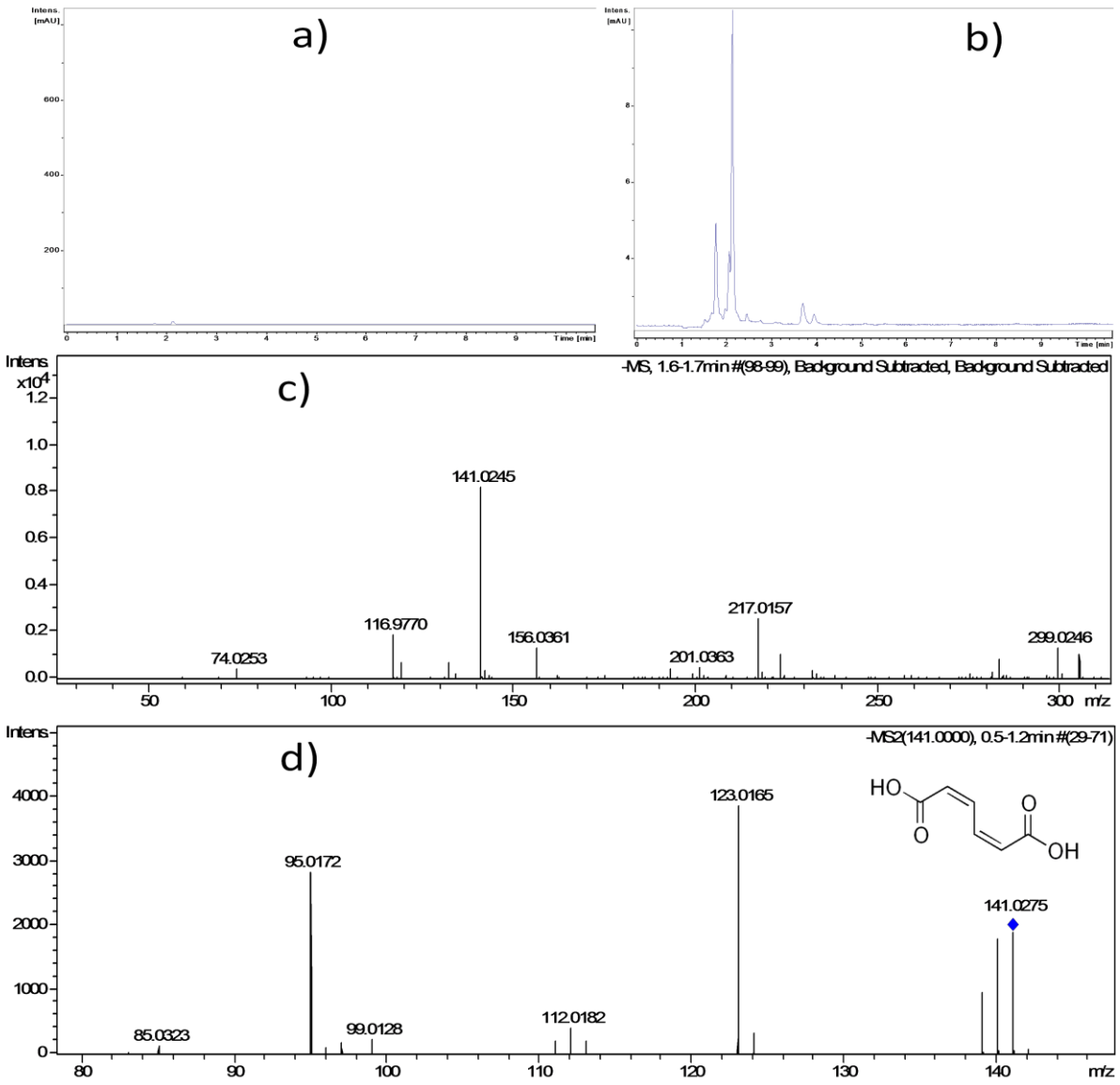


Fig 5.6. LC-MS analysis of metabolites collected at 8 hrs time. Panels a, b, c and d represent LC profile, enlarged LC profile, MS pattern and MS/MS pattern of a compound with retention time 1.6 min.

enzymes catechol 1,2 dioxygenase or catechol 2,3 dioxygenase leading to the formation of *cis-cis* muconate respectively and 2-hydroxymuconic semialdehyde (HMSA) respectively (Fig. 5. 7) (Loh and Chua, 2002). In the *ortho* pathway of catechol degradation, intradiol cleavage occurs leading to the formation of *cis,cis*- muconic acid (Fig. 5. 6) which is converted to β-ketoadipate-enol-lactone and finally to acetyl-CoA and

succinyl-CoA. In the *meta* cleavage pathway, extradiol cleavage of catechol occurs yielding 2-hydroxymuconic semialdehyde (HMSA) which is further converted to pyruvate and acetyl-CoA. Therefore, the existence of *cis, cis* –muconate and absence of HMSA in the metabolites indicate that degradation of benzoate in *Acinetobacter* sp. DS002 occurs through the *ortho* pathway.

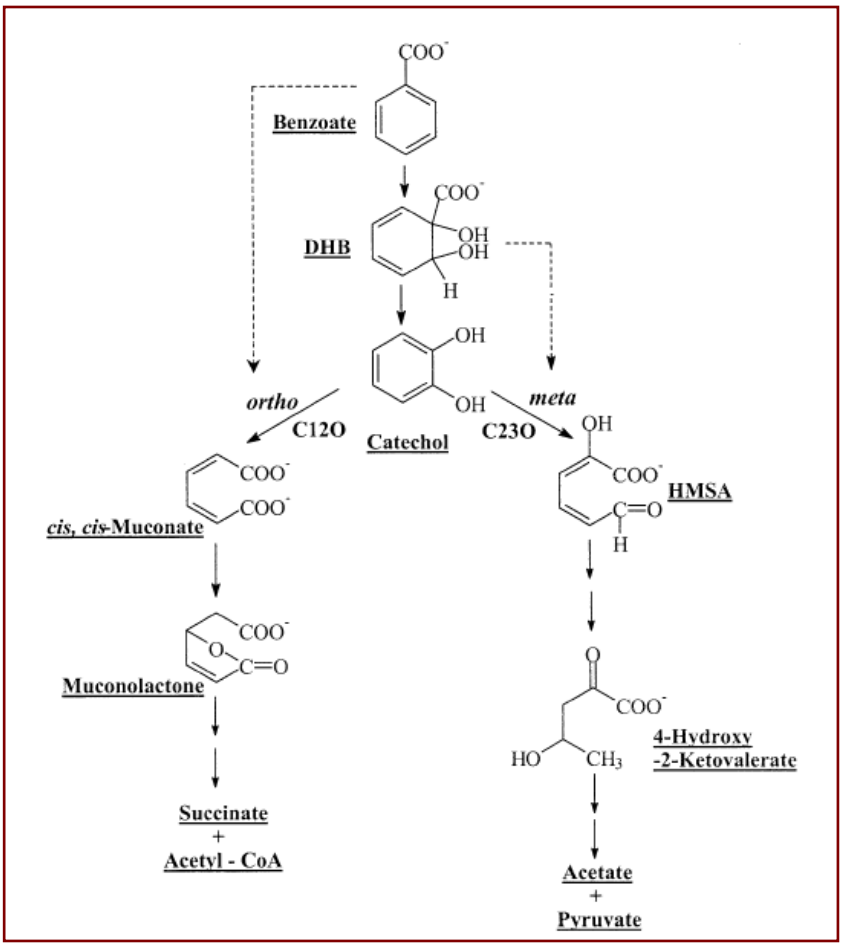


Fig. 5. 7. Schematic representation of o*rtho* and *meta* degradation pathways of benzoate.

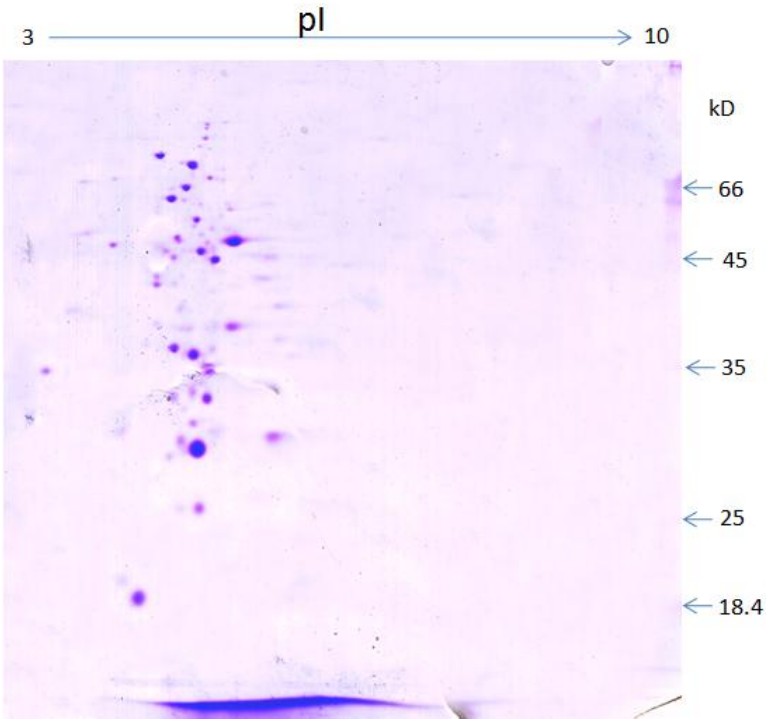
The metabolites generated in this study provide a clear understanding on pathway operational for aromatic compounds degradation in *Acinetobacter* sp. DS002.

5.3 Proteome analysis of *Acinetobacter* sp. DS002

In order to substantiate the data generated through identification metabolites parallel experiments were done to identify corresponding enzymes and to know their

regulation. Here in this study a proteomics approach was followed to gain information pertaining to i) upregulation of enzymes involved in degradation of benzoate ii) generate genome-wide expression profiling in response to shift in carbon source. In order to identify differently and differentially expressed proteins, initially the cells were grown in succinate where the enzymes involved in benzoate degradation are in highly repressed state. In this repressed state the proteins were extracted and basic proteome map was established by performing 2D electrophoresis as described in materials and methods section (Fig. 5. 8). The basic proteome map thus established was then compared with similar maps generated for the soluble proteins extracted from benzoate (5 and 50 mM) grown cultures (Fig.5.9, 5.10). Image analysis was performed using ImageMaster2D platinum software for normalization, spot detection, spot quantification, comparison of gels and for identification of differently and differentially expressed proteins due to shift in carbon source. Upon comparison of the proteome profiles, nearly 75 protein spots were found to be common both in benzoate (5 mM and 50mM) and succinate grown cultures (Fig. 5. 11). Further, the proteome profiles of *Acinetobacter* sp. DS002 grown in 5 mM and 50mM were almost identical matching more than 98.5%, except that the concentration dependent increase was seen in certain spots (Table. 5.1). Further, the protein spots that were showing significant intensities were picked for MALDI-TOF analysis. The 13 proteins that have shown significant score and sequence coverage with the proteins found in database were presented in table (Table 5. 1). Most of the proteins identified through MALDI-MS were essentially involved in degradation of benzoate via *ortho* pathway. Identification of electron transport component of benzoate 1,2 dioxygenase (spot 4344) and catechol 1,2 dioxygenase (spot 4364) through MS/MS, provided primary evidence to show that benzoate is degraded through *ortho* pathway in

Acinetobacter sp. DS002. Studies pertaining to benzoate degradation occasionally resulted



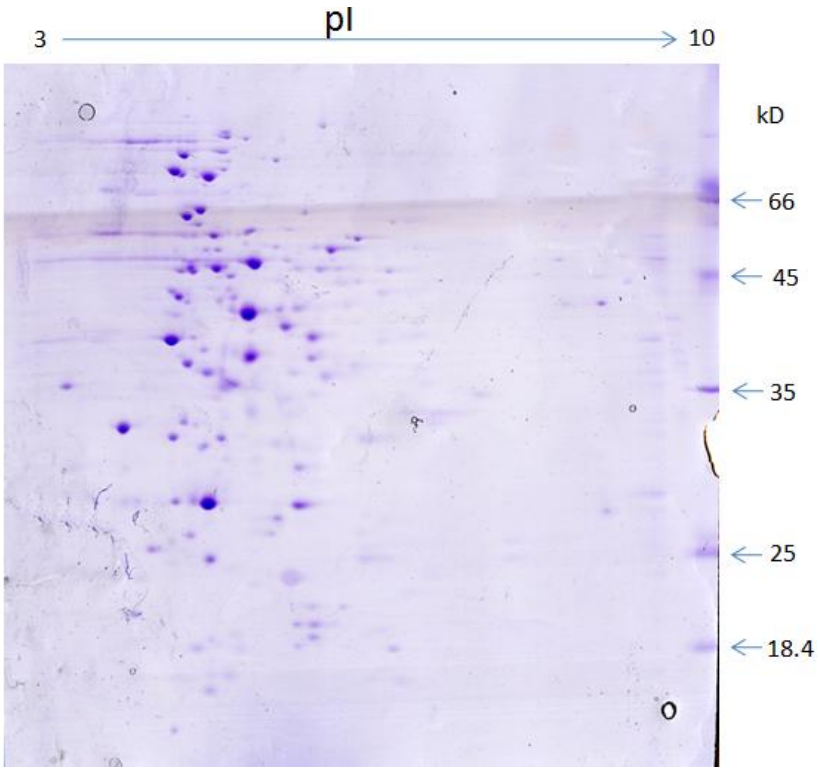


Fig. 5.8. 2D Proteome map generated for soluble proteins of *Acinetobacter* sp. DS002 grown in 10mM succinate

Fig. 5.9. 2D Proteome map generated for soluble proteins of *Acinetobacter* sp. DS002 grown in 5 mM benzoate

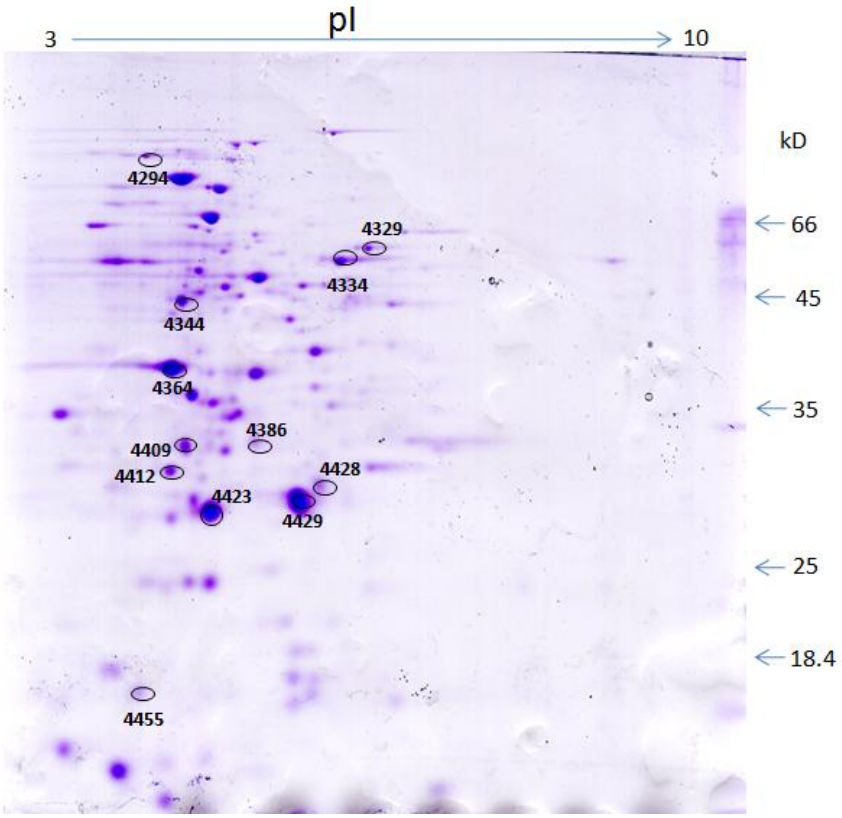


Fig. 5.10 2D Proteome map generated for soluble proteins of *Acinetobacter* sp. DS002 grown in 50 mM benzoate. Differently and differentially expressed spots that were subjected to MALDI-TOF were shown with an open circle. The number adjacent to the spot indicate the spot ID

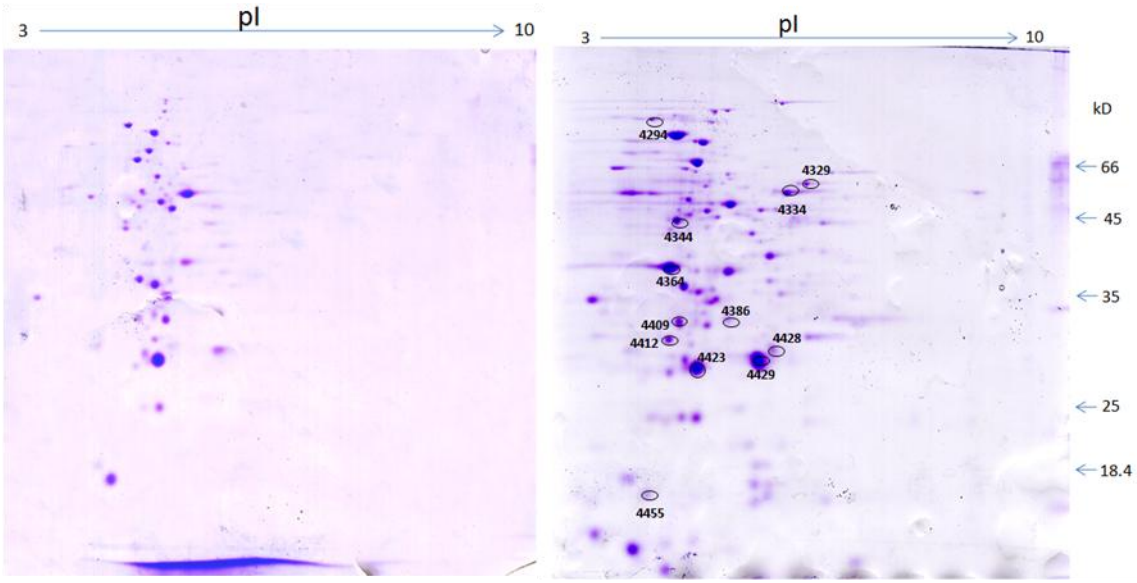
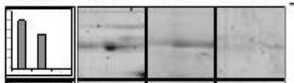


Fig. 5.11. Comparison of 2D proteome profiles of succinate (10 mM) and benzoate (50 mM) grown cultures of *Acinetobacter* sp. DS002. Differently and differentially expressed spots that were subjected to MALDI-TOF were shown with an open circle. The number adjacent to the spot indicate the spot ID.

in identification of isoforms of catechol 1,2 dioxygenase (Nakai et al, 1990). When searches were made at different pI points in the molecular mass range of 34kD no protein spots matched with mass fingerprint pattern of catechol 1,2 dioxygenase suggesting absence of its isoforms in *Acinetobacter* sp. DS002. Usually, benzoate and catechol degrading genes are organized as operons. Expression of these operons is strictly dependent on transcriptional regulator proteins such as BenM and CatM which belong to LysR family of transcriptional regulators (Collier et al, 1998; Romero-Arroyo et al, 1995). Interestingly the MS data of spot 4412 showed considerable similarity to LysR protein suggesting transcriptional regulation of benzoate degrading genes in *Acinetobacter* sp. DS002. Further, spots 4386 and 4428 have shown significant homology to ABC transport related proteins involved in unidirectional movement of a solute across the membranes (Saurin et al, 1999). Benzoate dependent induction of ABC transporters is rather unusual. The involvement of such transporters in unidirectional transport of

aromatic compounds such as phenols is well established (Kurbatov et al, 2006). Instead involvement of novel porin like proteins in benzoate transport is reported (Clark et al, 2002). Significant increase in the concentrations of ABC transporters only in benzoate grown cultures suggests existence of novel transport mechanism in *Acinetobacter* sp. DS002. However, further studies are required to validate this observation.

In addition to the aforementioned protein spots that are directly involved in catabolism of benzoate we have also seen a number of other protein spots in benzoate grown cultures that have no obvious link to benzoate catabolism. One of them is catalase (4294) and its induction very well correlates with the reports of catechol induced oxidative stress in microbes involved in biodegradation of aromatic compounds (Benndorf et al, 2001). The other two protein spots whose identity was established are

Spot No	Histogram	50 m mol L-1 Benzoate	5m mol L-1 Benzoate	10 m mol L-1 succinate	Protein ID	Mowse Score	% Sequence coverage*	Accession No
4294					Catalase *Acinetobacter baumannii* ACICU	76	20	gi|184157746
4329					Dihydrolipoamide dehydrogenase *Acinetobacter baumannii*	48	18	gi|169794948
4334					Hypothetical protein R2601_23970 *Roseovarius* sp. HTCC2601	77	23	gi|114766962
4344					Benzoate 1,2-dioxygenase electron transfer component *Acinetobacter baumannii* AYE	50	20	gi|169796581
4364					Catechol 1,2-dioxygenase *Acinetobacter baumannii*	89	36	gi|90018513
4386					ABC transporter related *Anaeromyxobacter* sp. Fw109-5	61	42	gi|153004317
4409					Putative adenylate/guanylate cyclase *Silicibacter* sp. TM1040	61	14	gi|99082352
4412					Transcriptional regulator, LysR family protein *Burkholderia mallei* PRL-20	44	80	gi|167004065
4423					Alkyl hydroperoxide reductase, C22 subunit, *Acinetobacter* sp.	83	48	gi|169633181
4428					ABC transporter related [*Burkholderia phymatum* STM815]	62	17	gi|186471060
4429					Major outer membrane protein P44-4 *Anaplasma phagocytophilum*	62	48	gi|19223945
4455					Hydroxylase for synthesis of 2-methylthio-cis-ribozeatin in tRNA *Klebsiella pneumoniae*	57	18	gi|152973133

* Percentage of protein sequence covered by the matching peptides

Table 5. 1. Magnified regions of 2D gel images with their histogram are represented in panel A. In the histogram the 1st, 2nd and 3rd bars represent protein intensities of a spot in 50 mM benzoate, 5 mM benzoate and 10 mM succinate grown cultures respectively. Panel B represent the hits obtained from MALDI TOF analyses and their accession numbers.

adenylate/guanylate cyclase (4409) and dihydrolipoamide dehydrogenase (4329). The role of cAMP in signaling mechanism is well known. In fact, it is part of the global switch that turns on several operons involved in metabolism of alternate carbon sources (Harman, 2001). Induction of adenylate/guanylate cyclase signifies synthesis of elevated cAMP/cGMP which might be needed for inducing genes required for benzoate catabolism. Dihydrolipoamide dehydrogenase (4329) and hydroxylase (4455) both are needed for operation of TCA cyles. One of them, dihydrolipoamide dehydrogenase plays a key role in formation of succinyl co-A (Kornfeld et al, 1977) and the enzyme hydroxylase has shown high homology to the product of *miaE* involved in conversion of of 2-methylthio-N-6 isopentenyl adenosine (ms2i6A) in to 2-methylthio-N-6 (cis-hydroxyl) isopentenyl adenosine ms2io6A). The ms2io6A is a modified base found adjacent to the anticodon of tRNAs that read codons beginning with "U". In *Salmonella typhimurium* presence of ms2io6A is a prerequisite to facilitate growth on TCA cycle intermediates such as succinate, fumerate or

Spot No	**MW of Peptides**	**Sequence**	**Protein**	**Accession No.**
4344	1044.494	FPWFEYR	Benzoate 1, 2-dioxygenase electron transfer component	gi|126641261
4344	2011.007	RSPGSGGLFSLAVNPYTCK	Flavodoxin/ Ferredoxin Oxidoreductase domain protein	gi|121996876
4364	1934.2730	RTIEGPLYVAGAPESVGFARM	Catechol 1,2 dioxygenase	Q43984
4364	2012.3720	KVEVWHANSLGNYSFFDKS	Catechol 1,2 dioxygenase	Q43984
4364	2081.3300	RHGNRPSHVHYFVSAPGYR.K	Catechol 1,2 dioxygenase	Q43984
4364	2780.7030	RKLTTQFNIEGDEYLWDDFAFATRD	Catechol 1,2 dioxygenase	Q43984

Table 5. 2. MS/MS hits obtained for protein spots 4344 and 4364. Sequence and MW of peptides that matched with either benzoate 1,2 dioxygenase or catechol 1,2 dioxygenase are provided.

malate (Persson et al, 1998). Upregulation of MaiE homologue in benzoate grown culture points towards enhanced biosynthesis of tRNA molecules having ms2io6A at 37th position. As end product of benzoate catabolism is a TCA intermediate, induction of MaiE during benzoate catabolism might be to enhance tRNA population with ms2io6A at 37th position which might be required for optimal operation of TCA cycle in *Acinetobacter* sp. DS002.

5.4 Cloning of *cat* operon

Organization of genes responsible for conversion of catechol to TCA intermediates was reported in several microbes (Harwood and Parales, 1996). A reverse genetic approach was employed in order to identify genes responsible for catechol degradation. As mentioned earlier, peptide mass fingerprint (PMF) of protein spot 4364 has shown significant similarity to catechol 1, 2 dioxygenase. The PMF hits were further confirmed by MS/MS analysis. The PMF of protein spot 4364 and its MS/MS data matched with catechol 1, 2 dioxygenase of *Acinetobacter* sp. (Q43984). A blast analysis was performed taking the *de novo* sequence of the protein spot 4364 and the catechol 1, 2 dioxygenase sequences of *Acinetobacter* genus found in the database. The alignment has shown high degree of sequence conservation throughout the length protein; except that the per cent homology was found to be little lower at the C-terminus of these proteins (Fig. 5. 12). Among these conserved regions two blocks that showed absolute sequence identity were identified. These two conserved sequence blocks (TPRTIEGPLYVAGA and DDFAFATRD) were taken to generate degenerate primers COF2 (5'-ACNCCNMGNACNATHGARGG-3') and COR1 (5'-CKNG TNGCRAANGCRAARTCRTC -3') for amplification of catechol 1, 2 dioxygenase gene from *Acinetobacter* sp. DS002 (Fig. 5. 12). When a typical PCR reaction was performed using genomic DNA of *Acinetobacter* sp. DS002 as template and COF2 and COR1 as primers an

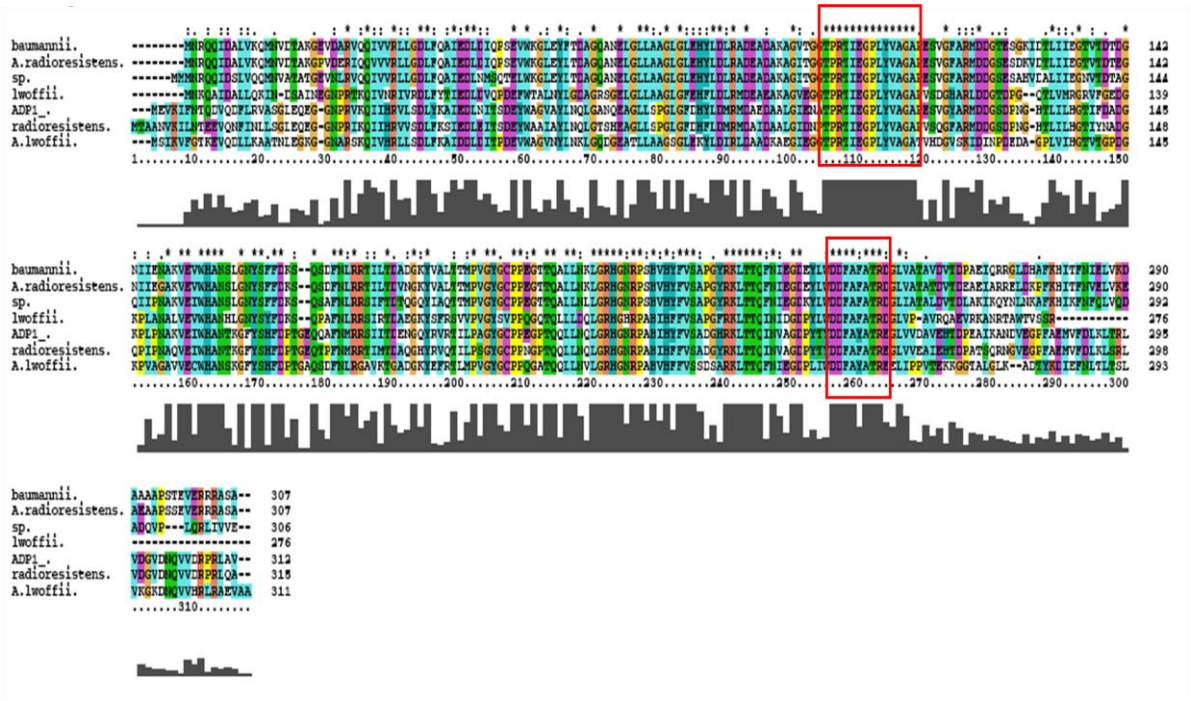


Fig 5. 12. Multiple alignment of various catechol 1,2 dioxygenases of the genus *Acinetobacter*. Conserved blocks of amino acids marked with red boxes were used to design degenerate primers.

amplicon of 500 bp was generated (Fig. 5. 13A). No amplicon was observed in the control reactions where genomic DNA of *E. coli* DH5α was used as template. Finally the identity of the 500 bp amplicon was confirmed by generating complete sequence. The sequence showed 98% identity with catechol 1,2 dioxygenase of *Acinetobacter baumannii* ATCC 17978 (Fig. 5. 14). The partial *catA* gene was then used as a probe to identify a fosmid clone containing complete *cat* / *ben* operons. When colony hybridization was performed five independent clones have given positive signals (Fig. 5.13B). Among these five positives, clone 912 was selected for further studies. In order to identify a fragment having the catechol 1,2 dioxygenase gene, a restriction profile of 912 clone was generated by digesting with restriction enzymes *Bam*HI, *Eco*RI, *Hind*III and *Sal*I . These fragments were again hybridized with partial *catA* gene to identify the profile that gives *cat* operon as a single restriction fragment. An 8kb *Sal*I fragment has been shown to have *cat* operon and was then

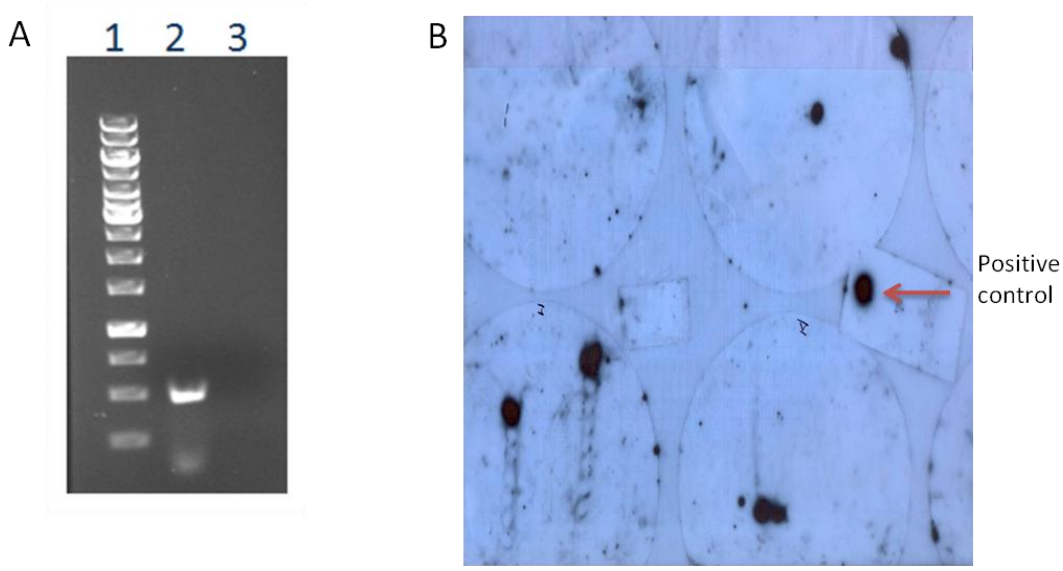


Fig. 5. 13. A) Amplification of *catA* gene using degenerate primers. Lane 1 represents 1 kb DNA ladder and lanes 2 and 3 represent amplicons obtained using genomic DNA of *Acinetobacter* sp. DS002 and *E.coli.* B*)* Identification of *catA* containing genomic clones by colony hybridization. The positive control is shown with an arrow mark.

```
>gb|CP000521.1|  Acinetobacter baumannii ATCC 17978, complete genome
Length=3976747

 Features in this part of subject sequence:
   CatA3

 Score =  789 bits (427),  Expect = 0.0
 Identities = 455/469 (98%), Gaps = 0/469 (0%)
 Strand=Plus/Plus

Query  71       TACGCCGCGGACGATCGAGGGTCCACTTTATGTTGCTGGCGCACCTGAATCAGTTGGCTT  130
                |||  ||  ||  ||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2147156  TACACCACGTACTATCGAAGGTCCACTTTATGTTGCTGGCGCACCTGAATCAGTTGGCTT  2147215

Query  131      TGCACGTATGGATGACGGAACCGAGACTGGCAAAATCGATACCTTAATTATTGAAGGTAC  190
                ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2147216  TGCACGTATGGATGACGGAACTGAGACTGGCAAAATCGATACCTTAATTATTGAAGGTAC  2147275

Query  191      GGTAACCGACACTGATGGCAATATTATTGAAAATGCCAAAGTTGAAGTATGGCATGCCAA  250
                |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2147276  GGTAACCGACACTAATGGCAATATTATTGAAAATGCCAAAGTTGAAGTATGGCATGCCAA  2147335

Query  251      CAGTTTAGGTAACTATTCATTCTTTGATAAGTCACAATCTGACTTTAACTTACGCCGTAC  310
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Sbjct  2147336  CAGTTTAGGTAACTATTCATTCTTTGATAAGTCACAATCTGACTTTAACTTACGTCGTAC  2147395

Query  311      CATTTTCACTGATGCAGATGGTAAATATGTAGCGTTAACCACTATGCCAGTTGGTTATGG  370
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2147396  CATTTTCACTGATGCAGATGGTAAATATGTAGCGTTAACCACTATGCCAGTTGGTTATGG  2147455

Query  371      TTGCCCTCCTGAAGGTACAACACAGGCTCTTCTTAACAAGTTAGGCCGTCATGGTAACCG  430
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2147456  ATGCCCTCCTGAAGGTACAACACAGGCTCTTCTTAACAAGTTAGGCCGTCATGGTAACCG  2147515

Query  431      TCCATCTCACGTTCACTACTTTGTATCTGCACCGGGTTACCGCAAGCTGACTACTCAATT  490
                ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Sbjct  2147516  TCCATCTCACGTTCACTACTTTGTATCTGCACCAGGTTACCGCAAGCTGACTACTCAATT  2147575

Query  491      CAACATTGAGGGTGATGAGTATTTATGGGACGACTTTGCGTTCGCAACT  539
                ||||||||||||||||||||  ||||||||  ||||||||  |||||  |||
Sbjct  2147576  CAACATTGAGGGTGATGAGTACTTATGGGATGACTTTGCCTTCGCTACT  2147624
```

Fig. 5. 14. BLASTN analysis of *catA* of *Acinetobacter* sp. DS002. Pairwise alignment of *catA* of *Acinetobacter* sp. DS002 with *catA3* of *Acinetobacter baumannii* 17978 is shown.

used for further studies by sub-cloning it in pBluescript vector (Fig. 5.15). Sequencing of the 8kb *Sal*I fragment of the clone 912 showed presence of genes which code for catechol 1, 2 dioxygenase, 3-oxo-acid-CoA transferase and β-ketoadipate thiolase (Fig. 5. 16).

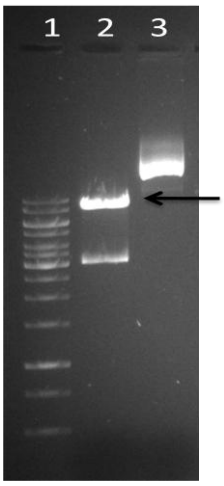


Fig. 5. 15. Sub-cloning of *catA* containing fragment into pBluescript II KS vector.Lane 1. represents 1kb DNA ladder. Lane 2 represents recombinant pBluescript plasmid having catA gene as a *Sal*I fragment. Lane 3 represents uncut recombinant plasmid.

Based on the identification of catabolic intermediates, PMF and MS/MS data and sequence information a pathway has been constructed for degradation of benzoate through the ortho pathway (Fig. 5. 16)

5.6 Purification of Catechol 1,2 dioxygenase

Catechol and substituted catechols occupy central position in biodegradation of aromatic compounds. Degradation of catechols further proceeds through either *ortho* or *meta* cleavage pathways depending on the availability of microbial dioxygenases. Catechol 1,2 dioxygenase, responsible for the intradiol cleavage of catechol, channels catechol degradation through ortho pathway, whereas catechol 2,3 dioxygenase diverts it towards *meta* cleavage pathway. The catachol 1, 2 dioxygenase is therefore a key enzyme in

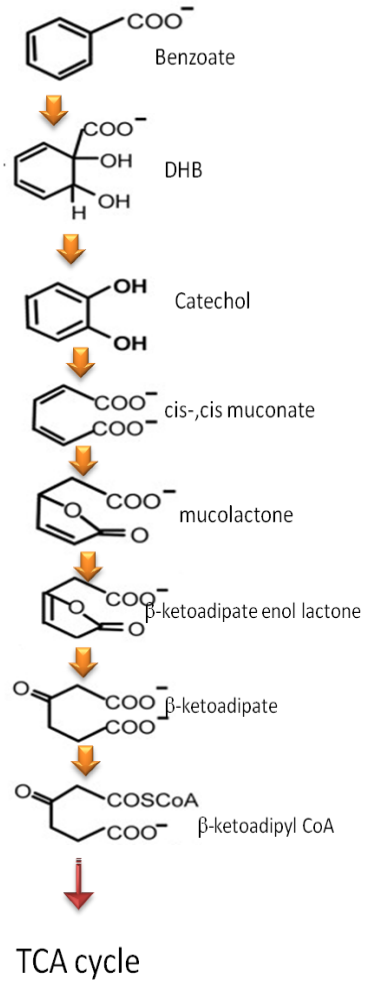


Fig 5. 16. Degradation pathway of benzoate in *Acinetobacter* sp. DS002

mineralization of catachols and substituted catechols. Phenolic and nirophenolic compounds generated during biodegradation of OP compounds will be converted to either catechols or nitrocatechols through the action of *p*-nitrophenol monoxygenase (Spain et al., 1979; Spain, 1994; Spain and Gibson, 1991; Zeyer and Kocher,1988). The nitrocatechol generated will be converted to benzentriol which is then serves as substrate for ring fission oxygneases (Hanne et al, 1993; Jain et al, 1994; Kadiyala and Spain, 1998; Chauhan et al, 2000). Alternatively, in certain cases the *p*-nitrophenol is converted to benzene triol *via* generation of hydroquinone. A detailed figure showing details of 4-nitrophenol degradation is given for quick reference (Fig. 5. 18). If fair assessment has to be made for further

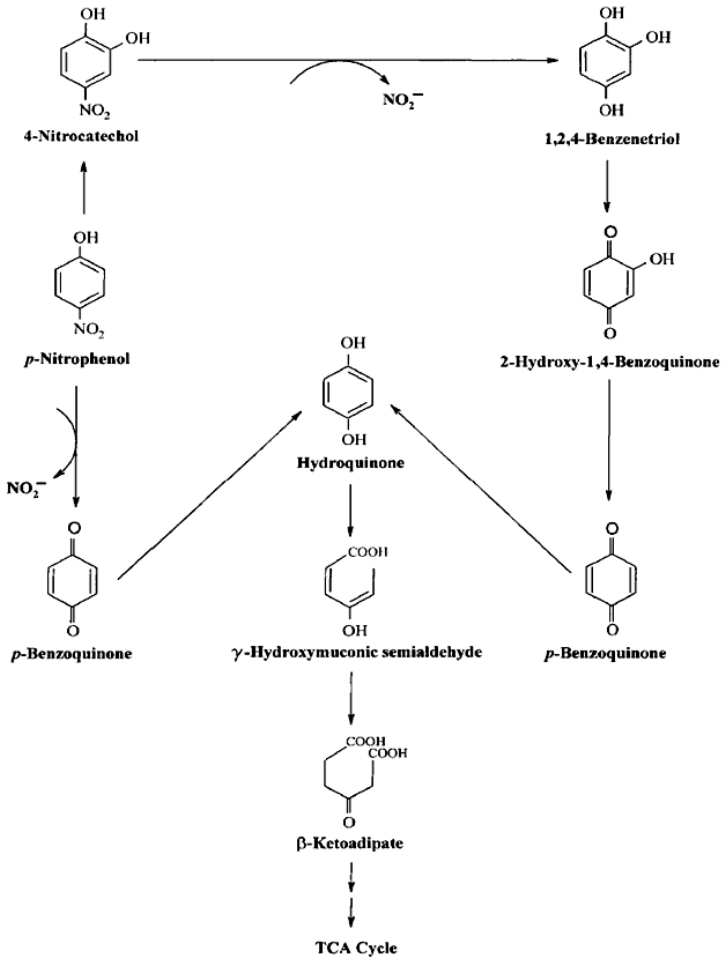


Fig. 5. 18. Degradation pathways of *p*-nitrophenol via 1,24 benzenetriol and benzoquinone

manipulation of *Acinetobacter* sp. DS002 it is necessary to assess the substrate specificity of catechol 1,2 dioxygenase using the intermediates of PNP degradation pathway as substrates. While attempting to do such experiments under *in vitro* conditions pure enzyme is necessary. Therefore an attempt was made to purify catechol 1,2 dioxygenase to electrophoretic homogeneity by following conventional protein purification techniques described in materials and methods section. The purification procedure gave 1 mg of protein from 20 g of cell pellet. Fractions containing more than 50% catechol 1,2 dioxygenase was taken at each stage of purification and the specific activity and fold purification was determined. Though the fold purification has increased with every stage of

purification, the gel permeation chromatography has given virtually homogenous catechol 1, 2 dioxygenase (Fig. 5. 19). Further, the protein fraction having activity showed a native mass of 66 kDa. The very same protein when analyzed on SDS-PAGE showed a molecular mass of 34 kDa (Fig. 5. 19d). Most of the C12Os with few exceptions are dimmers of identical or non-identical subunits (Aoki et al, 1984; Nakai et al, 1990). Gel permeation results indicate that C12O of *Acinetobacter* sp. DS002 is a homodimer with a molecular mass of 66 kDa.

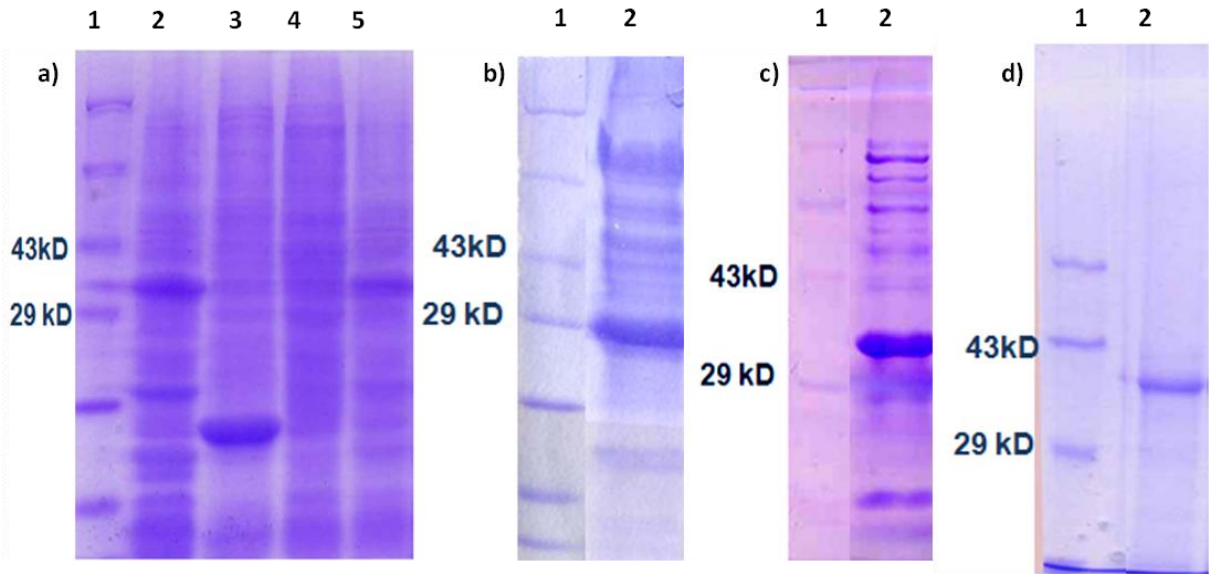


Fig. 5. 19. Purification of catechol 1, 2 dioxygenase of *Acinetobacter* sp. DS002. Panel a) represents ammonium sulphate fractionation. Lane 1 represents protein molecular weight marker. Lane 2 represents cytoplasmic fraction. Lanes 3-5 represent proteins obtained through 0-20% (Lane 3), 20-40% (Lane 4) and 40-60% (lane 5) saturation of ammonium sulphate. Panels b, c, d represent C12O purified using anion exchange (b), hydrophobic interaction (c) and gel permeation chromatography, respectively. In all panels lane 1 represents molecular weight markers, Lane 2 represents C12O. The C12O purified to electrophoretic homogeneity is seen in panel d.

5.7 Catechol 1, 2 dioxygenase assay

Substrate specificity was determined for C12O of *Acinetobacter* sp. DS002 using catechol and substituted catechols generated during OPH mediate hydrolysis of methyl parathion/ parathion degradation pathway by following procedures described in methods

section. In general the C12Os have showed relaxed specificity for catechols and methyl catechols but they have not shown any activity on nitro-catechols (Patel et al, 1976). However, C12O of *Acinetobacter* sp. DS002 has shown considerable activity when 1, 2, 4-benzenetriol and 4-nitrocatechol were used as substrate (Fig. 5. 20). Existence of C12Os with relaxed substrate specificity is not uncommon in literature. C12Os showing activity on methyl and halocatechols were seen in *Rhodococcus*, *Ralstonia* and *Pseudomonas arvilla* (Cha, 2006; Patel et al, 1997; Briganti et al, 1976; Briganti et al, 2000; Wang et al, 2006). The substituted phenols like 4-nitrophenol, 4-nitrocatechol and 1,2,4 benzenetriol are degradation products of methyl parathion and parathion (Pakala et al, 2007; Chauhan et al, 2000; Jain et al, 1994; Kadiyala and Spain, 1998). As benzenetriol and 4-nitrocatechol served

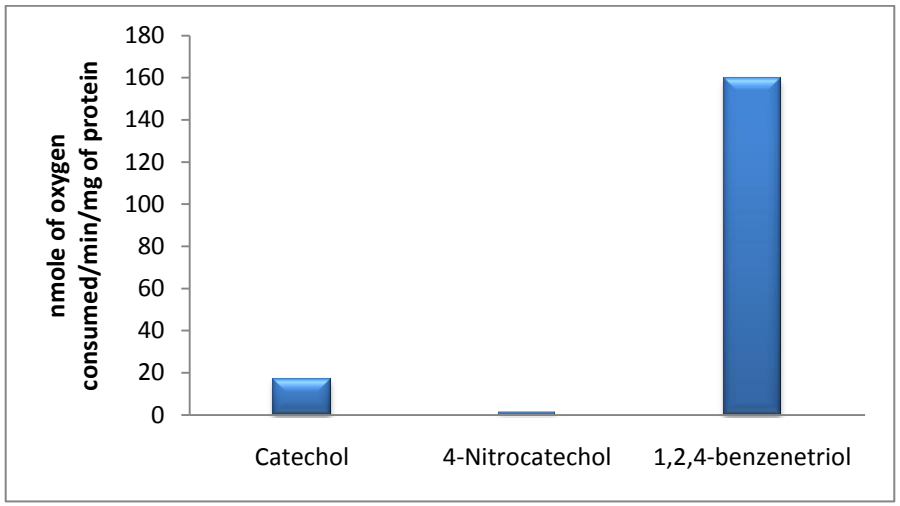


Fig. 5. 20. Activity of catechol 1,2 dioxygenase on catechol, 4-nitrocatechol and 1,2,4-benzenetriol respectively.

as substrate for C12O of *Acinetobacter* sp. DS002 channeling of these PNP intermediates appears to be possible through the *ortho* pathway.

5.8 Manipulation of *Acinetobacter* sp. DS002

Though *Acinetobacter* sp. DS002 was isolated from OP-polluted soils existence of *opd* gene, which codes for organophosphate hydrolase (OPH) involved hydrolytic cleavage of structurally diverse group of OP compounds was not apparent. The *Acinetobacter* strains were not tested positive for OPH activity. In agreement of this observation no amplification was seen when PCR was performed using *opd* specific primers. If the relaxed substrate specificity of C12O has to be exploited for complete mineralization of OP-compounds, expression of *opd* gene in *Acinetobacter* sp. DS002 is inevitable (Cha, 2006; Patel et al, 1997; Briganti et al, 1976; Briganti et al, 2000; Wang et al, 2006). The OPH activity alone can generate nitrophenols from OP-compounds, which serve as substrate for C12O. Therefore in the present study an attempt was made to mobilize a derivative of organophosphate degrading (*opd*) plasmid, pPDL2 of *Flavobacterium* sp. ATCC 27551 into *Acinetobacter* sp. DS002.

5.9 Mobilization of pPDL2 Tn*5*<R6Kγori/KAN-2>into *Acinetobacter* sp. DS002

As described in the previous chapter plasmid pPDL2 is a mobilizable plasmid. The derivative of pPDL2 generated by inserting mini-transposon Tn*5*<R6Kγori/KAN-2> was shown to be mobilizable. Therefore a triparental mating experiment was performed to transfer pPDL2:: Tn*5*<R6Kγori/KAN-2> into the aromatic compound degrading *Acinetobacter* sp. DS002. After mobilization of pPDL2- Tn*5*<R6Kγori/KAN-2> into *Acinetobacter* sp. DS002, its stability was frequently tested by monitoring the presence of PCR using *opd* specific primers (Fig. 5. 21). Interestingly plasmid pPDL2::Tn5<R6Kγori/Kan2> was found to be highly stable. Even in the absence of selection on kanamycin the plasmid was found to be highly stable. In consistence of its presence high amounts of OPH activity was found in all

Acinetobacter sp. DS002 (pPDL2::Tn*5*<R6Kγ*ori*/KAN-2>) clones (Fig. 5. 22).After establishing stable maintenance of plasmid pPDL2::Tn5<R6Kγ*ori*/Kan2> in *Acinetobacter* sp. DS002,

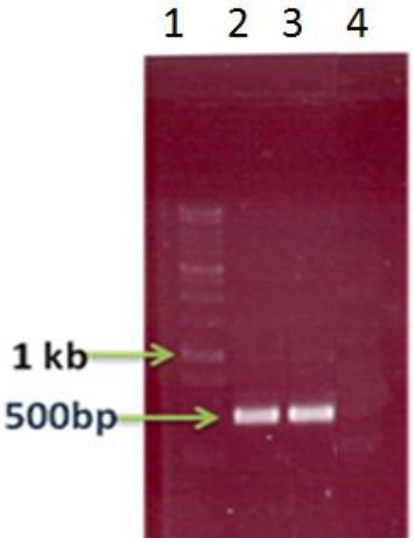


Fig. 5. 21. Confirmation of pPDL2 mobilization into *Acinetobacter* sp. DS002 by colony PCR using *opd* specific primers.Lane 1 represents 1 kb DNA ladder. Lanes 2-3 represent amplicons obtained from pPDL2::Tn5<R6Kγori/Kan2> harbouring *Acinetobacter* sp. DS002 and *E.coli pir*116 respectively. Lane 4 represents *E.coli pir*-116.

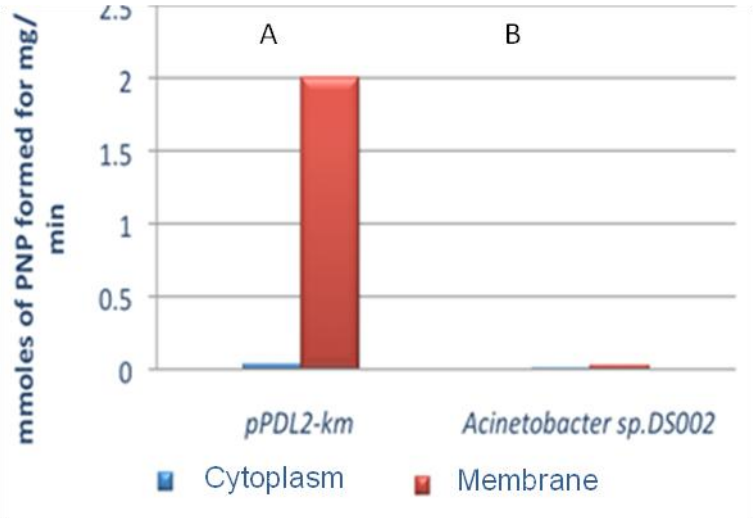


Fig. 5. 22. Assay of Phosphotriesterase activity in cytoplasmic and membrane fractions of A) *Acinetobacter* sp. DS002 having plasmid pPDL2::Tn5<R6Kγ/Kan2> and B) wild type *Acinetobacter* sp. DS002.

experiments were conducted to test their ability to degrade OP compound, paraoxon. Our lab has recently shown presence of OPH in the inner membrane of *B. diminuta* and its

dependence on Twin Arginine Transport (Tat) pathway for membrane targeting (Gorla et al, 2009). Therefore the *Acinetobacter* cells harboring plasmid pPDL2:: Tn*5*<R6Kγori/KAN-2>were fractionated to cytoplasmic and particulate fractions and were assayed for OPH activity as described in methods. Most of the OPH activity was found in membrane fraction and very little activity was seen cytoplasmic fraction, (Fig. 5. 22).

5.10 Degradation of methyl parathion

After establishing the stability and expression of OPH in *Acinetobacter* sp. DS002 (pPDL2:: Tn*5*<R6Kγ*ori*/KAN-2>), the manipulated strain was tested for its ability to degrade organophosphates like methyl parathion and parathion and their catabolic intermediates such as 4-nitrocatechol and 1, 2, 4-benzenetriol generated during their biodegradation. Immediately after adding parathion to the culture medium containing *Acinetobacter* sp. DS002 (pPDL2:: Tn*5*<R6Kγ*ori*/KAN-2>), it quickly turned into yellow colour indicating OPH mediated hydrolysis of methyl parathion. Such change of colour was not observed in control cultures having wild type strains of *Acinetobacter* sp. DS002. Subsequently the yellow colour of the medium generated due to the formation of *p*-nitrophenol from methyl parathion got slowly disappeared from the culture medium. When the culture medium was extracted for identification of *p*-nitrophenol metabolites both 4-nitrocatechol and 1,2,4-benzenetriol were identified indicating that the *p*-nitrophenol generated due to OPH mediated hydrolytic cleavage is further metabolized to generate nitrocatechol and 1,2,4-benzenetriol . As shown in the aforementioned sections benzenetriol has served as one of the substrates for the ring cleavage enzyme C12O purified from *Acinetobacter* sp. DS002. Such observation supports channelization of benzenetriol into TCA cycle indicating the possibility of *p*-nitrophenol serving as carbon source in *Acinetobacter* sp. DS002 (pPDL2:: Tn*5*<R6Kγori/KAN-2>). As shown in figure 5.18 PNP monooxygenase plays a critical role in conversion of PNP into 4-

nitrocatechol / 1, 2, 4-benzenetriol. The total sequence of pPDL2 presented in the first chapter gave no indication of PNP monooxygenase. As formation of nitrocatechol and 1, 2, 4-benzenetriol were found in the culture medium of *Acinetobacter* sp. DS002 (pPDL2:: Tn*5*<R6Kγori/KAN-2>) it suggests existence of such monooxygenases on the chromosome of *Acinetobacter* DS002. *Acinetobacter* sp. DS002 is a soil isolate and shown to grow on number of aromatic compounds. Though PNP at higher concentrations is shown to be toxic to *Acinetobacter* sp. DS002, it has supported for the growth of the strain at low concentrations. In the light of these observations degradation of PNP generated from OP-compounds is an understandable consequence. A schematic degradation pathway found to be operational in *Acinetobacter* sp. DS002 has been presented in Fig. 5. 23.

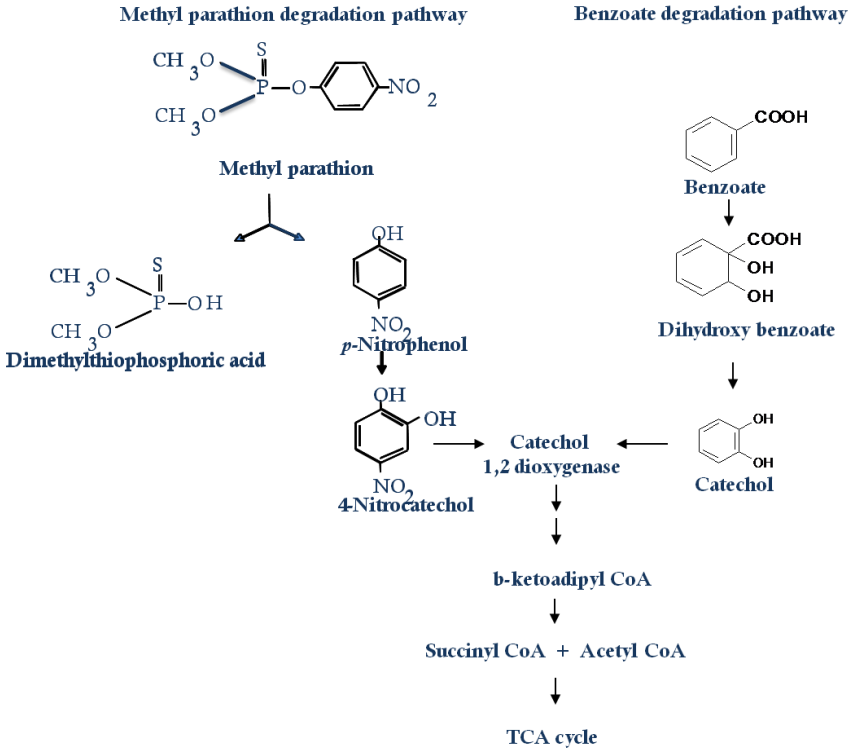


Fig. 5. 23. Schematic representation of methyl parathion and benzoate degradation pathways in *Acinetobacter* sp. DS002 (pPDL2:: Tn*5*<R6K*γori*/KAN-2>)

Conclusions

1. *Acinetobacter* sp. DS002 utilizes benzoate as sole source of carbon. The optimum concentration of benzoate for growth is 5 mM.
2. Catechol and *cis,cis*-muconic acid were identified as intermediates of benzoate catabolism in *Acinetobacter* sp. DS002.
3. Basic proteome maps of soluble proteins of *Acinetobacter* sp. DS002 grown in 10 mM succinate, 5 mM and 50 mM benzoate were generated.
4. Proteome maps of 5 mM and 50 mM benzoate grown cultures were identical whereas the proteome map of benzoate grown cultures sowed significant differences with the proteome map of succinate grown cultures.
5. Benzoate 1,2 dioxygenase and catechol 1,2 dioxygenase, key enzymes of benzoate degradation were identified through MALDI-TOF and MS/MS analysis.
6. Degenerate primers were designed for *catA* gene and *catA* gene amplified from *Acinetobacter* sp. DS002 showed 98% identity with *catA* gene of *Acinetobacter baumannii* 17978.
7. Genomic library clones of *Acinetobacter* sp. DS002 with *cat* operon were identified. The *cat* operon was sub-cloned and sequenced.
8. Catechol 1,2 dioxygenase (C12O) of *Acinetobacter* sp. DS002 was purified to electrophoretic homogeneity. C12O shows significant enzyme activity towards PNP intermediates 4-nitrocatechol and 1, 2, 4- benzenetriol.
9. Plasmid pPDL2:: Tn*5*<*R6Kγori*/KAN-2> was successfully mobilized into *Acinetobacter* sp. DS002. *Acinetobacter* (pPDL2::Tn*5*<*R6Kγori*/KAN-2>) has successfully degraded OP compounds Paraoxon and methyl parathion.

Abremski, K. E. & Hoess, R. H. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Protein Eng* **5**, 87-91 (1992).

Adriaens, P., Kohler, H. P., Kohler-Staub, D. & Focht, D. D. Bacterial dehalogenation of chlorobenzoates and coculture biodegradation of 4,4'-dichlorobiphenyl. *Appl Environ Microbiol* **55**, 887-892 (1989).

Afriat, L., Roodveldt, C., Manco, G. & Tawfik, D. S. The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. *Biochemistry* **45**, 13677-13686, doi:10.1021/bi061268r (2006).

Allpress, J.D., Gowland, C. P. (2010) Dehalogenases: Environmental defence mechanism and model of enzyme evolution. DOI: 10.1016/S0307-4412(98)00090-9

Alting-Mees, M. A. & Short, J. M. pBluescript II: gene mapping vectors. *Nucleic Acids Res* **17**, 9494 (1989).

Antunes, P., Machado, J., Sousa, J. C. & Peixe, L. Dissemination of sulfonamide resistance genes *(sul1, sul2*, and *sul3*) in Portuguese *Salmonella enterica* strains and relation with integrons. *Antimicrob Agents Chemother* **49**, 836-839, doi:10.1128/AAC.49.2.836-839.2005 (2005).

Aoki K, Konohana T, Shinke R, Nishira H. Two catechol 1,2-dioxygenases from aniline-assimilating bacterium, *Frateuria* species ANA-18. *Agric Biol Chem* **48.** 2097–2104 (1984).

Aravind, L. An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico Biol* **1**, 69-91 (1999).

Arber, W. Biological evolution: lessons to be learned from microbial population biology and

genetics. *Res in microbio* **155**, 297-300, doi:10.1016/j.resmic.2004.01.009 (2004).

Arber, W. Evolution of prokaryotic genomes. *Gene* **135**, 49-56 (1993).

Arber, W. Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS microbiol rev* **24**, 1-7 (2000).

Arber, W. The evolutionary strategy of DNA acquisition as a possible reason for a universal genetic code. *Hist Philos Life Sci* **28**, 525-532 (2006).

Arber, W. The evolutionary strategy of DNA acquisition as a possible reason for a universal genetic code. *Hist Philos Life Sci* **28**, 525-532 (2006).

Argos, P. *et al.* The integrase family of site-specific recombinases: regional similarities and global diversity. *Embo J* **5**, 433-440 (1986).

Armengaud, J., Happe, B. & Timmis, K. N. Genetic analysis of dioxin dioxygenase of *Sphingomonas sp*. Strain RW1: catabolic genes dispersed on the genome. *J Bacteriol* **180**, 3954-3966 (1998).

Ashelford, K. E., Day, M. J. & Fry, J. C. Elevated abundance of bacteriophage infecting bacteria in soil. *Appl Environ Microbiol* **69**, 285-289 (2003).

Assinder, S. J. & Williams, P. A. The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Microb Physiol* **31**, 1-69 (1990).

Aukema, K. G., Kron, E. M., Herdendorf, T. J. & Forest, K. T. Functional dissection of a conserved motif within the pilus retraction protein PilT. *J Bacteriol* **187**, 611-618, doi:10.1128/JB.187.2.611-618.2005 (2005).

Austin, S. & Abeles, A. Partition of unit-copy mini-plasmids to daughter cells. I. P1 and F mini-plasmids contain discrete, interchangeable sequences sufficient to promote

equi-partition. *J Mol Biol* **169**, 353-372 (1983).

Austin, S. & Abeles, A. The partition functions of P1, P7, and F miniplasmids. *Basic Life Sci* **30**, 215-226 (1985).

Austin, S. & Abeles, A. The partition functions of P1, P7, and F mini-plasmids. *Basic Life Sci* **30**, 215-226 (1985).

Bahl, M. I., Burmolle, M., Meisner, A., Hansen, L. H. & Sorensen, S. J. All IncP-1 plasmid subgroups, including the novel epsilon subgroup, are prevalent in the influent of a Danish wastewater treatment plant. *Plasmid* **62**, 134-139, doi:10.1016/j.plasmid.2009.05.004 (2009).

Bahl, M. I., Hansen, L. H., Licht, T. R. & Sorensen, S. J. Conjugative transfer facilitates stable maintenance of IncP-1 plasmid pKJK5 in *Escherichia coli* cells colonizing the gastrointestinal tract of the germfree rat. *Appl Environ Microbiol* **73**, 341-343, doi:10.1128/AEM.01971-06 (2007).

Barbour, A. G. & Garon, C. F. Linear plasmids of the bacterium *Borrelia burgdorferi* have covalently closed ends. *Science* **237**, 409-411 (1987).

Basta, T., Keck, A., Klein, J. & Stolz, A. Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading *Sphingomonas strains*. *J Bacteriol* **186**, 3862-3872, doi:10.1128/JB.186.12.3862-3872.2004 (2004).

Beaber, J. W., Hochhut, B. & Waldor, M. K. SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**, 72-74, doi:10.1038/nature02241 (2004).

Bellanger, X. *et al.* Conjugative transfer of the integrative conjugative elements ICESt1 and ICESt3 from *Streptococcus thermophilus*. *J Bacteriol* **191**, 2764-2775,

doi:10.1128/JB.01412-08 (2009).

Bellanger, X., Morel, C., Decaris, B. & Guedon, G. Derepression of excision of integrative and potentially conjugative elements from *Streptococcus thermophilus* by DNA damage response: implication of a *cI*-related repressor. *Journal of bacteriology* **189**, 1478-1481, doi:10.1128/JB.01125-06 (2007).

Bellanger, X., Morel, C., Decaris, B. & Guedon, G. Regulation of excision of integrative and potentially conjugative elements from *Streptococcus thermophilus*: role of the *arp1* repressor. *Journal of molecular microbiology and biotechnology* **14**, 16-21, doi:10.1159/000106078 (2008).

Benndorf, D., Loffhagen, N. and Babel, W. Protein synthesis patterns in *Acinetobacter calcoaceticus* induced by phenol and catechol show specificities of responses to chemostress. *FEMS Microbiol Lett*. **200.** 247-52 (2001).

Bennett PM. Genome plasticity *Methods in Molecular Biology, Vol 266, Genomics, Proteomics and Clinical Bacteriology*. Humana Press Inc.: Totowa, NJ; 71–113.In: Woodford N, Johnson A (eds) (2005).

Bennett, P. M. Integrons and gene cassettes: a genetic construction kit for bacteria. *J Antimicrob Chemother* **43**, 1-4 (1999).

Bennett, P. M. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br J Pharmacol* **153 Suppl 1**, S347-357, doi:10.1038/sj.bjp.0707607 (2008).

Benning, M. M., Kuo, J. M., Raushel, F. M. & Holden, H. M. Three-dimensional structure of phosphotriesterase: an enzyme capable of detoxifying organophosphate nerve agents.

Biochemistry **33**, 15001-15007 (1994).

Berg, D.E., and M.M. Howe, Mobile DNA. American Society of Microbiology Publication, Washington, D.C. (1989)

Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* **226**, 735-745 (1992).

Bhat, R.V. and Krishnamachary, K.A.V.R. Endemic familial arthritis of Malnad -An Epidemiological study. *Ind J Med Res* **66**. 777-785 (1977).

Bignell, C. & Thomas, C. M. The bacterial ParA-ParB partitioning proteins. *J Biotechnol* **91**, 1-34 (2001).

Bird SB, Sutherland TD, Gresham C, Oakeshott J, Scott C, and Eddleston M. *OpdA*, a bacterial organophosphorus hydrolase, prevents lethality in rats after poisoning with highly toxic organophosphorus pesticides. Toxicology 247(2-3): 88-92 (2008).

Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**, 1513-1523 (1979).

Boltner, D. & Osborn, A. M. Structural comparison of the integrative and conjugative elements R391, pMERPH, R997, and SXT. *Plasmid* **51**, 12-23 (2004).

Bonheyo, G. T., Hund, B. D., Shoemaker, N. B. & Salyers, A. A. Transfer region of a Bacteroides conjugative transposon contains regulatory as well as structural genes. *Plasmid* **46**, 202-209, doi:10.1006/plas.2001.1545 (2001).

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

Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of

protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254 (1976).

Bramhill, D. & Kornberg, A. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52**, 743-755 (1988).

Briganti F, Pessione E, Giunta C, Mazzoli R, Scozzafava A. Purification and catalytic properties of two catechol 1,2- dioxygenase isozymes from benzoate grown cells of *Acinetobacter radioresistens*. *J Protein Chem* **19**. 709–716 (2000).

Brochet, M. *et al.* Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. *Mol Microbiol* **71**, 948-959, doi:10.1111/j.1365-2958.2008.06579.x (2009).

Brochet, M., Couve, E., Glaser, P., Guedon, G. & Payot, S. Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. *J Bacteriol* **190**, 6913-6917, doi:10.1128/JB.00824-08 (2008).

Broda, P. Plasmids. Oxford &co San Francisco: W, H. Freeman (1979).

Brussow, H., Canchaya, C. & Hardt, W. D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Molecul Biol rev* **68**, 560-602, doi:10.1128/MMBR.68.3.560-602.2004 (2004).

Buchrieser, C., Prentice, M. & Carniel, E. The 102-kilobase unstable region of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement. *J Bacteriol* **180**, 2321-2329 (1998).

Burrus, V. & Waldor, M. K. Control of SXT integration and excision. *J Bacteriol* **185**, 5045-5054 (2003).

Burrus, V. & Waldor, M. K. Shaping bacterial genomes with integrative and conjugative

elements. *Res Microbiol* **155**, 376-386, doi:10.1016/j.resmic.2004.01.012 (2004).

Burrus, V., Pavlovic, G., Decaris, B. & Guedon, G. The ICESt1 element of *Streptococcus thermophilus* belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**, 77-97 (2002).

Bushman, W., Thompson, J. F., Vargas, L. & Landy, A. Control of directionality in lambda site specific recombination. *Science* **230**, 906-911 (1985).

Cai, M. & Xun, L. Organization and regulation of pentachlorophenol-degrading genes in *Sphingobium chlorophenolicum* ATCC 39723. *J Bacteriol* **184**, 4672-4680 (2002).

Camper, D. N., Effects of Pesticide Degradation Products on Soil Microflora. Pesticide Transformation Products Chapter 15, pp 205–216 DOI: 10.1021/bk-1991-0459.ch015. ACS Symposium Series, Vol. 459 (1991).

Canchaya, C., Proux, C., Fournous, G., Bruttin, A. & Brussow, H. Prophage genomics. *Microbiol Molecular Biol rev* **67**, 238-276, (2003).

Carattoli, A. Plasmid-mediated antimicrobial resistance in *Salmonella enterica*. *Curr Issues Mol Biol* **5**, 113-122 (2003).

Cha, C. J. Catechol 1,2-Dioxygenase from *Rhodococcus rhodochrous* N75 Capable of metabolizing Alkyl-Substituted catechols. *J Microbiol Biotechnol* 16(5): 778-785 (2006).

Chakrabarty, A. M. Genetic basis of the biodegradation of salicylate in *Pseudomonas. J Bacteriol* **112**, 815-823 (1972).

Chakrabarty, A. M. Genetic fusion of incompatible plasmids in *Pseudomonas*. *Proc Natl Acad Sci U S A* **70**, 1641-1644 (1973).

Chakrabarty, A. M., Friello, D. A. & Bopp, L. H. Transposition of plasmid DNA segments specifying hydrocarbon degradation and their expression in various microorganisms. *Proc Natl Acad Sci U S A* **75**, 3109-3112 (1978).

Chauhan, A., Chakraborti, A. K. & Jain, R. K. Plasmid-encoded degradation of *p*-nitrophenol and 4-nitrocatechol by *Arthrobacter protophormiae*. *Biochem Biophys Res Commun* **270**, 733-740, doi:10.1006/bbrc.2000.2500 (2000).

Cheng, T. C., DeFrank, J. J. & Rastogi, V. K. *Alteromonas* prolidase for organophosphorus G-agent decontamination. *Chem Biol Interact* **119-120**, 455-462 (1999).

Cheng, T. C., Harvey, S. P. & Stroup, A. N. Purification and Properties of a Highly Active Organophosphorus Acid Anhydrolase from *Alteromonas undina*. *Appl Environ Microbiol* **59**, 3138-3140 (1993).

Chiang, P. & Burrows, L. L. Biofilm formation by hyperpiliated mutants of *Pseudomonas aeruginosa*. *J Bacteriol* **185**, 2374-2378 (2003).

Chieko Nakai, S., Kihachiro Horiike, S., Seiki Kuramitsug, Hiroyuki Kagamiyamap, and Mitsuhiro Nozaki. Three Isozymes of Catechol l,Z-Dioxygenase (Pyrocatechase), αα, αβ and ββ from *Pseudomonas arvilla* C-l. *J Bio Chem* **265.** 660-665 (1990).

Cho, C. M., Mulchandani, A. & Chen, W. Altering the substrate specificity of organophosphorus hydrolase for enhanced hydrolysis of chlorpyrifos. *Appl Environ Microbiol* **70**, 4681-4685, doi:10.1128/AEM.70.8.4681-4685.2004 (2004).

Chow, J. Y., Wu, L. & Yew, W. S. Directed evolution of a quorum-quenching lactonase from *Mycobacterium avium subsp*. *paratuberculosis* K-10 in the amidohydrolase superfamily. *Biochemistry* **48**, 4344-4353, doi:10.1021/bi9004045 (2009).

Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. & Cascales, E. Biogenesis, architecture, and function of bacterial type IV secretion systems. *Ann Rev Microbiol* **59**, 451-485, doi:10.1146/annurev.micro.58.030603.123630 (2005).

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

Clarke, M., Maddera, L., Harris, R. L. & Silverman, P. M. F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci U S A* **105**, 17978-17981, doi:10.1073/pnas.0806786105 (2008).

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

Collis, C. M., Kim, M. J., Stokes, H. W. & Hall, R. M. Integron-encoded IntI integrases preferentially recognize the adjacent cognate *attI* site in recombination with a 59-be site. *Mol Microbiol* **46**, 1415-1427 (2002).

Cornelis, G., Ghosal, D. & Saedler, H. Tn*951*: a new transposon carrying a lactose operon. *Mol Gen Genet* **160**, 215-224 (1978).

Craig, N. L. Target site selection in transposition. *Annu Rev Biochem* **66**, 437-474, doi:10.1146/annurev.biochem.66.1.437 (1997).

Crowder, M. W., Spencer, J. & Vila, A. J. Metallo-beta-lactamases: novel weaponry for antibiotic resistance in bacteria. *Acc Chem Res* **39**, 721-728, doi:10.1021/ar0400241 (2006).

Currier, T. C. & Nester, E. W. Isolation of covalently closed circular DNA of high molecular weight from bacteria. *Anal Biochem* **76**, 431-441 (1976).

Dahlberg, C. & Chao, L. Amelioration of the cost of conjugative plasmid carriage in *Eschericha coli* K12. *Genetics* **165**, 1641-1649 (2003).

Daiyasu, H., Osaka, K., Ishino, Y. & Toh, H. Expansion of the zinc metallo-hydrolase family of the beta-lactamase fold. *FEBS Lett* **503**, 1-6 (2001).

Darnell, J., H. Lodish, and D. Baltimore. Molecular Cell Biology New York: Scientific American, (1986).

Davis, J. M. *et al.* Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. *Immunity* **17**, 693-702 (2002).

de la Campa, A. G., del Solar, G. H. & Espinosa, M. Initiation of replication of plasmid pLS1. The initiator protein RepB acts on two distant DNA regions. *J Mol Biol* **213**, 247-262, doi:10.1016/S0022-2836(05)80188-3 (1990).

de Vries, J. & Wackernagel, W. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc Natl Acad Sci USA* **99**, 2094-2099, doi:10.1073/pnas.042263399 (2002).

de Vries, J., Herzfeld, T. & Wackernagel, W. Transfer of plastid DNA from tobacco to the soil bacterium *Acinetobacter* sp. by natural transformation. *Mol Microbiol* **53**, 323-334, doi:10.1111/j.1365-2958.2004.04132.x (2004).

DeFrank, J. J. *et al.* Screening of halophilic bacteria and *Alteromonas* species for organophosphorus hydrolyzing enzyme activity. *Chem Biol Interact* **87**, 141-148 (1993).

del Solar, G. & Espinosa, M. The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: the antisense RNA II and the repressor protein, RepA. *Mol Microbiol* **6**, 83-94 (1992).

del Solar, G., Acebo, P. & Espinosa, M. Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II. *Mol Microbiol* **18**, 913-924 (1995).

del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. & Diaz-Orejas, R. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**, 434-464 (1998).

Dempsey, L. A. & Dubnau, D. A. Localization of the replication origin of plasmid pE194. *J Bacteriol* **171**, 2866-2869 (1989).

Dicke, K. E., Skrlin, S. M. & Murphy, S. E. Nicotine and 4-(methylnitrosamino)-1-(3-pyridyl) -butanone metabolism by cytochrome P450 2B6. *Drug Metab Dispos* **33**, 1760-1764, doi:10.1124/dmd.105.006718 (2005).

Dimopoulou, I. D., Kartali, S. I., Harding, R. M., Peto, T. E. & Crook, D. W. Diversity of antibiotic resistance integrative and conjugative elements among haemophili. *Journal of medical microbiol* **56**, 838-846, doi:10.1099/jmm.0.47125-0 (2007).

Dobrindt, U., Hochhut, B., Hentschel, U. & Hacker, J. Genomic islands in pathogenic and environmental microorganisms. *Nature Rev Microbiology* **2**, 414-424, doi:10.1038/nrmicro884 (2004).

Dobzhansky, T. Is genetic diversity compatible with human equality? *Social biology* **20**, 280-288 (1973).

Dodd, H. M., Horn, N. & Gasson, M. J. Analysis of the genetic determinant for production of the peptide antibiotic nisin. *J Gen Microbiol* **136**, 555-566 (1990).

Dong, Y. J., Bartlam, M., Sun, L., Zhou, Y.F., Zhang, Z.P., Zhang, C.G., Rao, Z., Zhang, X.E. Crystal structure of methyl parathion hydrolase from *Pseudomonas* sp. WBC-3. J Mol Biol 353,

655-663, doi:10.1016/j.jmb.2005.08.057 (2005).

Doublet, B., Boyd, D., Mulvey, M. R. & Cloeckaert, A. The Salmonella genomic island 1 is an integrative mobilizable element. *Mol Microbiol* **55**, 1911-1924, doi:10.1111/j.1365-2958.2005.04520.x (2005).

Doyle, M. *et al.* An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* **315**, 251-252, doi:10.1126/science.1137550 (2007).

Draghi, J. A. & Turner, P. E. DNA secretion and gene-level selection in bacteria. *Microbiol* **152**, 2683-2688, doi:10.1099/mic.0.29013-0 (2006).

Drenkard, E. & Ausubel, F. M. *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* **416**, 740-743, doi:10.1038/416740a (2002).

Dubnau, D. DNA uptake in bacteria. *Ann Rev Microbiol* **53**, 217-244, doi:10.1146/annurev.micro.53.1.217 (1999).

Dumas, D. P., Caldwell, S. R., Wild, J. R. & Raushel, F. M. Purification and properties of the phosphotriesterase from *Pseudomonas diminuta*. *J Biol Chem* **264**, 19659-19665 (1989).

Eddleston, M, Phillips, MR. Self-poisoning with pesticides. *BMJ* 328:42(2004).

Eric L. Taylor, A. Gordon Holley, Melanie Kirk. Pesticide Development: a Brief Look at the History, *Southern Regional Extension Forestry,* SREF-FM-010 (Texas A & M Publication 805-124) (2007).

Feng, G., Tsui, H. C. & Winkler, M. E. Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J Bacteriol* **178**, 2388-2396 (1996).

Figurski, D. H. & Helinski, D. R. 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 (1979).

Fluit, A. C. & Schmitz, F. J. Resistance integrons and super-integrons. *Clin Microbiol Infect* **10**, 272-288, doi:10.1111/j.1198-743X.2004.00858.x (2004).

Francuski, D. & Saenger, W. Crystal structure of the antitoxin-toxin protein complex RelB-RelE from *Methanococcus jannaschii*. *J Mol Biol* **393**, 898-908, doi:10.1016/j.jmb.2009.08.048 (2009).

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

Frantz, B., Aldrich, T. & Chakrabarty, A. M. Microbial degradation of synthetic recalcitrant compounds. *Biotechnol Adv* **5**, 85-99 (1987).

Frosch, M. & Meyer, T. F. Transformation-mediated exchange of virulence determinants by co-cultivation of pathogenic *Neisseriae*. *FEMS Microbiol lett* **79**, 345-349 (1992).

Funchain, P., Yeung, A., Stewart, J., Clendenin, W. M. & Miller, J. H. Amplification of mutator cells in a population as a result of horizontal transfer. *J Bacteriol* **183**, 3737-3741, doi:10.1128/JB.183.12.3737-3741.2001 (2001).

Gaillard, M. *et al.* The *clc* element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. *J Bacteriol* **188**, 1999-2013, doi:10.1128/JB.188.5.1999-2013.2006 (2006).

Gaillard, M., Pernet, N., Vogne, C., Hagenbuchle, O. & van der Meer, J. R. Host and invader

impact of transfer of the *clc* genomic island into *Pseudomonas aeruginosa* PAO1. *Proc Natl Acad Sci USA* **105**, 7058-7063, doi:10.1073/pnas.0801269105 (2008).

Galas, D. J., and M. Chandler. Bacterial insertion sequences, p. 109-163. In D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C. (1989).

Gal-Mor, O. & Finlay, B. B. Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cell Microbiol* **8**, 1707-1719, doi:10.1111/j.1462-5822.2006.00794.x (2006).

García-Trigueros, C., Dardanelli, M.S., Manyani, H., and Megías, M.(2007) Direct submission.

Garnier, F., Taourit, S., Glaser, P., Courvalin, P. & Galimand, M. Characterization of transposon Tn*1549*, conferring VanB-type resistance in *Enterococcus* sp. *Microbiol* **146 (Pt 6)**, 1481-1489 (2000).

Gasson, M.J., Godon, J.J., Pillidge, C.J., Eaton, T.J., Jury, K., and Shearman, C.A. Characterization and exploitation of conjugation in *Lactococcus lactis*. Int Dairy J 5: 757–762 (1995).

Gerdes, K. *et al.* Mechanism of postsegregational killing by the *hok* gene product of the *parB* system of plasmid R1 and its homology with the *relF* gene product of the *E. coli* relB operon. *Embo J* **5**, 2023-2029 (1986).

Gerdes, K., Rasmussen, P. B. & Molin, S. Unique type of plasmid maintenance function: post-segregational killing of plasmid-free cells. *Proc Natl Acad Sci U S A* **83**, 3116-3120 (1986).

Gill, R. E., Heffron, F. & Falkow, S. Identification of the protein encoded by the transposable element Tn*3* which is required for its transposition. *Nature* **282**, 797-801 (1979).

Giraldo, R. Common domains in the initiators of DNA replication in Bacteria, Archaea and Eukarya: combined structural, functional and phylogenetic perspectives. *FEMS Microbiol*

Rev **26**, 533-554 (2003).

Gomez-Valero, L., Rocha, E. P., Latorre, A. & Silva, F. J. Reconstructing the ancestor of *Mycobacterium leprae*: the dynamics of gene loss and genome reduction. *Genome research* **17**, 1178-1185, doi:10.1101/gr.6360207 (2007).

Gorla, P., Pandey, J. P., Parthasarathy, S., Merrick, M. & Siddavattam, D. Organophosphate hydrolase in *Brevundimonas diminuta* is targeted to the periplasmic face of the inner membrane by the twin arginine translocation pathway. *J Bacteriol* **191**, 6292-6299, doi:10.1128/JB.00824-09 (2009).

Gotfredsen, M. & Gerdes, K. The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. *Mol Microbiol* **29**, 1065-1076 (1998).

Greated, A., Lambertsen, L., Williams, P. A. & Thomas, C. M. Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida*. *Environ Microbiol* **4**, 856-871 (2002).

Grigoriev, A. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res* **26**, 2286-2290 (1998).

Grindley, N. D. & Reed, R. R. Transpositional recombination in prokaryotes. *Annu Rev Biochem* **54**, 863-896, doi:10.1146/annurev.bi.54.070185.004243 (1985).

Grindley, N. D. Transposition of Tn*3* and related transposons. *Cell* **32**, 3-5 (1983).

Grindley, N.D.F. The movement of Tn*3*-like elements: transposition and cointegrate resolution. In Mobile DNA II, Craig, N., Craigie, R., Gellert, M. and Lambowitz, A. (ed.), pp272-302. ASM Press, Washington, DC, USA (2002).

Grinsted, J., de la Cruz, F. & Schmitt, R. The Tn*21* subgroup of bacterial transposable elements. *Plasmid* **24**, 163-189 (1990).

Grohmann, E. Autonomous plasmid-like replication of Bacillus ICEBs1: a general feature of integrative conjugative elements? *Mol Microbiol* **75**, 261-263, doi:10.1111/j.1365-2958.2009.06978.x (2010).

Grohmann, E., Muth, G. & Espinosa, M. Conjugative plasmid transfer in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 277-301, (2003).

Groisman, E. A. & Ochman, H. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* **87**, 791-794 (1996).

Gronlund, H. & Gerdes, K. Toxin-antitoxin systems homologous with *relBE* of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes. *J Mol Biol* **285**, 1401-1415, doi:10.1006/jmbi.1998.2416 (1999).

Grozdanov, L. *et al.* Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *J Bacteriol* **186**, 5432-5441, doi:10.1128/JB.186.16.5432-5441.2004 (2004).

Gunderson, C. H., Lehmann, C. R., Sidell, F. R. & Jabbari, B. Nerve agents: a review. *Neurology* **42**, 946-950 (1992).

Hacker, J. & Carniel, E. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO reports* **2**, 376-381, doi:10.1093/embo-reports/kve097 (2001).

Hacker, J. & Kaper, J. B. Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**, 641-679, doi:10.1146/annurev.micro.54.1.641 (2000).

Hacker, J. *et al.* Deletions of chromosomal regions coding for fimbriae and hemolysins occur *in vitro* and *in vivo* in various extraintestinal *Escherichia coli* isolates. *Microb Pathog* **8**,

213-225 (1990).

Hacker, J., Blum-Oehler, G., Hochhut, B. & Dobrindt, U. The molecular basis of infectious diseases: pathogenicity islands and other mobile genetic elements. A review. *Acta Microbiol Immunol Hung* **50**, 321-330, doi:10.1556/AMicr.50.2003.4.1 (2003).

Hacker, J., Blum-Oehler, G., Hochhut, B. & Dobrindt, U. The molecular basis of infectious diseases: pathogenicity islands and other mobile genetic elements. A review. *Acta Microbiol Immunol Hung* **50**, 321-330, doi:10.1556/AMicr.50.2003.4.1 (2003).

Haenni, M. *et al.* Diversity and mobility of integrative and conjugative elements in bovine isolates of *Streptococcus agalactiae*, *S. dysgalactiae* subsp. *dysgalactiae*, and *S. uberis*. *App Environ Microbiol* **76**, 7957-7965, doi:10.1128/AEM.00805-10 (2010).

Halary, S., Leigh, J. W., Cheaib, B., Lopez, P. & Bapteste, E. Network analyses structure genetic diversity in independent genetic worlds. *Proc Natl Acad Sci USA* **107**, 127-132, doi:10.1073/pnas.0908978107 (2010).

Hall, R. M. & Collis, C. M. Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Mol Microbiol* **15**, 593-600 (1995).

Hamoen, L. W., Haijema, B., Bijlsma, J. J., Venema, G. & Lovett, C. M. The *Bacillus subtilis* competence transcription factor, ComK, overrides LexA-imposed transcriptional inhibition without physically displacing LexA. *J Biol chem* **276**, 42901-42907, doi:10.1074/jbc.M104407200 (2001).

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

Hanne, L.F., Kirk, L.L., Appel, S.M., Narayan, A.D. and Bains, K.K. Degradation and induction specificity in *Actinomycetes* that degrade *p*-nitrophenol. *Appl. Environ. Microbiol.* **59.** 3505-3508 (1993).

Harman, J.G., Allosteric regulation of the cAMP receptor protein. *Biochem Biophys Acta*. **1547:**1-17(2001).

Harwood, C. S. & Parales, R. E. The beta-ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* **50**, 553-590, doi:10.1146/annurev.micro.50.1.553 (1996).

Hayes, F. & Barilla, D. The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat Rev Microbiol* **4**, 133-143, doi:10.1038/nrmicro1342 (2006).

Heffron, F., McCarthy, B. J., Ohtsubo, H. & Ohtsubo, E. DNA sequence analysis of the transposon Tn*3*: three genes and three sites involved in transposition of Tn*3*. *Cell* **18**, 1153-1163 (1979).

Hernandez-Arriaga, A. M., Rubio-Lepe, T. S., Espinosa, M. & del Solar, G. Repressor CopG prevents access of RNA polymerase to promoter and actively dissociates open complexes. *Nucleic Acids Res* **37**, 4799-4811, doi:10.1093/nar/gkp503 (2009).

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

Hinerfeld, D. & Churchward, G. Xis protein of the conjugative transposon Tn*916* plays dual opposing roles in transposon excision. *Mol Microbiol* **41**, 1459-1467 (2001).

Hiraga, S. Chromosome and plasmid partitioning in *Escherichia coli*. Annu. Rev. Biochem. 61:283-306 (1992).

Holmes, R.K., Jobling, M.G. Genetics: Conjugation. in: Baron's Medical Microbiology (Baron S et al., eds.) (4th ed.). Univ of Texas Medical Branch. ISBN 0-9631172-1-1 (1996).

Horne, I., Qiu, X., Russell, R. J. & Oakeshott, J. G. The phosphotriesterase gene *opdA* in *Agrobacterium radiobacter* P230 is transposable. *FEMS Microbiol Lett* **222**, 1-8 (2003).

Horne, I., Sutherland, T. D., Harcourt, R. L., Russell, R. J. & Oakeshott, J. G. Identification of an *opd* (organophosphate degradation) gene in an *Agrobacterium* isolate. *Appl Environ Microbiol* **68**, 3371-3376 (2002).

Hsieh, P. Molecular mechanisms of DNA mismatch repair. *Mutation research* **486**, 71-87 (2001).

Hughes, E. J., Bayly, R. C. & Skurray, R. A. Characterization of a TOL-like plasmid from *Alcaligenes eutrophus* that controls expression of a chromosomally encoded *p*-cresol pathway. *J Bacteriol* **158**, 73-78 (1984).

Hui, F. M. & Morrison, D. A. Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. *J Bacteriol* **173**, 372-381 (1991).

Hulter, N. & Wackernagel, W. Double illegitimate recombination events integrate DNA segments through two different mechanisms during natural transformation of *Acinetobacter baylyi*. *Mol Microbiol* **67**, 984-995, doi:10.1111/j.1365-2958.2007.06096.x (2008).

Ilyina, T. V. & Koonin, E. V. Conserved sequence motifs in the initiator proteins for rolling circle

DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res* **20**, 3279-3285 (1992).

Ish-Horowicz, D. & Burke, J. F. Rapid and efficient cosmid cloning. *Nucleic Acids Res* **9**, 2989-2998 (1981).

Ishiguro, N., Sato, G., Sasakawa, C., Danbara, H. & Yoshikawa, M. Identification of citrate utilization transposon Tn*3411* from a naturally occurring citrate utilization plasmid. *J Bacteriol* **149**, 961-968 (1982).

Jacoby, G. A., Rogers, J. E., Jacob, A. E. & Hedges, R. W. Transposition of *Pseudomonas* toluene-degrading genes and expression in *Escherichia coli. Nature* **274**, 179-180 (1978).

Jaffe, A., Ogura, T. & Hiraga, S. Effects of the ccd function of the F plasmid on bacterial growth. *J Bacteriol* **163**, 841-849 (1985).

Jahnke, M., Lehmann, F., Schoebel, A., Auling, G. Transposition of the TOL catabolic genes (Tn*4651*) into the degradative plasmid pSAH of *Alcaligenes* sp. 0-1 ensures simultaneous mineralization of sulpho- and methyl-substituted aromatics. *J. Gen. Microbiol.* 139, 1959–1966 (1993).

Jain, R. K., Dreisbach, J. H. & Spain, J. C. Biodegradation of *p*-nitrophenol via 1,2,4-benzenetriol by an *Arthrobacter* sp. *Appl Environ Microbiol* **60**, 3030-3032 (1994).

Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer in microbial genome evolution. *Theoretical population biology* **61**, 489-495 (2002).

Jouanneau, Y., Meyer, C., Jakoncic, J., Stojanoff, V. & Gaillard, J. Characterization of a naphthalene dioxygenase endowed with an exceptionally broad substrate specificity toward polycyclic aromatic hydrocarbons. *Biochemistry* **45**, 12380-12391,

doi:10.1021/bi0611311 (2006).

Jovanovic, O. S., Ayres, E. K. & Figurski, D. H. Host-inhibitory functions encoded by promiscuous plasmids. Transient arrest of *Escherichia coli* segregants that fail to inherit plasmid RK2. *J Mol Biol* **237**, 52-64, doi:10.1006/jmbi.1994.1208 (1994).

Juhas, M. *et al.* Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33**, 376-393, doi:10.1111/j.1574-6976.2008.00136.x (2009).

Juhas, M. *et al.* Sequence and functional analyses of *Haemophilus* sp. genomic islands. *Genome Biol* **8**, R237, doi:10.1186/gb-2007-8-11-r237 (2007).

Kadiyala, V. & Spain, J. C. A two-component monooxygenase catalyzes both the hydroxylation of *p*-nitrophenol and the oxidative release of nitrite from 4-nitrocatechol in *Bacillus sphaericus* JS905. *Appl Environ Microbiol* **64**, 2479-2484 (1998).

Karlin, S., Mrazek, J. & Campbell, A. M. Compositional biases of bacterial genomes and evolutionary implications. *J Bacteriol* **179**, 3899-3913 (1997).

Karpouzas, D. G. & Singh, B. K. Microbial degradation of organophosphorus xenobiotics: metabolic pathways and molecular basis. *Adv Microb Physiol* **51**, 119-185 (2006).

Kawahara, K., Tanaka, A., Yoon, J. & Yokota, A. Reclassification of a parathione-degrading *Flavobacterium* sp. ATCC 27551 as *Sphingobium fuliginis*. *J Gen Appl Microbiol* **56**, 249-255 (2010).

Khajamohiddin, S. *et al.* A novel meta-cleavage product hydrolase from *Flavobacterium s*p. ATCC27551. *Biochem Biophys Res Commun* **351**, 675-681, doi:10.1016/j.bbrc.2006.10.080 (2006).

Khersonsky, O., Roodveldt, C. & Tawfik, D. S. Enzyme promiscuity: evolutionary and mechanistic

aspects. *Curr Opin Chem Biol* **10**, 498-508, doi:10.1016/j.cbpa.2006.08.011 (2006).

Kitts, P. A., Symington, L. S., Dyson, P. & Sherratt, D. J. Transposon-encoded site-specific recombination: nature of the Tn*3* DNA sequences which constitute the recombination site res. *Embo J* **2**, 1055-1060 (1983).

Klockgether, J., Reva, O., Larbig, K. & Tummler, B. Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J Bacteriol* **186**, 518-534 (2004).

Kobs, G. pGEM®-T Vector: Cloning of modified blunt-ended DNA fragments. *Promega Note*s 55, 28–29 (1995).

Kornberg A, and Baker TA DNA Replication, 2nd Edition WH Freeman and Company, New York, New York (1992).

Kornfeld, S., Benziman M. and Milner, Y. Alpha-ketoglutarate dehydrogenase complex of *Acetobacter xylinum*. Purification and regulatory properties. *J biol chem* **252**. 2940-2947 (1977).

Kues, U. & Stahl, U. Replication of plasmids in gram-negative bacteria. *Microbiol Rev* **53**, 491-516 (1989).

Kuroda, M. *et al.* Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* **357**, 1225-1240 (2001).

Lagomarsino, M. C., Bassetti, B., Castellani, G. & Remondini, D. Functional models for large-scale gene regulation networks: realism and fiction. *Molecular bioSystems* **5**, 335-344, doi:10.1039/b816841p (2009).

Lanka, E. & Wilkins, B. M. DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* **64**, 141-169, doi:10.1146/annurev.bi.64.070195.001041 (1995).

Larbig, K., Kiewitz, C. & Tummler, B. Pathogenicity islands and PAI-like structures in *Pseudomonas* species. *Curr Top Microbiol Immunol* **264**, 201-211 (2002).

Lawley, T. D., Klimke, W. A., Gubbins, M. J. & Frost, L. S. F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* **224**, 1-15 (2003).

LeClerc, J. E., Li, B., Payne, W. L. & Cebula, T. A. High mutation frequencies among *Escherichia coli* and *Salmonella pathogens*. *Science* **274**, 1208-1211 (1996).

Lee, C. A. & Grossman, A. D. Identification of the origin of transfer (*oriT*) and DNA relaxase required for conjugation of the integrative and conjugative element ICEBs1 of *Bacillus subtilis*. *J Bacteriol* **189**, 7254-7261, doi:10.1128/JB.00932-07 (2007).

Lee, C. A., Auchtung, J. M., Monson, R. E. & Grossman, A. D. Identification and characterization of *int* (integrase), *xis* (excisionase) and chromosomal attachment sites of the integrative and conjugative element ICEBs1 of *Bacillus subtilis*. *Mol Microbiol* **66**, 1356-1369, doi:10.1111/j.1365-2958.2007.06000.x (2007).

Leisinger, T. Microorganisms and xenobiotic compounds. *Experientia* **39**, 1183-1191 (1983).

Leonid Kurbatov, Dirk Albrecht, Heidrun Herrmann, and Lothar Petruschka. Analysis of the proteome of *Pseudomonas putida* KT2440 grown on different sources of carbon and energy. *Envirol Microbiol* **8**. 466-478 (2006).

Lesic, B. *et al.* Excision of the high-pathogenicity island of *Yersinia pseudotuberculosis* requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol Microbiol* **52**, 1337-1348, doi:10.1111/j.1365-2958.2004.04073.x (2004).

Lessard, P. A., O'Brien, X. M., Currie, D. H. & Sinskey, A. J. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. *BMC microbiol* **4**, 15,

doi:10.1186/1471-2180-4-15 (2004).

Levin, B. R. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. *Genetics* **99**, 1-23 (1981).

Levings, R. S., Lightfoot, D., Elbourne, L. D., Djordjevic, S. P. & Hall, R. M. New integron-associated gene cassette encoding a trimethoprim-resistant DfrB-type dihydrofolate reductase. *Antimicrob Agents Chemother* **50**, 2863-2865, doi:10.1128/AAC.00449-06 (2006).

Lewis, J. A. & Hatfull, G. F. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. *Nucleic Acids Res* **29**, 2205-2216 (2001).

Li, L. Y. *et al.* The mobilization regions of two integrated *Bacteroides* elements, NBU1 and NBU2, have only a single mobilization protein and may be on a cassette. *J Bacteriol* **177**, 3940-3945 (1995).

Li, L. Y., Shoemaker, N. B. & Salyers, A. A. Characterization of the mobilization region of a *Bacteroides* insertion element (NBU1) that is excised and transferred by *Bacteroides* conjugative transposons. *J Bacteriol* **175**, 6588-6598 (1993).

Liebert, C. A., Hall, R. M. & Summers, A. O. Transposon Tn*21*, flagship of the floating genome. *Microbiol Mol Biol Rev* **63**, 507-522 (1999).

Liu, D. *et al.* Mechanism of the quorum-quenching lactonase (AiiA) from *Bacillus thuringiensis*. Product-bound structures. *Biochemistry* **47**, 7706-7714, doi:10.1021/bi800368y (2008).

Liu, H., Zhang, J. J., Wang, S. J., Zhang, X. E. & Zhou, N. Y. Plasmid-borne catabolism of methyl parathion and *p*-nitrophenol in *Pseudomonas* sp. strain WBC-3. *Biochem Biophys Res*

Commun **334**, 1107-1114, doi:10.1016/j.bbrc.2005.07.006 (2005).

Liu, H., Zhang, J., Wang., S., Zhang, X., Zhou,N. Plasmid- borne catabolism of methyl parathion and *p*-nitrophenol in *Pseudomonas* sp. strain WBC-3. *Biochem Biophys Res Commun* **334.** 1107-1114 (2005).

Llosa, M., Gomis-Ruth, F. X., Coll, M. & de la Cruz Fd, F. Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45**, 1-8 (2002).

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

Mahillon, J. & Chandler, M. Insertion sequences. *Microbiol Mol Biol Rev* **62**, 725-774 (1998).

Majewski, J. & Cohan, F. M. The effect of mismatch repair and heteroduplex formation on sexual isolation in *Bacillus*. *Genetics* **148**, 13-18 (1998).

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F. M. & Dowson, C. G. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J Bacteriol* **182**, 1016-1023 (2000).

Mandrich, L. & Manco, G. Evolution in the amidohydrolase superfamily: substrate-assisted gain of function in the E183K mutant of a phosphotriesterase-like metal-carboxylesterase. *Biochemistry* **48**, 5602-5612, doi:10.1021/bi801932x (2009).

Marrero, J. & Waldor, M. K. The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. J B*acteriol* **189**, 3302-3305, doi:10.1128/JB.01902-06 (2007).

Masai, E. *et al.* Genetic and biochemical characterization of a 2-pyrone-4, 6-dicarboxylic acid hydrolase involved in the protocatechuate 4, 5-cleavage pathway of *Sphingomonas*

paucimobilis SYK-6. *J Bacteriol* **181**, 55-62 (1999).

Matic, I. *et al.* Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **277**, 1833-1834 (1997).

Matsubara, M. & Mizuno, T. The SixA phospho-histidine phosphatase modulates the ArcB phosphorelay signal transduction in *Escherichia coli*. *FEBS Lett* **470**, 118-124 (2000).

McCormick, M., Wishart, W., Ohtsubo, H., Heffron, F. & Ohtsubo, E. Plasmid cointegrates and their resolution mediated by transposon Tn*3* mutants. *Gene* **15**, 103-118 (1981).

Meier, P. & Wackernagel, W. Monitoring the spread of recombinant DNA from field plots with transgenic sugar beet plants by PCR and natural transformation of *Pseudomonas stutzeri. Transgenic research* **12**, 293-304 (2003).

Mercer, A. A., Morelli, G., Heuzenroeder, M., Kamke, M. & Achtman, M. Conservation of plasmids among *Escherichia coli* K1 isolates of diverse origins. *Infect Immun* **46**, 649-657 (1984).

Merone, L., Mandrich, L., Rossi, M. & Manco, G. A thermostable phosphotriesterase from the archaeon *Sulfolobus solfataricus*: cloning, overexpression and properties. *Extremophiles* **9**, 297-305, doi:10.1007/s00792-005-0445-4 (2005).

Miyauchi, K., Suh, S. K., Nagata, Y. & Takagi, M. Cloning and sequencing of a 2,5-dichlorohydroquinone reductive dehalogenase gene whose product is involved in degradation of gamma-hexachlorocyclohexane by *Sphingomonas paucimobilis*. *J Bacteriol* **180**, 1354-1359 (1998).

Mohd-Zain, Z. *et al.* Transferable antibiotic resistance elements in *Haemophilus influenzae* share a common evolutionary origin with a diverse family of syntenic genomic islands. *J*

Bacteriol **186**, 8114-8122, doi:10.1128/JB.186.23.8114-8122.2004 (2004).

Moon, K., Shoemaker, N. B., Gardner, J. F. & Salyers, A. A. Regulation of excision genes of the *Bacteroides* conjugative transposon CTnDOT. *J Bacteriol* **187**, 5732-5741, doi:10.1128/JB.187.16.5732-5741.2005 (2005).

Morand, P. C. *et al.* Type IV pilus retraction in pathogenic *Neisseria* is regulated by the PilC proteins. *Embo J* **23**, 2009-2017, doi:10.1038/sj.emboj.7600200 (2004).

Mulbry, W. W. & Karns, J. S. Parathion hydrolase specified by the *Flavobacterium* opd gene: relationship between the gene and protein. *J Bacteriol* **171**, 6740-6746 (1989).

Mulbry, W. W. & Karns, J. S. Purification and characterization of three parathion hydrolases from gram-negative bacterial strains. *Appl Environ Microbiol* **55**, 289-293 (1989).

Mulbry, W. W. *et al.* Identification of a plasmid-borne parathion hydrolase gene from *Flavobacterium* sp. by southern hybridization with *opd* from *Pseudomonas diminuta*. *Appl Environ Microbiol* **51**, 926-930 (1986).

Mulbry, W. W., Kearney, P. C., Nelson, J. O. & Karns, J. S. Physical comparison of parathion hydrolase plasmids from *Pseudomonas diminuta* and *Flavobacterium* sp. *Plasmid* **18**, 173-177 (1987).

Mulbry, W. W., Kearney, P. C., Nelson, J. O. & Karns, J. S. Physical comparison of parathion hydrolase plasmids from *Pseudomonas diminuta* and *Flavobacterium* sp. *Plasmid* **18**, 173-177 (1987).

Munnecke, D. M. Chemical, physical, and biological methods for the disposal and detoxification of pesticides. *Residue Rev* **70**, 1-26 (1979).

Munnecke, D. M. Hydrolysis of organophosphate insecticides by an immobilized-enzyme

system. *Biotechnol Bioeng* **21**, 2247-2261, doi:10.1002/bit.260211207 (1979).

Nagata, Y., Miyauchi, K. & Takagi, M. Complete analysis of genes and enzymes for gamma-hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26. *J Ind Microbiol Biotechnol* **23**, 380-390, doi:10.1038/sj/jim/2900736 (1999).

Nakatsu, C. H. & Wyndham, R. C. Cloning and expression of the transposable chlorobenzoate-3,4-dioxygenase genes of *Alcaligenes* sp. strain BR60. *Appl Environ Microbiol* **59**, 3625-3633 (1993).

Nakatsu, C. H., Providenti, M. & Wyndham, R. C. The *cis*-diol dehydrogenase *cbaC* gene of Tn*5271* is required for growth on 3-chlorobenzoate but not 3,4-dichlorobenzoate. *Gene* **196**, 209-218 (1997).

Nakatsu, C., Ng, J., Singh, R., Straus, N. & Wyndham, C. Chlorobenzoate catabolic transposon Tn*5271* is a composite class I element with flanking class II insertion sequences. *Proc Natl Acad Sci U S A* **88**, 8312-8316 (1991).

Nakazawa, T., Hayashi, E., Yokota, T., Ebina, Y. & Nakazawa, A. Isolation of TOL and RP4 recombinants by integrative suppression. *J Bacteriol* **134**, 270-277 (1978).

Nash, H. A. & Robertson, C. A. Purification and properties of the *Escherichia coli* protein factor required for lambda integrative recombination. *J Biol Chem* **256**, 9246-9253 (1981).

Nojiri, H., Shintani, M. & Omori, T. Divergence of mobile genetic elements involved in the distribution of xenobiotic-catabolic capacity. *App Microbiol and Biotechnol* **64**, 154-174, doi:10.1007/s00253-003-1509-y (2004).

Nordstrom, K. & Austin, S. J. Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet* **23**, 37-69, doi:10.1146/annurev.ge.23.120189.000345 (1989).

Norman, A., Hansen, L. H. & Sorensen, S. J. Conjugative plasmids: vessels of the communal gene pool. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**, 2275-2289, doi:10.1098/rstb.2009.0037 (2009).

O'Brien, T. F. Emergence, spread, and environmental effect of antimicrobial resistance: how use of an antimicrobial anywhere can increase resistance to any antimicrobial anywhere else. *Clin Infect Dis* **34 Suppl 3**, S78-84, doi:10.1086/340244 (2002).

Osborn, A. M. & Boltner, D. When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid* **48**, 202-212 (2002).

O'Toole, G. A. & Kolter, R. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol Microbiol* **30**, 295-304 (1998).

Ou, L.T and Sharma, A. Degradation of methyl parathion by a mixed bacterial culture and a *Bacillus* sp. isolated from different soils. *J Agric Food Chem.,* **37.** 1514-1518 (1989).

Pakala, S. B. *et al.* Biodegradation of methyl parathion and *p*-nitrophenol: evidence for the presence of a p-nitrophenol 2-hydroxylase in a Gram-negative *Serratia* sp. strain DS001. *Appl Microbiol Biotechnol* **73**, 1452-1462, doi:10.1007/s00253-006-0595-z (2007).

Palmen, R. & Hellingwerf, K. J. Uptake and processing of DNA by *Acinetobacter calcoaceticus*--a review. *Gene* **192**, 179-190 (1997).

Pandeeti, E. V., Chinnaboina, M. R. & Siddavattam, D. Benzoate-mediated changes on expression profile of soluble proteins in *Serratia* sp. DS001. *Lett Appl Microbiol* **48**, 566-571, doi:10.1111/j.1472-765X.2009.02564.x (2009).

Pansegrau, W. & Lanka, E. Common sequence motifs in DNA relaxases and nick regions from a

variety of DNA transfer systems. *Nucleic Acids Res* **19**, 3455 (1991).

Pansegrau, W. & Lanka, E. Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. *Nucleic Acids Res* **19**, 3455 (1991).

Park, H. S. *et al.* Design and evolution of new catalytic activity with an existing protein scaffold. *Science* **311**, 535-538, doi:10.1126/science.1118953 (2006).

Parkhill, J. *et al.* Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nature genetics* **35**, 32-40, doi:10.1038/ng1227 (2003).

Parkinson, J. S. & Kofoid, E. C. Communication modules in bacterial signaling proteins. *Annu Rev Genet* **26**, 71-112, doi:10.1146/annurev.ge.26.120192.000443 (1992).

Partridge, S. R. & Hall, R. M. Gene cassettes potentially encoding fosfomycin resistance determinants. *Antimicrob Agents Chemother* **49**, 860-861, doi:10.1128/AAC.49.2.860-861.2005 (2005).

Patel RN, Hou CT, Felix A, Lillard MO. Catechol 1,2- dioxygenase from *Acinetobacter calcoaceticus*: purification and properties. *J Bacteriol* **127**. 536–544 (1976).

Pernodet, J. L., Simonet, J. M. & Guerineau, M. Plasmids in different strains of *Streptomyces ambofaciens*: free and integrated form of plasmid pSAM2. *Mol Gen Genet* **198**, 35-41 (1984).

Pikkemaat, M. G. & Janssen, D. B. Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox. *Nucleic Acids Res* **30**, E35-35 (2002).

Pimentel, D and Leviton, L. Pesticides: Amounts applied and amounts reaching pests. *Bioscience,* **36.** 86-91 (1986).

Poelarends, G. J. *et al.* Haloalkane-utilizing *Rhodococcus* strains isolated from geographically distinct locations possess a highly conserved gene cluster encoding haloalkane catabolism. *J Bacteriol* **182**, 2725-2731 (2000).

Poelarends, G. J., Serrano, H., Johnson, W. H., Jr. & Whitman, C. P. Inactivation of malonate semialdehyde decarboxylase by 3-halopropiolates: evidence for hydratase activity. *Biochemistry* **44**, 9375-9381, doi:10.1021/bi050296r (2005).

Preston, G. M., Haubold, B. & Rainey, P. B. Bacterial genomics and adaptation to life on plants: implications for the evolution of pathogenicity and symbiosis. *Curr Opin Microbiol* **1**, 589-597 (1998).

Providenti, M. A., Mampel, J., MacSween, S., Cook, A. M. & Wyndham, R. C. *Comamonas testosteroni* BR6020 possesses a single genetic locus for extradiol cleavage of protocatechuate. *Microbiology* **147**, 2157-2167 (2001).

Prudhomme, M., Libante, V. & Claverys, J. P. Homologous recombination at the border: insertion-deletions and the trapping of foreign DNA in *Streptococcus pneumoniae*. *Proc Natl Acad Sci USA* **99**, 2100-2105, doi:10.1073/pnas.032262999 (2002).

Qiu, X., Gurkar, A. U. & Lory, S. Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* **103**, 19830-19835, doi:10.1073/pnas.0606810104 (2006).

Quandt, J. & Hynes, M. F. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. *Gene* **127**, 15-21 (1993).

Ramsay, J. P., Sullivan, J. T., Stuart, G. S., Lamont, I. L. & Ronson, C. W. Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel

recombination directionality factor RdfS, and a putative relaxase RlxS. *Mol Microbiol* **62**, 723-734, doi:10.1111/j.1365-2958.2006.05396.x (2006).

Rani, N. L. and Lalitha Kumari, D. Degradation of methyl parathion by *Pseudomonas putida*. *Can J Microbiol.,* **40**. 1000-1006 (1994).

Ravatn, R., Studer, S., Zehnder, A. J. & van der Meer, J. R. Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase clc element of *Pseudomonas* sp. Strain B13. *J Bacteriol* **180**, 5505-5514 (1998).

Ravatn, R., Zehnder, A. J. & van der Meer, J. R. Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from *Pseudomonas* sp. strain B13 to *Pseudomonas putida* F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms. *Appl Environ Microbiol* **64**, 2126-2132 (1998).

Reiner, A. M. & Hegeman, G. D. Metabolism of benzoic acid by bacteria. Accumulation of (-)-3,5-cyclohexadiene-1,2-diol-1-carboxylic acid by mutant strain of *Alcaligenes eutrophus*. *Biochemistry* **10**, 2530-2536 (1971).

Reiner, A. M. Metabolism of aromatic compounds in bacteria. Purification and properties of the catechol-forming enzyme, 3,5-cyclohexadiene-1,2-diol-1-carboxylic acid (NAD+) oxidoreductase (decarboxylating). *J Biol Chem* **247**, 4960-4965 (1972).

Rheinwald, J. G., Chakrabarty, A. M. & Gunsalus, I. C. A transmissible plasmid controlling camphor oxidation in *Pseudomonas putida*. *Proc Natl Acad Sci U S A* **70**, 885-889 (1973).

Riley, M. A. & Gordon, D. M. A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col-plasmid lineages. *J Gen Microbiol* **138**, 1345-1352

(1992).

Roberts, A. P. & Mullany, P. A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* **17**, 251-258, doi:10.1016/j.tim.2009.03.002 (2009).

Robson, J., McKenzie, J. L., Cursons, R., Cook, G. M. & Arcus, V. L. The *vapBC* operon from *Mycobacterium smegmatis* is an autoregulated toxin-antitoxin module that controls growth *via* inhibition of translation. *J Mol Biol* **390**, 353-367, doi:10.1016/j.jmb.2009.05.006 (2009).

Romero-Arroyo, C.E., Schell, M.A., Gaines, G.L., Neidle, E.L. *catM* encodes a LysR -type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. *J Bacteriol* **177**. 5891-5898 (1995).

Romine, M. F., Fredrickson, J. K. & Li, S. M. Induction of aromatic catabolic activity in *Sphingomonas aromaticivorans* strain F199. *J Ind Microbiol Biotechnol* **23**, 303-313, doi:10.1038/sj/jim/2900750 (1999).

Roodveldt, C. & Tawfik, D. S. Directed evolution of phosphotriesterase from *Pseudomonas diminuta* for heterologous expression in *Escherichia coli* results in stabilization of the metal-free state. *Protein Eng Des Sel* **18**, 51-58, doi:10.1093/protein/gzi005 (2005).

Ruiz-Maso, J. A., Lopez-Zumel, C., Menendez, M., Espinosa, M. & del Solar, G. Structural features of the initiator of replication protein RepB encoded by the promiscuous plasmid pMV158. *Biochim Biophys Acta* **1696**, 113-119 (2004).

Russi et al. "Molecular Machinery for DNA Translocation in Bacterial Conjugation". Plasmids: Current Research and Future Trends. Caister Academic Press. ISBN 978-1-904455-35-6 (2008).

Rutherford, K. *et al.* Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944-945 (2000).

Salyers, A. A., Shoemaker, N. B., Stevens, A. M. & Li, L. Y. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiol Rev* **59**, 579-590 (1995).

Saurin, W., Hofnung, M., Dassa, E. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol* **48**. 22-41 (1999).

Schmidt, H. & Hensel, M. Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* **17**, 14-56 (2004).

Sentchilo, V. S., Perebituk, A. N., Zehnder, A. J. & van der Meer, J. R. Molecular diversity of plasmids bearing genes that encode toluene and xylene metabolism in *Pseudomonas* strains isolated from different contaminated sites in Belarus. *Appl Environ Microbiol* **66**, 2842-2852 (2000).

Sentchilo, V., Zehnder, A. J. & van der Meer, J. R. Characterization of two alternative promoters for integrase expression in the clc genomic island of *Pseudomonas* sp. strain B13. *Mol Microbiol* **49**, 93-104 (2003).

Serdar, C. M., Gibson, D. T., Munnecke, D. M. & Lancaster, J. H. Plasmid Involvement in Parathion Hydrolysis by *Pseudomonas diminuta*. *Appl Environ Microbiol* **44**, 246-249 (1982).

Sethunathan, N. & Yoshida, T. A *Flavobacterium* sp. that degrades diazinon and parathion. *Can J Microbiol* **19**, 873-875 (1973).

Shapiro, J. A. Molecular model for the transposition and replication of bacteriophage Mu and

other transposable elements. *Proc Natl Acad Sci U S A* **76**, 1933-1937 (1979).

Sherratt, D. J. Tn*3* and related transposable elements: site-specific recombination and transposition. In Berg, D. E., Howe, M. (eds) Mobile DNA. American Society for Microbiology, Washinghton, DC pp. 163-184 (1989).

Shoemaker, N. B. & Salyers, A. A. Tetracycline-dependent appearance of plasmid like forms in *Bacteroides uniformis* 0061 mediated by conjugal *Bacteroides* tetracycline resistance elements. *J Bacteriol* **170**, 1651-1657 (1988).

Shoemaker, N. B., Vlamakis, H., Hayes, K. & Salyers, A. A. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl Environ Microbiol* **67**, 561-568, doi:10.1128/AEM.67.2.561-568.2001 (2001).

Shoemaker, N. B., Wang, G. R. & Salyers, A. A. NBU1, a mobilizable site-specific integrated element from *Bacteroides* spp., can integrate nonspecifically in *Escherichia coli*. *J Bacteriol* **178**, 3601-3607 (1996).

Siddavattam, D., Khajamohiddin, S., Manavathi, B., Pakala, S. B. & Merrick, M. Transposon-like organization of the plasmid-borne organophosphate degradation (*opd*) gene cluster found in *Flavobacterium* sp. *Appl Environ Microbiol* **69**, 2533-2539 (2003).

Silver, S. Bacterial resistances to toxic metal ions--a review. *Gene* **179**, 9-19 (1996).

Sinclair, M. I. & Holloway, B. W. Chromosomal insertion of TOL transposons in *Pseudomonas aeruginosa* PAO. *J Gen Microbiol* **137**, 1111-1120 (1991).

Sinclair, M. I., Maxwell, P. C., Lyon, B. R. & Holloway, B. W. Chromosomal location of TOL plasmid DNA in *Pseudomonas putida*. *J Bacteriol* **168**, 1302-1308 (1986).

Singh, B. K. Organophosphorus-degrading bacteria: ecology and industrial applications. *Nat Rev*

Microbiol **7**, 156-164, doi:10.1038/nrmicro2050 (2009).

Slayers, A. A., Shoemaker, N. B. and Li, L. Y. In the driver's seat: The *Bacteroides* conjugative transposons and the elements they mobilize. *J. Bacteriol*. 177:5727-5731 (1995).

Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. & de la Cruz, F. Mobility of plasmids. *Microbiol Molecul Biol Rev* **74**, 434-452, doi:10.1128/MMBR.00020-10 (2010).

Smith, I. M., D. G. McNamara, P. R. Scott, and K. M. Harris. Quarantine pests for Europe. 1992. CAB London.

Sogorb, M. A. & Vilanova, E. Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through hydrolysis. *Toxicol Lett* **128**, 215-228 (2002).

Somani, S. M. Chemical Warfare Agents USA: Academic Press Inc (1992).

Soskine, M. & Tawfik, D. S. Mutational effects and the evolution of new protein functions. *Nat Rev Genet* **11**, 572-582, doi:10.1038/nrg2808 (2010).

Spain, J. C. Environmental biotechnology research: an overview. *Aviat Space Environ Med* **65**, A131-137 (1994).

Spain, J. C., Wyss, O. & Gibson, D. T. Enzymatic oxidation of *p*-nitrophenol. *Biochem Biophys Res Commun* **88**, 634-641 (1979).

Spain, J.C. and Gibson, D.T. Pathway for biodegradation of *p*-nitrophenol in *Moraxella* sp. *Appl. Environ. Microbiol.* **57.** 812-819 (1991).

Stanier, R. Y., Wachter, D., Gasser, C. & Wilson, A. C. Comparative immunological studies of two *Pseudomonas* enzymes. *J Bacteriol* **102**, 351-362 (1970).

Stanisich, V. A. (1988). Identification and analysis of plasmids at the genetic level. *Methods Microbiol* **21**, 11–48.

Steinmoen, H., Knutsen, E. & Havarstein, L. S. Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a sub-fraction of the cell population. *Proc Natl Acad Sci USA* **99**, 7681-7686, doi:10.1073/pnas.112464599 (2002).

Stevens, A. M., Sanders, J. M., Shoemaker, N. B. & Salyers, A. A. Genes involved in production of plasmid like forms by a *Bacteroides* conjugal chromosomal element share amino acid homology with two-component regulatory systems. *J Bacteriol* **174**, 2935-2942 (1992).

Stevens, A. M., Shoemaker, N. B. & Salyers, A. A. The region of a *Bacteroides* conjugal chromosomal tetracycline resistance element which is responsible for production of plasmid like forms from unlinked chromosomal DNA might also be involved in transfer of the element. *J Bacteriol* **172**, 4271-4279 (1990).

Stokes, H. W. *et al.* Class 1 integrons potentially predating the association with Tn*402*-like transposition genes are present in a sediment microbial community. *J Bacteriol* **188**, 5722-5730, doi:10.1128/JB.01950-05 (2006).

Stokes, H. W. *et al.* Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. *Appl Environ Microbiol* **67**, 5240-5246, doi:10.1128/AEM.67.11.5240-5246.2001 (2001).

Sullivan, J. T. & Ronson, C. W. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc Natl Acad Sci U S A* **95**, 5145-5149 (1998).

Sullivan, J. T. *et al.* Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J Bacteriol* **184**, 3086-3095 (2002).

Summers, D. K., Beton, C. W. & Withers, H. L. Multicopy plasmid instability: the dimer catastrophe hypothesis. *Mol Microbiol* **8**, 1031-1038 (1993).

Summers, D. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol* **29**, 1137-1145 (1998).

Taddei, F. *et al.* Role of mutator alleles in adaptive evolution. *Nature* **387**, 700-702, doi:10.1038/42696 (1997).

Tan, H. M. Bacterial catabolic transposons. *Appl Microbiol Biotechnol* **51**, 1-12 (1999).

Tawfik, D. S. Biochemistry. Loop grafting and the origins of enzyme species. *Science* **311**, 475-476, doi:10.1126/science.1123883 (2006).

Taylor, P. Anticholinesterase agents. In Goodman and Gilman's The Pharmacological Basis of Therapeutics, 10th ed. (J. G. Hardman, L. E. Limbird, and A. G. Goodman, Eds.), pp. 175–192. McGraw-Hill, New York (2001).

te Poele, E. M. *et al.* Actinomycete integrative and conjugative pMEA-like elements of *Amycolatopsis* and *Saccharopolyspora* decoded. *Plasmid* **59**, 202-216, doi:10.1016/j.plasmid.2008.01.003 (2008).

Thomas CM. Promiscuous plasmids of Gram-negative bacteria. Academic Press: London; (1989).

Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev. Microbiol* **3**, 711-721, doi:10.1038/nrmicro1234 (2005).

Thompson, J and Abbot, D.C. Pesticide Residues, Lecture Series No. 3, The Royal Institute of Chemistry, 30, Russel Square, London, .W.C.1. (1966).

Tolmasky, M. E., Colloms, S., Blakely, G. & Sherratt, D. J. Stability by multimer resolution of pJHCMW1 is due to the Tn1331 resolvase and not to the *Escherichia coli* Xer system.

Microbiol **146 (Pt 3)**, 581-589 (2000).

Top EM, Moënne-Loccoz Y, Pembroke T, Thomas CM. Phenotypic traits conferred by plasmids. In: Thomas CM (ed) The horizontal gene pool: bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam, pp 249–285 (2000).

Top, E. M., Holben, W. E. & Forney, L. J. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. *Appl Environ Microbiol* **61**, 1691-1698 (1995).

Toussaint, A. *et al.* The biphenyl- and 4-chlorobiphenyl-catabolic transposon Tn*4371*, a member of a new family of genomic islands related to IncP and Ti plasmids. *Appl Environ Microbiol* **69**, 4837-4845 (2003).

Townsend, J. P., Nielsen, K. M., Fisher, D. S. & Hartl, D. L. Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. *Genetics* **164**, 13-21 (2003).

Tsuchimoto, S., Nishimura, Y. & Ohtsubo, E. The stable maintenance system pem of plasmid R100: degradation of PemI protein may allow PemK protein to inhibit cell growth. *J Bacteriol* **174**, 4205-4211 (1992).

Tsuda, M. & Iino, T. Genetic analysis of a transposon carrying toluene degrading genes on a TOL plasmid pWW0. *Mol Gen Genet* **210**, 270-276 (1987).

Tsuda, M. & Iino, T. Identification and characterization of Tn*4653*, a transposon covering the toluene transposon Tn*4651* on TOL plasmid pWW0. *Mol Gen Genet* **213**, 72-77 (1988).

Tsuda, M. & Iino, T. Naphthalene degrading genes on plasmid NAH7 are on a defective transposon. *Mol Gen Genet* **223**, 33-39 (1990).

Tsuda, M., Minegishi, K. & Iino, T. Toluene transposons Tn*4651* and Tn*4653* are class II transposons. *J Bacteriol* **171**, 1386-1393 (1989).

Turner, L. R., Lara, J. C., Nunn, D. N. & Lory, S. Mutations in the consensus ATP-binding sites of XcpR and PilB eliminate extracellular protein secretion and pilus biogenesis in *Pseudomonas aeruginosa*. *J Bacteriol* **175**, 4962-4969 (1993).

Ubeda, C. *et al.* SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. *Mol Microbiol* **67**, 493-503, doi:10.1111/j.1365-2958.2007.06027.x (2008).

Ubeda, C., Barry, P., Penades, J. R. & Novick, R. P. A pathogenicity island replicon in *Staphylococcus aureus* replicates as an unstable plasmid. *Proc Natl Acad Sci U S A* **104**, 14182-14188, doi:10.1073/pnas.0705994104 (2007).

Vallaeys, T., Courde, L., McGowan, C., Wright, A. D. & Fulthorpe, R. R. Phylogenetic analyses indicate independent recruitment of diverse gene cassettes during assemblage of the 2,4-D catabolic pathway. FEMS Microbiol Ecol 28, 373–382 (1999).

van der Meer, J. R., Eggen, R. I., Zehnder, A. J. & de Vos, W. M. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. *J Bacteriol* **173**, 2425-2434 (1991).

van der Meer, J. R., Ravatn, R. & Sentchilo, V. The *clc* element of *Pseudomonas* sp. strain B13 and other mobile degradative elements employing phage-like integrases. *Arch Microbiol* **175**, 79-85 (2001).

van der Meer, J. R., Zehnder, A. J. & de Vos, W. M. Identification of a novel composite

transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. *J Bacteriol* **173**, 7077-7083 (1991).

Van Ham, R. C. *et al.* Reductive genome evolution in *Buchnera aphidicola*. *Proc Natl Acad Sci USA* **100**, 581-586, doi:10.1073/pnas.0235981100 (2003).

Van Melderen, L., Bernard, P. & Couturier, M. Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol Microbiol* **11**, 1151-1157 (1994).

Vernikos, G. S. & Parkhill, J. Resolving the structural features of genomic islands: a machine learning approach. *Genome Res* **18**, 331-342, doi:10.1101/gr.7004508 (2008).

Vijayakumar, M. N. & Ayalew, S. Nucleotide sequence analysis of the termini and chromosomal locus involved in site-specific integration of the streptococcal conjugative transposon Tn*5252*. *J Bacteriol* **175**, 2713-2719 (1993).

Vulic, M., Dionisio, F., Taddei, F. & Radman, M. Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc Natl Acad Sci U S A* **94**, 9763-9767 (1997).

Wall, D. & Kaiser, D. Type IV pili and cell motility. *Mol Microbiol* **32**, 1-10 (1999).

Wang C, You S, Wang, S., Purification and characterization of a novel catechol 1,2-dioxygenase from *Pseudomonas aeruginosa* with benzoic acid as a carbon source. *Proc Biochem* **41.** 1594–1601 (2006).

Wang, H. *et al.* Characterization of the ends and target sites of the novel conjugative transposon Tn*5397* from *Clostridium difficile*: excision and circularization is mediated by the large resolvase, TndX. *J Bacteriol* **182**, 3775-3783 (2000).

Wang, X. *et al.* Phytodegradation of organophosphorus compounds by transgenic plants expressing a bacterial organophosphorus hydrolase. *Biochem Biophys Res Commun* **365**, 453-458, doi:10.1016/j.bbrc.2007.10.193 (2008).

Wei, M. *et al.* A transposable class I composite transposon carrying *mph* (methyl parathion hydrolase) from *Pseudomonas* sp. strain WBC-3. *FEMS Microbiol Lett* **292**, 85-91, doi:10.1111/j.1574-6968.2008.01468.x (2009).

Wernegreen, J. J., Harding, E. E. & Riley, M. A. *Rhizobium* gone native: unexpected plasmid stability of indigenous *Rhizobium leguminosarum*. *Proc Natl Acad Sci U S A* **94**, 5483-5488 (1997).

Whitchurch, C. B., Hobbs, M., Livingston, S. P., Krishnapillai, V. & Mattick, J. S. Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. *Gene* **101**, 33-44 (1991).

Williams, D. R. & Thomas, C. M. Active partitioning of bacterial plasmids. *J Gen Microbiol* **138**, 1-16 (1992).

Williams, K. P. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* **30**, 866-875 (2002).

Williams, P. A. & Murray, K. Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (arvilla) mt-2: evidence for the existence of a TOL plasmid. *J Bacteriol* **120**, 416-423 (1974).

Wolfgang, M., Park, H. S., Hayes, S. F., van Putten, J. P. & Koomey, M. Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function mutations in pilT, a twitching motility gene in *Neisseria gonorrhoeae*. *Proc Natl Acad Sci USA* **95**, 14973-14978

(1998).

Wommack, K. E. & Colwell, R. R. Virioplankton: viruses in aquatic ecosystems. *Microbiol mol boil rev* **64**, 69-114 (2000).

Wozniak, R. A. & Waldor, M. K. A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genet* **5**, e1000439, doi:10.1371/journal.pgen.1000439 (2009).

Wozniak, R. A. & Waldor, M. K. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* **8**, 552-563, doi:10.1038/nrmicro2382 (2010).

Wu, X. P., Liu, W. D., Cao, H., Li, S. P. & Cui, Z. L. Determination of the catalytic structures of methyl parathion hydrolase. *Sheng Wu Gong Cheng Xue Bao* **21**, 998-1002 (2005).

Wyndham, R. C. *et al.* Distribution of the catabolic transposon Tn*5271* in a groundwater bioremediation system. *Appl Environ Microbiol* **60**, 86-93 (1994).

Wyndham, R. C., Cashore, A. E., Nakatsu, C. H. & Peel, M. C. Catabolic transposons. *Biodegradation* **5**, 323-342 (1994).

Yamanaka, K., Fang, L. & Inouye, M. The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. *Molecular microbiology* **27**, 247-255 (1998).

Yamane, K. *et al.* Signal transduction, competence development and protein secretion in *Bacillus subtilis*. *Tanpakushitsu kakusan koso. Protein nucleic acid enzyme* **44**, 1467-1474 (1999).

Yang, H. L., Zubay, G. & Levy, S. B. Synthesis of an R plasmid protein associated with tetracycline resistance is negatively regulated. *Proc Natl Acad Sci USA* **73**, 1509-1512 (1976).

Yarmolinsky, M. B. Programmed cell death in bacterial populations. *Science* **267**, 836-837 (1995).

Yen, K. M. & Serdar, C. M. Genetics of naphthalene catabolism in *Pseudomonads*. *Crit Rev Microbiol* **15**, 247-268, doi:10.3109/10408418809104459 (1988).

Yew, W. S., Akana, J., Wise, E. L., Rayment, I. & Gerlt, J. A. Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry* **44**, 1807-1815, doi:10.1021/bi047815v (2005).

Young, D. M. & Ornston, L. N. Functions of the mismatch repair gene mutS from *Acinetobacter* sp. strain ADP1. *J Bacteriol* **183**, 6822-6831, doi:10.1128/JB.183.23.6822-6831.2001 (2001).

Zawadzki, P., Riley, M. A. & Cohan, F. M. Homology among nearly all plasmids infecting three *Bacillus* species. *J Bacteriol* **178**, 191-198 (1996).

Zhang, Z., Hong, Q., Xu, J., Zhang, X. & Li, S. Isolation of fenitrothion-degrading strain *Burkholderia* sp. FDS-1 and cloning of *mpd* gene. *Biodegradation* **17**, 275-283, doi:10.1007/s10532-005-7130-2 (2006).

Zhongli, C., Shunpeng, L. & Guoping, F. Isolation of methyl parathion-degrading strain M6 and cloning of the methyl parathion hydrolase gene. *Appl Environ Microbiol* **67**, 4922-4925 (2001).

Zylstra, G. J., and E. Kim. Aromatic hydrocarbon degradation by *Sphingomonas yanoikuyae* B1. *J Ind Microbiol Biotechnol*. 19:408-414 (1997).