# Biodegradation of organophosphates: Studies on membrane targeting of organophosphate hydrolase (OPH) and regulation of *opd* gene expression in *Brevundimonas diminuta* MG


**THESIS SUBMITTED FOR THE DEGREE OF**

*Doctor of Philosophy*
*in*
*Animal Sciences*

**Supervisor:**
*Prof. Dayananda Siddavattam*

**By:**
*Purushotham Gorla*

**Department of Animal Sciences**
**School of Life Sciences**
**University of Hyderabad**
**Hyderabad - 500 046**
**India**

**February, 2010**

**Enrolment No. 04LAPH08**


**University of Hyderabad**

(Central University established in 1974 by act of parliament)

**HYDERABAD – 500 046, INDIA**

## CERTIFICATE

This is to certify that **Mr. Purushotham Gorla** 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 "**Biodegradation of organophosphates: Studies on membrane targeting of organophosphate hydrolase (OPH) and regulation of *opd* gene expression in *Brevundimonas diminuta* MG"** 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 act of parliament)**
**HYDERABAD – 500 046, INDIA**

I hereby declare that the work embodied in this thesis entitled **"Biodegradation of organophosphates: Studies on membrane targeting of organophosphate hydrolase (OPH) and regulation of *opd* gene expression in *Brevundimonas diminuta* MG"** 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)**

**Purushotham Gorla**
**(Research Scholar)**

*Dedicated to……*

*Srmt. Challagali Amaravani*

*Sri. Challagali Prabhakar Raju*

# *Acknowledgments*

*I take this opportunity to express my sincere gratitude to my teacher and research supervisor,* ***Prof. Dayananda Siddavattam*** *for his unwavering support, guidance, critical judgment, and sustained interest throughout the course of investigation and preparation of the manuscripts that helped to reach my research goal.*

*I Thank the Dean* ***Prof. M. Ramanatham*** *and former Dean* ***Prof. A. S. Raghavendra*** *for allowing me to use the school facilities.*

*I would like to acknowledge the Head, Department of Animal Sciences,* ***Prof. S. Dayananda*** *and former Head* ***Prof. Aparna Dutta Gupta*** *for allowing me to use the Department facilities.*

*I would like to thank my Doctoral Committee members* ***Prof. Appa Rao Podile*** *and* ***Prof. Aparna Dutta Gupta*** *for their valuable suggestions and I always treasure all the discussions especially with* ***Prof. Aparna Dutta Gupta*** *on academic as well as other aspects of life.*

*I thank* ***Prof. P. Reddanna, Prof. Manjula Sridharan, Prof. B. Senthilkumaran, Dr. Anitha Jagota, Dr. Yenugu Suresh, Dr. R. Mourya, Dr. K. Arun Kumar*** *and* ***Dr. P. Jagan*** *for their valuable suggestions whenever needed.*

*I thank* ***Dr. Gopinath Kodetham, Dr. Brahmanandam Manavathi*** *and* ***Dr. Krishnaveni Misra*** *for their valuable suggestions in green fluorescent protein expression studies.*

*I thank* ***Dr. Mike Merrick*** *and* ***Dr. Tracy Palmer,*** *Dept. of Molecular Microbiology, John Innes Centre, Norwich, UK for providing me E. coli MC4100 and its tat mutants.*

*I would like to thank* ***all the faculty members*** *of School of Life Sciences for their help whenever needed.*

*I consider myself to be very fortunate to have such a wonderful lab environment during my stay. I would like to extend my heartfelt thanks and gratitude to the entire team or rather friends: Aleem, Elish, Gopi, Paul, Prasanna, Toshi, Swetha, Sunil and Ravi.*

*I would like to extend thanks to my senior colleagues Dr. Bramanandam Manavathi, Dr. Pakala Suresh Babu, Dr. Syed Khajamohiddin, Dr. Vinay Kumar for their valuable suggestions during my stay and special thanks to Dr. Jay Prakash Pandey who not only helped in part of my research work but also a good friend.*

*I specially thank my well wisher **Dr. Pakala Suresh Babu** who gave me an immense moral support throughout my Doctoral studies and in personal life as well.*

*I thank **DST, UGC-SAP** and **UGC-RGNF** for giving me financial assistance throughout my Doctoral studies.*

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

*I thank Mr. Lallan, Mr. Ankinedu, Mr. Jagan, Mr. Gopi and Ms. Jyothi who helped me in different endeavors during my work here and I extend my thanks to Ms. Rama Devi, Dr. Madhura Rekha, Mrs. Leena Bhashyam for their technical help whenever needed.*

*I thank Mr. S. Murthy, Ms. Nalini (CIL) and Dr. Nandini Rangarajan (CCMB) in their technical help for confocal images.*

*I Thank Mr. Manikanta (lab Asst), Mr. Krishna (Lab Attendant) for their assistance in administrative work and I also thank Mr. Mallesh who is always there to provide MQ water whenever needed.*

*I thank my friends who gave me support in all the way especially in tough times, Balanagesh, Suri, Gnani, Aravind, Gopi, Ravi, Ramesh, Sridhar, Narendra, Sanjay, Karan Singh, Cnu, Raghuveer, Balu, Devi, Chaitanya, Prasad, Damu, Sunil, Anand, Veena, Mahi, Sravan, Pathi, Anil, Viswa, Onni, Sita, Ramu, Giri, Chalapai, Aproy and Santosh.*

*I wish to remember with love, my dearest friend Hari for invaluable support throughout Doctoral studies.*

*I greatly acknowledge **Mrs. Sujatha Siddavattam** and **family members** for a kind help especially at the end of my Doctoral studies.*

*I specially thank my parents, **Smt. G. Lakshmi Devi** and **Sri. G. Venkataramana** and family members **Samrajyam, Suresh, Chinnu, Gangadhar, Smt. Lakshmidevi, Sri.. Kamalamma, Sri.. Mangamma, Sri.. Sreenivasulu, Smt.. Lalitha, Sri. Reddamma** and **Late Sri.. Venkatappa** and **Sri.. Gangenna,** for their affection, Love, encouragement and unwavering support in all the tough times.*

*I greatly acknowledge my lovable brother **Mr. Purushotham (CDFD)** for constant support in all the way till the end of my Doctoral work.*

*I specially thank all **my friends in** School Life Sciences for helping me whenever required.*

***(Purushotham Gorla)***

# INTRODUCTION

# CHAPTER - I

## To elucidate the role of twin arginine (*tat*) motif in membrane targeting of OPH in *B. diminuta*

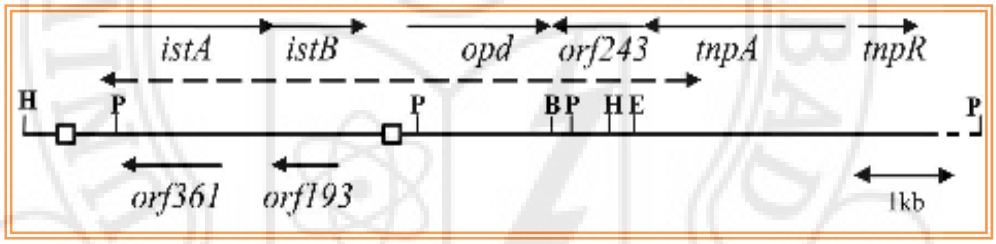
## 1.1. ORIGIN OF THE PROBLEM

The genesis of this work is identification of a structural motif that facilitates transport / targeting of prefolded proteins across / into the membrane in organophosphate hydrolase (OPH) isolated from a number of soil bacteria. The membrane associated OPH is a homodimer and possesses Zn ions as cofactor (Omburo et al., 1992; Benning et al., 1995). It hydrolyzes characteristic triester linkage found in structurally diverse group of neurotoxic organophosphates and nerve gas agents (Munnecke et al., 1976; Munnecke et al., 1982; Dumas et al., 1990). In view of its unusual catalytic properties and exceptional stability, a number of attempts have been made to produce recombinant OPH in large quantities. The recombinant OPH thus produced has successfully been used to develop biosensors and devices intending to use for detection and decontamination of op residues (Randhir et al., 2005).

The OPH producing bacteria have been isolated from diverse geographical regions (Sethunathan and Yoshida, 1973; Serdar et al., 1982; Rani and Lalitha kumari, 1994; Somara and Siddavattam, 1995; Horne et al., 2002). In all these isolates existence of highly conserved organophosphate degrading (*opd*) gene is shown to be responsible for coding OPH enzyme (Mulbry and Karns, 1989; Serdar and Gibson, 1985). The *opd* gene is constitutively expressed and in fact presence of op compounds in the culture medium failed to exert any influence on levels of OPH expression (Mulbry and Karns, 1986). In most of the soil bacteria the *opd* gene is found on large dissimilar plasmids (Mulbry et al., 1986; Harper et al., 1988; Somara and Siddavattam, 1995). As evidenced by hybridization

studies and restriction pattern a 7 kb DNA region containing *opd* gene was found to be conserved in plasmids pPDL2 and pCMS1 isolated from *Flavobacterium* sp. ATCC27551 and *Brevundimonas diminuta* respectively (Mulbry et al., 1986). However no homology was evident in these two plasmids outside this conserved 7 kb region. Such unusual observation has prompted us to undertake detailed sequence analysis of the conserved regions. In plasmid pPDL2, the conserved sequence showed feature of a complex transposon (Siddavattam et al., 2003). Mobile elements were found flanking *opd* gene and an adjacently located open reading frame (ORF) that codes for a protein of 306 amino acids (Fig. 1. 1).

**Fig. 1. 1**




**Map of the sequenced region *opd* gene cluster from pPDL2.** The eight ORFs identified are shown, together with the locations of the IR sequences that flank ISFlsp*1* ( ). (Siddavattam et al., 2003. *App. Environ. Microbiol.*)

An IS elements that show similarity to IS21 class of IS elements was identified upstream of the *opd* gene. Downstream of *orf306* a mobile element showing high similarity to transposon Tn*3* was identified. Existence of identical plasmid borne transposon-like *opd* gene cluster supports horizontal transfer of *opd* gene cluster among soil bacteria. Though considerable information is available on biochemistry and catalytic properties of OPH, information pertaining to the mechanism of membrane targeting is scarce. The present

study is a meticulously designed effort to throw light on these aspects. Further a moderate effort is also made to identify the *cis* elements that contribute for regulation of *opd* gene expression.

## 1.2. REVIEW OF EXISTING LITERATURE

In order to highlight the importance of the current work to the reviewer an attempt is made to provide a brief account on biodegradation of organophosphates and on membrane targeting /transport of bacterial proteins. The organophosphates are extensively used across the globe as insecticides to control number of insects that damage economically important crops (Brajesh, 2008). Though they are susceptible for enzymatic hydrolysis persistent and indiscriminate use resulted in accumulation of their residues in various components of environment. As op compounds exclusively inhibit acetyl cholinesterase (AChE) microbes and plants which do not have this target molecule are least affected. A good number of soil bacteria that use op compounds as source of carbon and energy have been isolated from diverse geographical regions (Singh and Walker, 2006). In fact, isolation of methyl parathion degrading *Flavobaterium* sp. from the rice fields of International Rice Research Institute is the first report on biodegradation of organophosphates (Sethunathan and Yoshida, 1973). After this report a number of parathion and methyl parathion degrading microorganisms were isolated (Munnecke et al., 1974; Rosenberg and Alexander, 1979; Nelson, 1982; Serdar et al., 1982; Lewis et al., 1985; Mulbry et al., 1986; Chaudhry et al., 1988; Ou et al., 1989; Misra et al., 1992; Rani and Lalithakumari 1994, Somara and Siddavattam 1995). In most of these isolates a 40 KD membrane associated homodimer designated as organophosphate hydrolase was shown to

be responsible for degradation of organophosphates (Serdar et al., 1989; Mulbry et al., 1986; Mulbry et al., 1986).

### 1.2.1. Organophosphorus hydrolase (OPH)

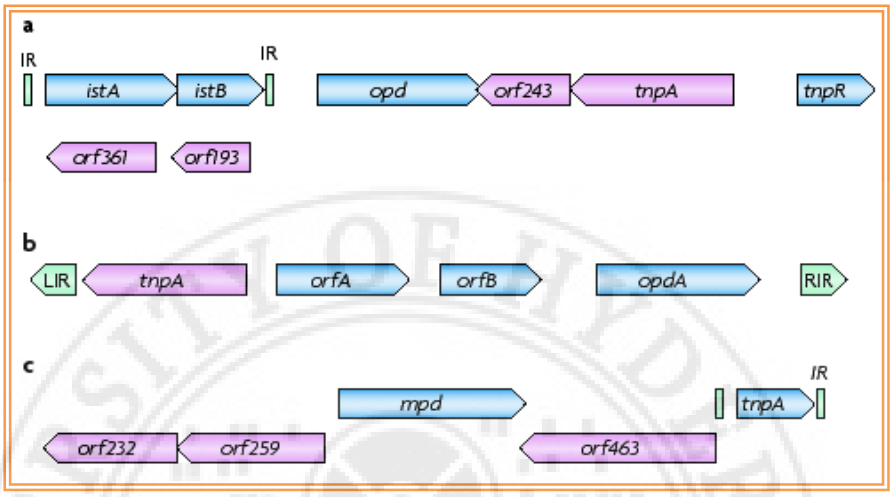
OPH hydrolyses triester linkage found in structurally diverse group of op compounds and nerve gas agents (Benning et al., 1994). Its potential in detoxification of organophosphate nerve agents and pesticides has been well documented (Munnecke et al., 1976; Munnecke et al., 1982; Dumas et al., 1990). The OPH has been purified independently and characterized from number of sources (Sethunathan and Yoshida, 1973, Serdar et al., 1982). It is identified as one of the members of the amidohydrolase super family (Holm and Sander, 1997). Members of this superfamily utilize one or two divalent metal ions to activate a hydrolytic water molecule for a nucleophilic attack at tetrahedral phosphorus or trigonal carbon centers. OPH is a homodimeric metalloprotein and the active site of the native OPH enzyme contains two zinc ions per monomer (Omburo et al., 1992; Benning et al., 1994). Though $Zn^{2+}$ ions serve as natural cofactors they can be replaced with $Co^{2+}$, $Cd^{2+}$, $Ni^{2+}$, or $Mn^{2+}$ ions without compromising on catalytic properties of the enzyme (Omburo et al., 1992).

### 1.2.2. Genetics of op compound degradation

*Flavobacterium* sp. ATCC27551, *Brevundimonas diminuta* were used as model organisms to gain insights on genetics and molecular biology of op compound degradation. In these two organisms almost identical organophosphate degrading (*opd*) genes were identified on otherwise dissimilar indigenous plasmids (Serdar et al, 1982; Mulbry and Karns, 1986). Among *opd* plasmids, pPDL2 (40kb, isolated from

*Flavobacteium* sp. ATCC27551) and pCMS1 (66Kb, isolated from *Brevundimonas diminuta*) were used for further characterization (Mulbry et al., 1987).

**Fig. 1. 2**




**Organizations of three known organophosphorus-degrading genes.** The shapes indicate different gene locations and the direction of transcription. **a)** Structure of the *opd* (organophosphorus degrading) gene from the *Flavobacterium* sp. genome, which includes a complete *istAB* operon, the *tnp*A and *tnp*R genes and *orf243*, a gene that encodes for metabolite utilization. Two *orfs* on a complementary strand encode for a protein of unknown function (Siddavattam et al., 2003). **b)** Genomic structure of the *opdA* gene from the *Agrobacterium radiobacter* genome, which includes *tnp*A and inverted repeats (IRs) (left inverted repeats (LIRs) and right inverted repeats (RIRs)), *opdA* and two *orfs* of unknown function (Horne et al., 2003). **c)** Genomic structure of the *mpd* gene cluster from the *Ochrobactrum* sp. genome, which includes IRs, *tnp*A and three *orfs* of unknown function (Zhang, 2006). (The Figure was adapted from Brajesh, 2008. *Nature Reviews Microbiology*).

In these two dissimilar plasmids *opd* gene was found as part of complex transposon (Siddavattam et al., 2003). Interestingly similar organization is seen for the *opd* gene reported from *Agrobacterium radiobacter* P230 isolated from Australian and Chinese agricultural soils (Horne et al., 2002). In a typical transposition assay Oakeshott and his associates have successfully demonstrated the event of *opd* element transposition in *E. coli* (Horne et al., 2002). A diagrammatic representation is shown below to draw a better understanding on organization of *opd* elements found in soil bacteria (Fig. 1. 2). If the structural organization of *opd* elements is clearly seen most of the *opd* associated ORFs are transposases. However in the *opd* element found in *Flavobacterium* sp. a novel esterase is

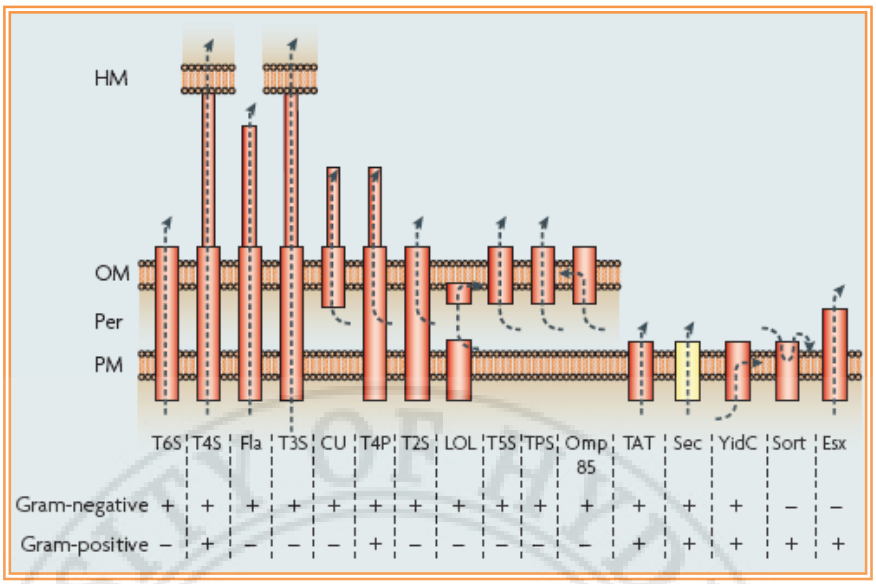
seen adjacent to *opd* gene and is later shown to hydrolyze meta fission products generated during degradation of aromatic compounds (Khajamohiddin et al., 2006).

### 1.2.3. Protein secretion in bacteria

A cell needs to keep the integrity of its essential processes and retain its reactants and enzymes. Bacteria often live in harsh environments such as the intestinal lumen, the apoplast of plant leaves and a variety of other locales where exposure to high temperature, saturating levels of salt and even raw sewage are not uncommon. To survive in these realms, all bacteria have evolved a resolute outer membrane or wall to protect themselves against these severe surroundings. This strategy not only prohibits permeation of hydrophobic substances and other potentially hazardous chemicals that might otherwise be lethal, it also creates new extracytoplasmic locations (e.g., the periplasm, the inner and outer membrane, the extracellular space) that function as both a buffer to extracellular stress and create a new environment for protein accumulation and protein chemistries (Bronstein et al., 2004).

The transport of proteins across the cell envelope is a basic function found in all groups of bacteria. Evolution has produced a remarkable array of mechanisms to export proteins. As shown in Fig. 1. 3 sixteen such systems, which handle protein secretion, sorting and membrane integration, are present in Bacteria (Holland, 2004; Economou et al., 2006). Among these pathways T6S, T4S, Fla, T3S, CU, T4P, T2S, LOL, T5S, TPS, Omp, TAT, Sec, YidC were found in gram negative bacteria, whereas T4S, T4P, TAT, Sec, YidC, Sort, Esx transport mechanisms were identified in gram positive bacteria (Effrosyni et al., 2007).

**Fig. 1. 3**

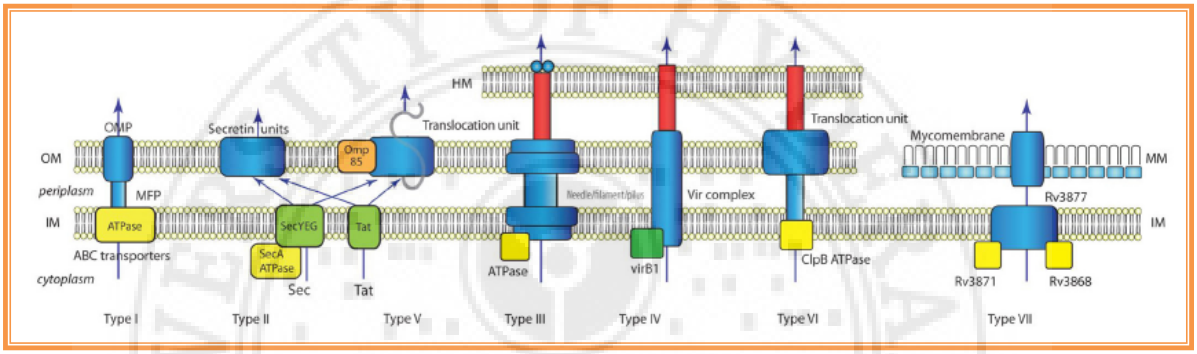



The protein-export and secretion-systems in Gram-negative and positive Bacteria. (Effrosyni et al., 2007. *Nature Reviews Microbiology*)

Analysis of a large number of bacterial genomes indicate that up to 17% of *Proteobacteria* genomes encode for proteins with signal sequences for the general secretory pathway (GSP) (Bendtsen et al., 2005), with many additional secretion systems and substrate proteins being present in most species. Secreted proteins perform few functions and most important of them are biogenesis of the cell envelope, acquisition of nutrients, motility, intercellular communication etc. (Bendtsen et al., 2005). The GSP and its components are found in all three kingdoms of life: in bacteria, archaea and eukaryotic organelles (chloroplasts but also in the endoplasmic reticulum) and provides a general mechanism for the transport of proteins across the cytoplasm or organelle membrane (Albert, 2002). While this transport process is sufficient for the secretion of proteins in Gram-positive species, Gram-negative species are posed with a specific problem, the transport across a second membrane system, the outer membrane (OM). In fact, the

difference found in OM secretion mechanisms are basis for classification of secretion pathways in Gram-negative bacteria. In Fig. 1. 4 a schematic representation is given showing connections among different secretion pathways in Gram-negative bacteria.

**Fig. 1. 4**




**Summary of known bacterial secretion systems**. In this simplified view only the basics of each secretion system are sketched. HM: Host membrane; OM: outer membrane; IM: inner membrane; MM: mycomembrane; OMP: outer membrane protein; MFP: membrane fusion protein. ATPases and chaperones are shown in yellow. (Tseng et al., 2009. *BMC Microbiology*)

### 1.2.3.1. Type I secretion systems

Type one secretion systems (T1SS) or ATP-binding cassette (ABC) transporters are heterotrimeric complexes consisting of an inner membrane (IM) ABC exporter, a membrane fusion protein (MFP) and a pore-forming, outer membrane protein (OMP) (Delepelaire, 2004; Holland et al., 2005). T1SS allow the secretion of a wide range of substrates (proteinaceous and nonproteinaceous) from the cytoplasm to the extracellular space in a single step, without a stable periplasmic intermediate. Most protein substrates

described so far possess a C-terminal signal sequence which is characterized by loosely conserved secondary structures (Stanley et al., 1991) and is not cleaved off during secretion. This implies that co-translational secretion is not possible (Delepelaire, 2004). The mechanism of type I secretion was studied in great detail on the basis of the α-hemolysin (HlyA) secretion found in some uropathogenic *Escherichia coli* (UPEC) (Thanabalu et al., 1998).

### 1.2.3.2. Type II secretion systems

The type II secretion system (T2SS) is also known as the Sec-dependent system as many proteins that pass through the T2SS must first reach the periplasm via the Sec pathway. Although Sec-dependent translocation is universal (Cao and Saier, 2003), the T2SS is found only in the Gram-negative proteobacteria phylum (Filloux, 2004; Cianciotto, 2005). It is found in species that extend from obligate symbionts (mutualistic, commensal and pathogenic) to free-living species, but is not universal among any particular group. It appears to be a specialized system that promotes functions specific to the interaction of a species with its biotic or abiotic environment. In Gram-negative bacteria, type II secretion is one the of five protein secretion systems that permit the export of proteins from within the bacterial cell to the extracellular milieu and/or into target host cells (Nicholas, 2005). The T2SS is required for virulence of the human pathogens *Vibrio cholerae*, *Legionella pneumonphila*, and enterotoxigenic *E. coli*, and of the plant pathogens *Ralstonia solanacearum, Pectobacterium atrosepticum* (*Erwinia carotovora* subsp. *atroseptica*) and *Xanthomonas campestris* pv.*campestris* (Filloux, 2004; Cianciotto, 2005).

### 1.2.3.3. Type III secretion system

Type III secretion system (T3SS) is most complex among all protein secretion systems. Its supramolecular structures span the inner membrane (IM), the periplasmic space, the outer membrane (OM), the extracellular space and a host cellular membrane (Daniela and Sheng, 2009). These complex assemblies are structurally and evolutionarily related to the flagella systems. T3SS have been isolated in species of several Gram-negative bacteria (*Salmonella, Yersinia, Shigella, Escherichia, Pseudomonas*) and were shown to consist at least 20 different subunits which enable these bacteria to translocate substrates (effectors) directly into the cytoplasm of the host cell to exert a broad range of virulence functions (Ghosh, 2004). Because of their shape and their ability to translocate proteins in a cell contact-dependent manner, they are also referred to as 'injectisomes' or 'molecular needles' (Cornelis, 2006). A characteristic of T3SS is the presence of cognate chaperones, small acidic proteins. These chaperones are considered to stabilize and prevent terminal folding of effector proteins. The energy of ATP hydrolysis by the ATPase is conducted by release of the chaperone from an effector chaperone complex and subsequent loading of the unfolded substrate into the T3SS apparatus (Akeda and Galan, 2005). Another function of the effector-specific chaperones could be the masking of domains needed for membrane targeting within the host cell (Letzelter et al., 2006).

### 1.2.3.4. Type IV secretion systems (T4SS)

T4SS are characterized by the ability to translocate proteins or complexes of protein and single-stranded DNA. Based on sequence similarities, T4SS are believed to have evolved from bacterial conjugation machineries (Cascales and Christie, 2003). The T-DNA

transfer system of *Agrobacterium tumefaciens* is the prototypical type A T4SS (Burns, 1999). This well-studied T4SS translocates protein–DNA complexes but is somehow distinct from the T4SS of pathogens of humans and animals that appear to translocate proteins only. Again by sequence comparison, T4SS are categorized into two subclasses: type IVA (T4ASS) and type IVB (T4BSS) (Christie and Vogel, 2000). For substrate recruitment and targeting to the inner membrane parts of the transenvelope protein complex, a homohexamer of VirD4 acts as the so-called 'coupling protein' (CP) in T4ASS. A stable interaction of the CP with homologs of VirB10, a part of the multi-subunit transenvelope protein complex, was demonstrated (Llosa et al., 2003). The components of the transenvelope complex are members of the mating-pair formation (Mpf) protein family. Different functions could be assigned to sets of VirB proteins which are responsible in formation of the channel traversing the periplasmic space, outer membrane pore (Ward et al., 2002) and pilus formation required in conjugation. Recent work demonstrated the roles of T4ASS in several important human bacterial pathogens (Henderson et al., 2000). Pertussis toxin of *Bordetella pertussis* is secreted in a contact independent manner, while CagA of *Helicobacter pylori* is a translocated T4SS effector protein that induces inflammatory responses and cytoskeletal alterations in the host cell (Henderson et al., 2000). T4ASS have also been identified in *Brucella* spp. and *Bartonella henselae*, and the translocated effectors have central functions in the intracellular lifestyle of these pathogens.

In contrast to the T4ASS, T4BSS are less well understood. One example of a T4BSS is the virulence associated dot/icm machinery of *Legionella pneumophila* (Segal et al., 1998; Vogel et al., 1998). The system was discovered by screening for mutants unable to survive within macrophages (Vogel and Isberg, 1999). Individual mutants were assigned

to dot (defect in organelle trafficking) or *icm* (intracellular multiplication) according to their respective phenotypes. To date only in *L. pneumophila* substrates of a T4BSS have been described (Segal et al., 2005).

### 1.2.3.5. Type V secretion systems

The type V secretion system is one of the most recently described pathways permitting the translocation of proteins into the extracellular milieu. Paradoxically, it possesses the simplest secretion apparatus and represents the largest family of protein-translocating outer membrane porins in Gram-negative bacteria (Yen et al., 2002). This system was first described for the IgA1 protease produced by *Neisseria gonorrhoeae* (Pohlner et al., 1987). Type V secretion system (T5SS) includes several mechanisms such as the two-partner system (TPS) and the oligomeric coiled-coil adhesin (Oca) system. A very large number of proteins are secreted via the T5SS, more even than the T2SS, over 500 in the T5aSS class alone (Jacob et al., 2004; Dautin and Bernstein, 2007; Bernstein, 2007). Most of the T5SS secreted proteins characterized to date contribute to the virulence of animal or human pathogens. Proteins secreted via the T5SS include adhesins such as AIDA-I and Ag43 of *E. coli*, Hia of *Haemophilus influenzae*, YadA of *Yersinia enteroliticola* and Prn of *Bordetella pertussis*; toxins such as VacA of *Helicobacter pylori*; proteases such as IgA proteases of *Neisseria gonorrheae* and *Neisseria meningitides*, SepA of *Shigella flexneri* and PrtS of *Serratia marcescens*; and S-layer proteins such as rOmpB of *Rickettsia* sp. and Hsr of *Helicobacter pylori* (Tsai-Tien et al., 2009).

### 1.2.3.6. Type VI secretion system (T6SS)

The type VI secretion machinery (T6SS) is a recently characterized secretion system that appears to constitute a phage-tail-spike-like injectisome that has the potential to introduce effector proteins directly into the cytoplasm of host cells (Bingle et al., 2008; Shrivastava and Mande, 2008; Cascales, 2008), analogous to the T3SS and T4SS machineries. Investigation of the virulence-associated secretion cluster in *V. cholerae* led to the identification of type VI secretion system. Type VI secretion system (T6SS) manages the export of substrates at least in the extracellular space without the requirement of hydrophobic N-terminal signal sequences. Using the model host *Dictyostelium discoideum*, a virulence function of the T6SS secreted substrates have been shown, suggesting a translocation of these proteins in the cytosol of the amoebae (Pukatzk et al., 2006). Recently, another T6SS was identified in *P. aeruginosa* where translocation of an ATPase, ClpV1, was identified, presumably energizing the secretion mechanism (Mougous et al., 2006). The T6SS is required for virulence in human and animal pathogens such as *Vibrio cholerae*, *Edwardsiella tarda*, *Pseudomonas aeruginosa*, *Francisella tularensis*, and *Burkholderia mallei*, and also in plant pathogens *such as Agrobacterium tumefaciens, Pectobacterium atrosepticum* and *Xanthomonas oryzae* (Wu et al., 2008).

### 1.2.3.7. Type VII secretion system

Despite of many Gram-positive bacteria have only a single membrane, in certain species of Gram-positive bacteria, most notably the *Mycobacteria*, have a cell wall that is heavily modified by lipids, called a mycomembrane (Abdallah et al., 2007). The genomes of these species have evolved with a novel and specialized secretion systems also called as

Type VII secretion systems (T7SS) for the transport of extracellular proteins across their hydrophobic, and highly impermeable, cell wall (Abdallah et al., 2007). The structure and operation of the T7SS and the various proteins involved in the inner membrane translocation channel is yet to be determined (Tsai-Tien et al., 2009).

Although most proteins are translocated via the Sec-pathway, proteins that are already folded in the cytoplasm, for example because they have to incorporate co-factors in cellular compartment, are translocated *via* the Tat-pathway also called as twin arginine translocation pathway (Berks et al., 1996). The proteins destined to the membrane or extracellular environment by using Tat secretion pathway depends on the characteristic features of the signal peptide and nature of substrate proteins (Driessen et al., 1998; Berks et al., 2000; Driessen et al., 2001; Mori and Ito, 2001; Berks et al., 2003). Though Tat pathway is presently best characterised for *E. coli*, but it was first described for thylakoids which is also designated as the ΔpH pathway in chloroplasts (Mould and Robinson, 1991; Chaddock et al., 1995). Four proteins, TatA, B, C, and E, are involved in Tat-dependent protein secretion in *E. coli*. The detailed mechanism of Sec and Tat pathways were described in the following sections.

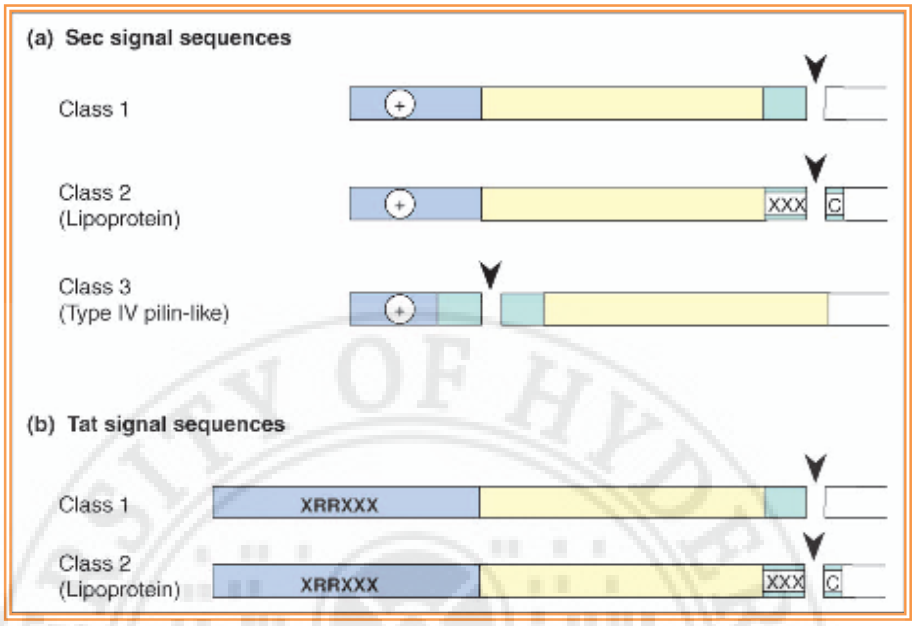
**1.2.3.8. Structural features - Signal peptide**

Signal peptides control the entry of virtually all proteins to the secretory pathway, both in eukaryotes and prokaryotes (Gierasch, 1989; von Heijne, 1990; Rapoport, 1992). The signal peptide of secretory proteins function as both the targeting and recognition signal and ranges in length from 18 to about 30 amino acid residues. It is composed of three domain: the positively charged amino terminus (N region); the nonpolar, hydrophobic

core region (H region); and the more polar cleavage region (C region). Although the amino acid sequences of these domains are not well conserved among the many signal sequences, but they demonstrate similar physical characteristics (Izard and Kendall, 1994). In fact, it is well reported that the shared features facilitate different signal peptides to interact with common elements in a 'general secretory pathway' in *Escherichia coli* (Pugsley, 1993). Sec substrates that are translocated across the hydrophobic membrane possess amino-terminal signal peptides that resemble a trans membrane segment (TM), but they also contain a short positively charged amino-terminus and often a signal peptidase recognition site (Von, 1990). Upon cleavage, the substrate is either released from the membrane or anchored to the lipid bilayer (Fig. 1. 5).

***N domain.*** The presence of a net positive charge in the N region, introduced by lysine or arginine residues, enhances the processing and translocation rates of a precursor protein but is not essential. Preproteins with signal sequences that carry a neutral or even negatively charged N region can be processed, although at reduced rates (Gennity et al., 1990). With increasing positive charge at the N region of the signal sequence, the SecA requirement for translocation is reduced while the interaction of the preprotein with SecA is enhanced in proteins involved in Sec-dependent translocation pathway (Akita et al., 1990). This suggests that the N region is involved in the targeting of the preprotein to the translocase.

**Fig. 1. 5**




Schematic representation of different classes of Sec and Tat signal sequences. Arrowheads indicate the signal peptide cleavage sites. Cleavage of predicted class 2 signal peptides of Tat substrates by SPII has not yet been experimentally confirmed. The LIPO-box is represented as XXXC in a white box. Signal sequences are colored as follows: blue, amino-domain; yellow, H-domain; green, carboxy-domain. The '+'symbols represent positive charges (Pohlschroder, 2005. *Current Opinion in Microbiology*)

The N region has been suggested to bind the negatively charged surface of the lipid bilayer of the membrane (Phoenix et al., 1993). A reduction in the number of positive charges in the N region results in inefficient interaction with the membrane; this phenomenon can be compensated for by an increased hydrophobicity of the H region (Phoenix et al., 1993). Strikingly, this is accompanied by a restoration of the translocation defect, suggesting that the interaction of the signal sequence domain with the membrane is an important step in targeting and/or translocation.

***N-domain – role in membrane topology***

The positive charges in the N region are thought to orient the signal sequence of secretory proteins or the stop-transfer signal of membrane proteins correctly within the

lipid bilayer. The transmembrane electrical potential (Dc, inside negative) prevents the translocation of positively charged residues and facilitates that of negatively charged residues (Anderson and Heijne, 1994). In this way, the Dc would contribute to the realization of the correct topology of integral membrane proteins, which obey the "positive-inside rule" (Von, 1990). On the other hand, in acidophilic bacteria and archaea, the Dc has a reversed polarity, i.e., inside positive versus outside. The topology of the inner membrane proteins of these bacteria also obeys the positive-inside rule (Vossenberg et al., 1998), whereas the signal sequences of the secretory proteins identified thus far are undistinguishable from those of neutro- or alkalophiles. This challenges the electrophoretic mechanism and suggests that Dc may affect protein translocation and membrane insertion by another mechanism that would involve the translocation apparatus in a more direct manner.

***H domain***

The H domain is the hydrophobic core of a signal sequence and varies in length from 7 to 15 amino acids. It is the most important part of the signal sequence; this is best illustrated by the observation that an increase in the total hydrophobicity of this domain can overcome mutations in the other regions of the signal sequence (Phoenix et al., 1993; MacFarlane and Muller, 1995). To some extent, the total hydrophobicity of the H region determines the efficiency of translocation and the translocation efficiency increases with the length and hydrophobicity of the H region (Chou and Kendall, 1990). This relation is sigmoidal, and a minimum hydrophobicity is required for translocation (Doud and Kendall, 1993). The residues in the H region are responsible for the α-helical conformation which

extends from the N region. Frequently, a so-called helix breaker, i.e., a glycyl or prolyl residue, is found in the middle of the H region. This may allow the signal sequence to form a hairpin-like structure that can insert into the lipid bilayer. According to the unlooping model, the signal sequence inserts into the membrane by extension through unlooping of this hairpin (Shinde et al., 1989; Vrije et al., 1990). Indeed, nuclear magnetic resonance spectroscopy studies have shown that in a membranous environment, signal sequences may adopt a two-domain conformation consisting of an amino-terminal α-helix and a more flexible carboxy-terminal domain (Rizo et al, 1993; Wang et al., 1993; Chupin et al., 1995). In addition, when two cysteines are introduced into the signal sequence, translocation under oxidized conditions is hampered (Nouwen et al., 1994), indicating that the formation of the loop prevents translocation. Unlooping would be facilitated by the Dc to orient the signal sequence within the electrical field. In this respect, some replacements of the α-helix breakers with α-helix-forming residues render preprotein translocation less dependent on the proton motive force (Nouwen et al., 1996), although the effects are rather subtle. The unlooping model does not explain how the signal sequence inserts into the membrane while bound to SecA. It is more likely that both SecA and the signal sequence insert simultaneously at the translocation site upon the binding of ATP.

***C domain*** The leader peptidase cleavage site (C domain) is the only part of the signal sequence that demands some primary sequence specificity. Two types of leader peptidases are known, type I, serving ordinary preproteins, and type II, cleaving the leaders of lipoproteins. For the signal sequences that rely on type I leader peptidases, the limitations are on the residues located at positions -1 and -3 relative to the start of the mature part (Von Heijne, 1984). This domain interacts with the leader peptidase which cleaves off the signal

sequence (Dalbey and Heijne, 1992). Usually, these residues have small neutral side chains, such as alanine, glycine, serine, and threonine (Von Heijne, 1984). Lipoproteins depend on type II leader peptidases, and the demands differ from those for type I peptidases only at the -3 and +1 positions (Sankaran and Wu, 1994). Precursors of lipoproteins contain larger hydrophobic amino acid residues at the -3 position, with a preference for leucine (Fekkes and Driessen, 1999). At the +1 position, a cysteine is always present and has to undergo modification prior to processing. After the signal sequence has been removed, it is degraded completely by a number of peptidases (Heijne, 1990).

**1.2.3.9. Classification of Signal peptides**

Based on the structural features the signal peptides have been classified into two major groups. These are designated as **Sec** signal peptides and **Tat** signal peptides (Pohlschroder, 2005).

**1.2.3.9.1. Sec signal peptides**

Though there are minor structural differences among Sec signal peptides, these are found mainly in proteins that are destined to target / translocate across the membrane in denatured form. Upon targeting / translocating they become active by acquiring native confirmation. However due to existence of minor differences they have been again classified into three sub groups.

- **Class 1 signal peptides** contain a class I signal peptidase (SPI) processing site following the hydrophobic stretch. These signal peptides are usually 18±35 amino

acids long and do not contain a strict consensus sequence. They share, however, a tripartite structure. The amino-terminal N domain, which is usually two to eight residues, contains one or more positively charged residues. The N-domain is important for interaction with the protein translocation machinery (Akita et al., 1990) and negatively charged lipid head groups on the cytoplasmic face of the lipid bilayer (de Vrije et al., 1990; Phoenix et al., 1993). The hydrophobic H-domain that follows the N-domain varies in length from 8 to 15 residues. This region has been proposed to form an α-helical conformation in the membrane (Briggs et al., 1986). The third domain (C domain) of the signal peptide contains the cleavage site for signal peptidase (SPase). The residues at positions -3 and -1 (relative to the start of the mature protein) are usually those with small neutral side chains, such as alanine, glycine, serine and threonine (von Heijne, 1984). Substrates processed by this universally conserved signal peptidase are either released from the membrane or anchored to the lipid bilayer by way of a carboxy-terminal membrane anchor (Pohlschroder et al., 2005).

- **Class 2 signal peptides** are found in bacterial lipoproteins. These signal peptides, which contain similar N and H-domains to Class I signal peptides, are characterized by a conserved lipobox in the C-domain with the consensus sequence L(A/S)(G/A)C (von Heijne, 1989). The invariable cysteine in this lipobox is lipid-modified by a diacylglyceryl transferase and becomes the first residue of the mature protein after cleavage by a lipoprotein-specific type II SPase. Due to the lipid-modified cysteine, the protein remains anchored to the cytoplasmic membrane.

These lipoproteins attach to the cytoplasm or outer membrane by way of a lipid-modified invariant amino-terminal cysteine (Pohlschroder et al., 2005).

- **Class 3 signal peptides** initially identified in bacterial type IV prepilin subunits, contain a prepilin peptidase processing site that precedes the hydrophobic stretch. These signal peptides also have a positively charged N-domain and a hydrophobic H-domain. In contrast to other signal peptides, prepilin signal peptides are cleaved just before the H-domain by a specific SPase that has its active site on the cytoplasmic face of the membrane (Lory, 1994). These are distinct from class 1 and class 2 signal peptides, this hydrophobic stretch remains part of the mature protein and is essential for initial membrane anchoring as well as for subunit–subunit interactions that are crucial for the biosynthesis of type IV pili and pilus like structures (Pohlschroder et al., 2005).

**1.2.3.9.2. Tat signal peptides**

As invariant twin arginines are found in the H region, they are designated as Tat (**t**win **a**rginine **t**ransport) signal peptides. These twin arginines are part of a consensus sequence of *tat* motif (S/T)**RR**xøø, where x is any residue (except polar residue or proline), and ø is a hydrophobic residue (Berks, 1996; Cristobal et al., 1999).

Substrates of the twin arginine transport (Tat) protein translocation system possess a signal sequence with n-, h-, and c-regions like those of Sec pathway proteins and shows several distinct Tat system-specific features. The most important feature is the presence of a twin arginine motif, which is part of a larger conserved sequence (S/T)-R-R-x-ø-ø (where

x is any amino acid and ⊠ is a hydrophobic amino acid) found at the n-region / h-region boundary (Cristobal et al., 1999; Berks et al. 2000). In Gram-negative Bacteria, this sequence is usually (S/T-R-R-x-F-L-K) and these signal peptides are longer, on average, by 14 amino acid residues than Sec signal peptides (Bardy et al., 2003).

Based on signal peptide cleavage by signal peptidases, Tat signal peptides were classified into two groups (Bolhuis, 2002; Bardy et al., 2003; Pohlschroder et al., 2004; Pohlschroder et al., 2005). **Class I Tat signal peptides** shows similar features of typical Tat signal peptide having n-termianl, h-hydrophobic and c-terminal region which possess highly conserved *tat* motif and a signal peptidase I cleavage site. Whereas **Class II Tat signal peptides** also possess similar features as that of the class I except that it contains LIPO box at the c-terminal region (Bolhuis, 2002; Albers and Driessen, 2002; Bardy et al., 2003; Albers et al., 2004) and reports suggesting that some of these LIPO-box-containing archaeal proteins have lipid modifications at the conserved terminal cysteines (Kokoeva et al., 2002).

### 1.2.4. Protein translocation - Evolution

The cellular membrane is a basic feature that separates the interior of the cell from its external environment. This membrane must combine two antagonistic features:

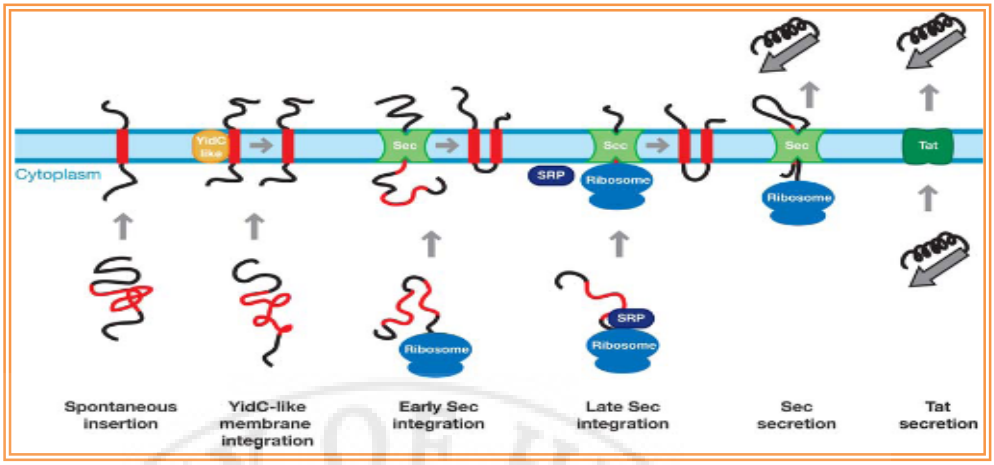
*(a)* It must guarantee the maintenance of both the composition and concentration of cytoplasmic molecules; and

*(b)* it must allow for the controlled exchange of matter and information between the cell and its environment.

This functional paradox is solved by combining two different types of molecules, amphiphilic lipids and proteins. The lipids provide a flexible and impermeable barrier, and the introduction of proteins in the form of receptors and channels enables the cell to transduce information and to regulate the transport of molecules. Thus, mechanisms allowing the insertion of proteins into the cellular membrane are a prerequisite for the origin of cells. This need for transport was no longer limited to smaller molecules, such as signaling molecules, catabolites, and a need to secrete large, enzymatically active proteins, enabling primitive cells to alter their external environments. As a result, protein export mechanisms evolved that allowed cells to digest organic matter and form extracytoplasmic protective structures, such as the cell wall and outer membrane (Pohlschroder et al., 2005). The earliest mechanism of protein translocation likely involved the spontaneous insertion of membrane proteins.

The next translocation system evolved is the YidC/Oxa1/Alb3 membrane integration system which is present in prokaryotic domains and is relatively simplistic and requires only the YidC protein to facilitate substrate translocation. But it alone is incapable of transferring large domains across the membrane. The other candidate is the universally conserved Sec system, which is minimally composed of SecY and SecE proteins. The Sec pathway facilitates protein integration into and translocation across membranes. Final and the recently evolved translocation system is Tat (**t**win **a**rginine **t**ranslocation) system and most likely evolved after the Sec pathway, since Tat signal sequences appear to be Sec-like and hence may have directly evolved from preexisting Sec signals (Pohlschroder et al., 2005). A theoretical timeline of protein translocation evolution was shown in Fig. 1. 6.

**Fig. 1. 6**




A theoretical timeline of protein translocation evolution. (Pohlschroder et al., 2005. *Annual Reviews of Microbiology*)

### 1.2.4.1. Sec-dependent protein translocation

In cells from all three domains of life, more than one third of the proteome is secreted across, or inserted into, biological membranes. Secretory proteins include hydrolytic enzymes, periplasmic lipoproteins, toxins and surface appendages such as pilli and flagella. Integral membrane proteins mediate selective transport, energy conversion, cell division, extracellular signal sensing, and membrane and cell-wall biogenesis. To successfully localize polypeptides extracytoplasmically, the cell must tackle five daunting tasks (Effrosyni et al., 2007).

1. Discriminate the cytoplasmic-resident proteins from those that are destined for export.
2. Deal with the inherent tendency of polypeptides to fold rapidly.
3. Target exported proteins to the membrane with specificity and fidelity.
4. Achieve transmembrane crossing of these elongated, heteropolymeric substrates, which are several times as long as the membrane is wide.

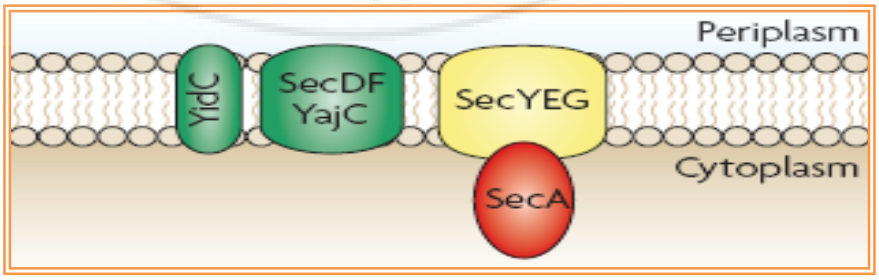
5. Manage a second sorting event that releases membrane proteins into the lipid bilayer and secretory proteins to the *trans* side of the membrane.

The Sec pathway is the only known universally conserved protein translocation pathway. Through the course of their evolutionary divergence, the three domains of life have reached disparate solutions to the fundamental problem of translocating proteins in an unfolded fashion to extracytoplasmic destinations via the Sec pathway (Pohlschroder et al., 2005).

**1.2.4.2.1. Structural features of the Sec translocase**

In all domains of life the bulk of the structure of the translocation channel comprises two universally conserved proteins, Sec61$\alpha$ and Sec61$\gamma$, in eukaryotes, with the corresponding proteins in the prokaryotic domains designated SecY and SecE, respectively (Hartmann et al., 1994). In addition to these Sec subunits, the pore complexe contain a third subunit that differs between domains is Sec61$\beta$, is conserved in eukaryotes and archaea, and a distinct protein, SecG, completes the heterotrimeric complex in bacteria (Van et al., 2004). Translocase consists of a peripheral ATPase, SecA, and at least five integral membrane proteins (i.e. SecY, SecE, SecG, SecD and SecF). As shown in Fig. 1. 7 this large complex most likely forms the translocation pore along which the preproteins pass the cytoplasmic membrane in an unfolded form (Eli et al., 2006).

**Fig. 1. 7**

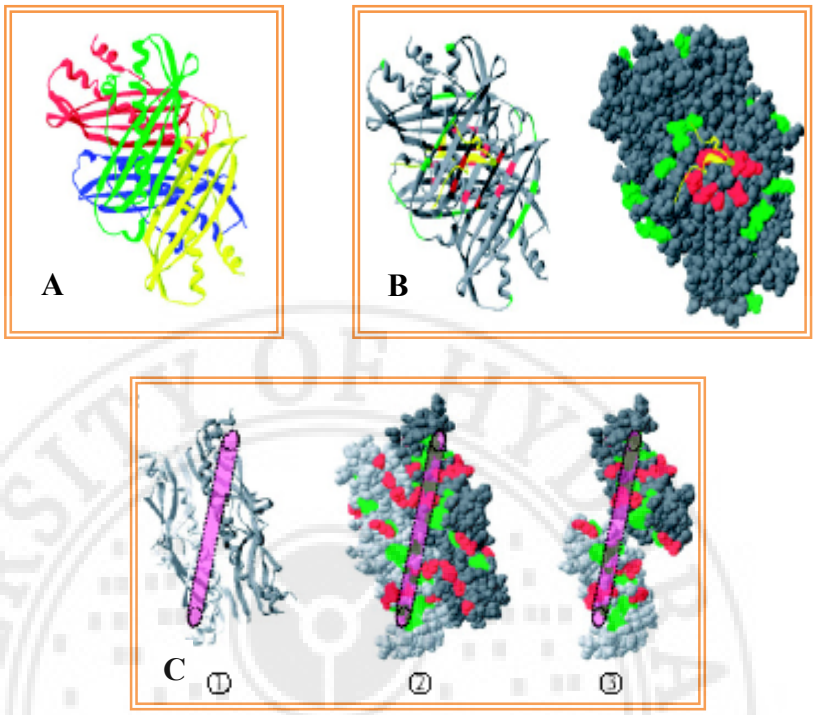



A schematic representation of the bacterial pre-protein translocase subunits (Effrosyni et al., 2007 *Nature Reviews Microbiology*).

### 1.2.4.2.2. The preprotein–SecB complex

SecB is a general cellular chaperone that appears to be important for the export of several secretory proteins, such as the outer-membrane porin lamb (Randall et al., 2002). It is not essential for viability and is not present in all bacteria, but has a dual role: it maintains secretory pre-proteins in a translocation-competent state and interacts specifically with membrane-bound SecA. SecB crystal structures (Xu, Z. et al., 2000; Zhou et al., 2003) revealed the structural features that could allow this chaperone to bind to extended polypeptides and SecA. SecB is a stable tetramer that probably binds to pre-proteins by recognizing exposed hydrophobic surfaces (Randall et al., 1998). A groove has been identified in the structures that could constitute the pre-protein binding site. Each SecB4 contains two such grooves, which, as these sites are mainly solvent exposed, could accommodate peptides without disturbing the stable tetramer (Fig. 1. 8c, pink cylinder). The grooves are lined by several aromatic residues at one end and hydrophobic residues at the opposing end (Fig. 1. 8c, part 2). The aromatic-residue-rich sub-site of SecB could accommodate the proposed SecB-binding motif in pre-proteins, which, it is suggested, is a nonapeptide that contains several aromatic and basic residues (Knoblauch et al., 1999). In an acidic region on the top surface of the SecB tetramer, several residues interact with the 25 C-terminal residues of the SecA C-terminal domain. (Fekkes et al., 1997; Woodbury et a., 2000) (Fig. 1. 8b). The structure of SecB and its interaction with SecA was shown in Fig. 1. 8.

**Fig. 1. 8**




**The structure of SecB. (A)** A ribbon Nature Reviews representation of the *Escherichia coli* SecB (*ecSecB*) tetramer, which is a dimer of dimmers (Dekker et al., 2003) (Protein Data Bank code: 1QYN). **(B) The SecB–SecA interaction**. A ribbon and space-filling model representation of the *ec*SecB tetramer (grey) (Dekker et al., 2003) in complex with the carboxy (C)-terminal peptide of *Haemophilus influenzae* SecA (yellow) (Xu et al., 2000; Zhou et al., 2003) is shown. **(C) The SecB–pre-protein complex.** A ribbon representation of the *ec*SecB4 tetramer is shown (Effrosyni et al., 2007. *Nature Reviews Microbiology*).

Several routes for how the polypeptide wraps its extended aminoacyl chain around SecB have been proposed, as would be expected for a chaperone that binds to numerous diverse substrates. SecB binds the preprotein at its mature domain (Randall and Hardy, 1995). The signal sequence provides no positive contribution to the binding energy or binding affinity of the interaction of the preprotein with SecB (Randall et al., 1998), but the signal sequence slows the folding of the mature domain (Park et al., 1988; Hardy and Randall, 1991). It has been suggested that this slowing process allows SecB to discriminate between the precursor proteins and other proteins in the cell, as formulated in the kinetic

partitioning model (Hardy and Randall, 1991). In this model, the cytosolic proteins would escape stable interaction by folding more rapidly than precursor proteins do. The final distribution of the precursor protein among the different pathways in the cell is then determined by partitioning that is dependent on the rate of folding or aggregation relative to the rate of binding to the chaperone (Randall and Hardy, 1986). Indeed, when the folding rate of the SecB-dependent precursors of maltose-binding protein (MBP) and galactose-binding protein (GBP) was increased, the amount of unbound ligand increased as well (Topping and Randall, 1997). However, several other observations are against the folding rate as the determining factor for SecB selectivity.

**1.2.4.2.3. The Sec YEG complex**

The translocase core consists of the integral membrane SecYEG heterotrimer. SecYEG retains limited but detectable sequence homology from bacteria to humans, and is remarkable in its ability to translocate proteins not only through, but also laterally into, the membrane (BOX 2), keeping the membrane barrier intact throughout these processes. SecYEG and SecA form the active holoenzyme (Brundage et al., 1990). SecA binds to the membrane with low affinity at acidic phospholipids but with high affinity (5–40 nM) at SecYEG (Hartl et al., 1990). Several cytoplasm-exposed loops of SecY are available for a possible interaction with SecA and/or the ribosome (Breyton et al., 2002; Van et al., 2004; Mitra et al., 2005; Bostina et al., 2005; Mori et al., 2006). SecYEG makes three connections with the ribosome — two between ribosomal-RNA hairpins and the cytoplasmic SecY loops and one between ribosomal proteins and the cytoplasmic region of SecG (Mitra et al., 2005). SecG, although not essential for the initial binding of SecA to

SecYEG, might interact at a later stage of the translocation reaction (Nishiyama et al., 1995; Nagamori et al., 2002). Pre-proteins cross the membrane, and a flexible molecular motor, the SecA ATPase, drives translocation at the expense of metabolic energy in the form of ATP and the protonmotive force (PMF). Other cytoplasmic and membrane subunits optimize the translocation reaction.

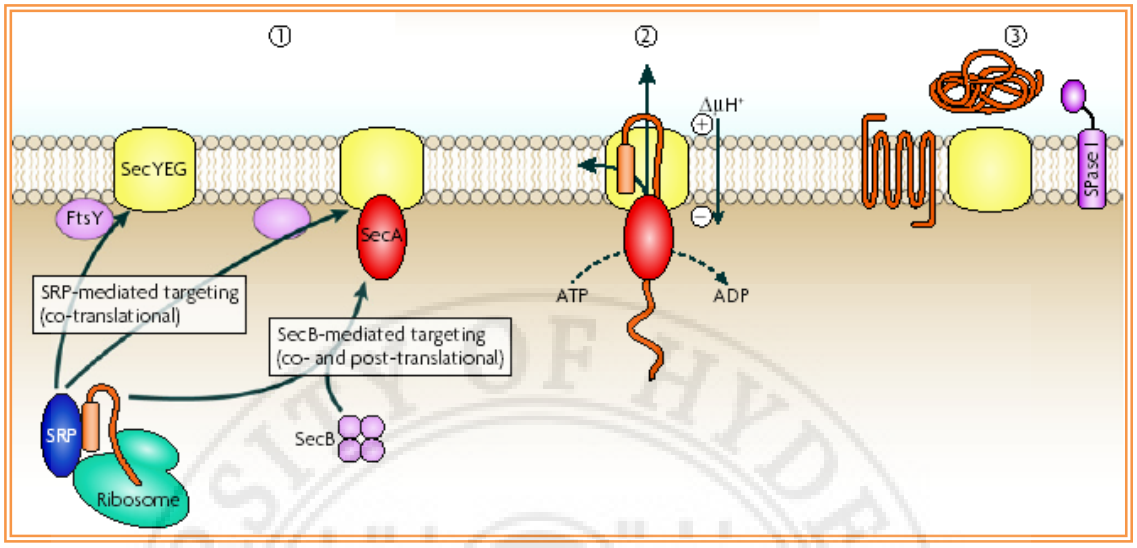
**1.2.4.2.4. Protein sorting and targeting**

Nascent pre-proteins are recognized directly by piloting factors, such as the ribonucleo protein signal-recognition particle (SrP) (Luirink and Sinning, 2004) or the SecB chaperone (Randall and Hardy. 2002; Ullers et al., 2004). Despite the fact that in bacteria the SrP is predominantly involved in the targeting of inner-membrane proteins (BOX 2), there are examples of long, strongly hydrophobic signal peptides of nascent secretory proteins, which are probably preferentially bound by the SrP (Schierle et al., 2003; Sijbrandi et al., 2003). Other signal peptides delay pre-protein folding and thus allow SecB to bind to the mature region of the pre-protein. This process can occur while the polypeptide chain is still being synthesized, but it is not mechanistically coupled to elongation (Randall, 1983). In both cases, the resulting SrP–pre-protein and SecB–pre-protein complexes are targeted to the translocase at the membrane. For SrP, this is achieved by docking to its membrane receptor FtsY3, and for SecB, by docking to the SecA subunit of the translocase (Hartl et al., 1990). Although SecA does not contribute to the SrP targeting route (Scotti et al., 1999), when long, hydrophilic segments are encountered, SecA is recruited to catalyse their export (Neumann et al., 2000).

**1.2.4.2.5. Translocation**

In the presence of preprotein, SecB binds SecA with higher affinity (Fekkes et al., 1997; Hartl et al., 1990). Only the signal sequence is needed to stimulate SecA for this high-affinity SecB recognition (Fekkes et al., 1998). Upon interaction of SecB with SecA, SecB releases the precursor protein, which is subsequently transferred to SecA (Fekkes et al., 1997; Fekkes et al., 1998). This reaction involves binding of the signal sequence of the preprotein to SecA. Upon binding SecB, SecA might lower the affinity of SecB for the preprotein by modulating the conformation of preprotein binding site of SecB. This site is located at the opposite face of the putative b strand that constitutes the SecA-binding site. According to this scenario, the SecA-bound SecB will be unable to interact with a new preprotein, and therefore the binding reaction precludes other proteins from entering the translocation pathway at a site that is already occupied. Only upon initiation of translocation, i.e., after the binding of ATP to SecA, SecB released from the membrane to bind a new preprotein in the cytosol (Fekkes et al., 1997). And finally pre-proteins are converted into mature proteins once after reaching the *trans* side of the membrane upon cleavage of signal peptide by signal peptidase (Paetzel et al., 2002; Mogensen and Otzen, 2005). A schematic representation of general Secretory pathway in *E. coli* was shown in Fig. 1. 9.

**Fig. 1. 9**




A schematic representation of general Secretory pathway (Effrosyni et al., 2007. *Nature Reviews Microbiology*).

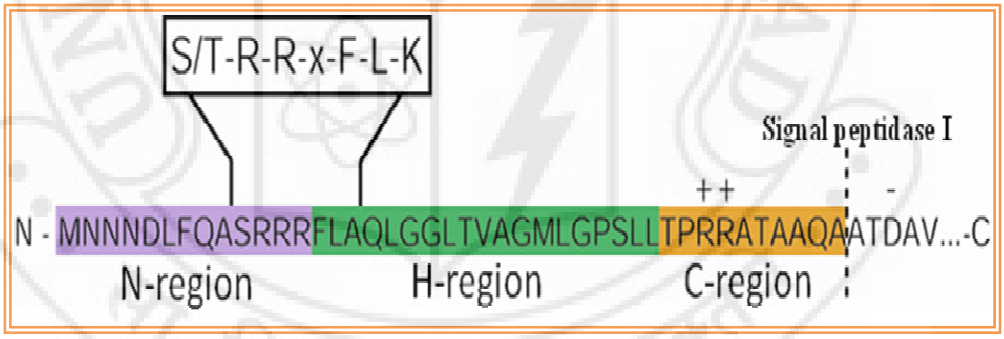
### 1.2.4.2. Twin Arginine Translocation (Tat) pathway

The discovery of the Tat pathway dates to the early 1990s, when it was noticed that a subset of polypeptides in chloroplasts could be translocated independently of ATP hydrolysis and instead depend exclusively on the proton gradient (Mould and Robinson, 1991; Cline et al., 1992). For this reason, it was initially designated the ΔpH pathway but more recently has been termed the cpTat pathway (chloroplast Tat/ΔpH pathway). In 1995 Creighton and his coworkers (Creighton et al., 1995) presented the first evidence that the cpTat pathway enables the translocation of prefolded proteins. Shortly thereafter, Berks (Berks, 1996) observed that a group of bacterial periplasmic proteins containing various cofactors share a unique type of signal peptide containing a consensus "double arginine" motif also found in substrates of the chloroplast pathway. The existence of a bacterial pathway analogous to the one in chloroplasts was thus established; it was initially termed

mtt (**m**embrane **t**argeting and **t**ranslocation) (Weiner et al., 1998) and then later Tat (**t**win-**a**rginine **t**ranslocation) (Sargent et al., 1998).

The bacterial cytoplasmic membrane and the chloroplast thylakoid membrane have a common evolutionary origin and contain equivalent protein transport pathways. The Sec pathway translocates proteins as unstructured chains (Rapoport, 2007), whereas the twin-arginine translocation (Tat) system is dedicated to transporting folded proteins (Berks et al., 2003; Berks et al., 2005; Cline and Theg, 2007). In both cases substrates must be translocated without compromising the membrane permeability barrier. This translocation is particularly challenging for the Tat apparatus because folded proteins are larger and more variable in size than the linear peptides transported by the Sec system.

**Fig. 1. 10**




Features of a typical Tat signal peptide, ssTorA from *E. coli*. (Philip et al., 2006. *Annual Reviews of Microbiology*)

The mechanism of Tat transport has not been established but appears unrelated to that of other membrane transporters (Leake et al., 2008). As shown in Fig. 1. 10 another unique feature of the Tat system is its substrates possess a conserved S**RR**xFLK "twin-

arginine" motif in their N-terminal signal sequences in which the twin arginines are highly invariant in nature (Berks et al., 1996).

**1.2.4.2.1. Organization and structural features of *tat* genes**

The first *tat* gene to be characterized was the maize hcf106 gene, following the characterization of a mutant line (Voelker and Barkan, 1995) defective in the targeting of a subset of thylakoid lumen proteins. It was well established that proteins were transported across the thylakoid membrane by two distinct pathways, one of which was related to bacterial Sec-type systems. The discovery of a chloroplastic Sec pathway was not unexpected given that these organelles evolved from endosymbiotic cyanobacteria, but the sequencing of the *hcf106* gene (Settles, 1997) provided the first structural information on the Sec-independent pathway used to transport other thylakoid lumen proteins. As expected, this gene was found to encode a novel component of the thylakoid protein transport apparatus. Most significantly, however, the *hcf106* gene was also found to be homologous to many open reading frames in bacterial genomes-the first concrete indication of a mainstream, Sec-independent bacterial protein export system. The first bacterial *tat* mutants were isolated in 1998 (Weiner et al., 1998; Sargent et al., 1998; Bogsch et al., 1998) and this field is just evolved when compared with most of the other protein transport systems.

The first *tatABC* operon was identified in *E. coli* since the initial isolation of *tat* mutants were discovered in this organism (Colin and Albert, 2004). As shown in Table. 1.1 the minimal integral membrane protein components TatA, TatB, and TatC were identified in *E. coli.*

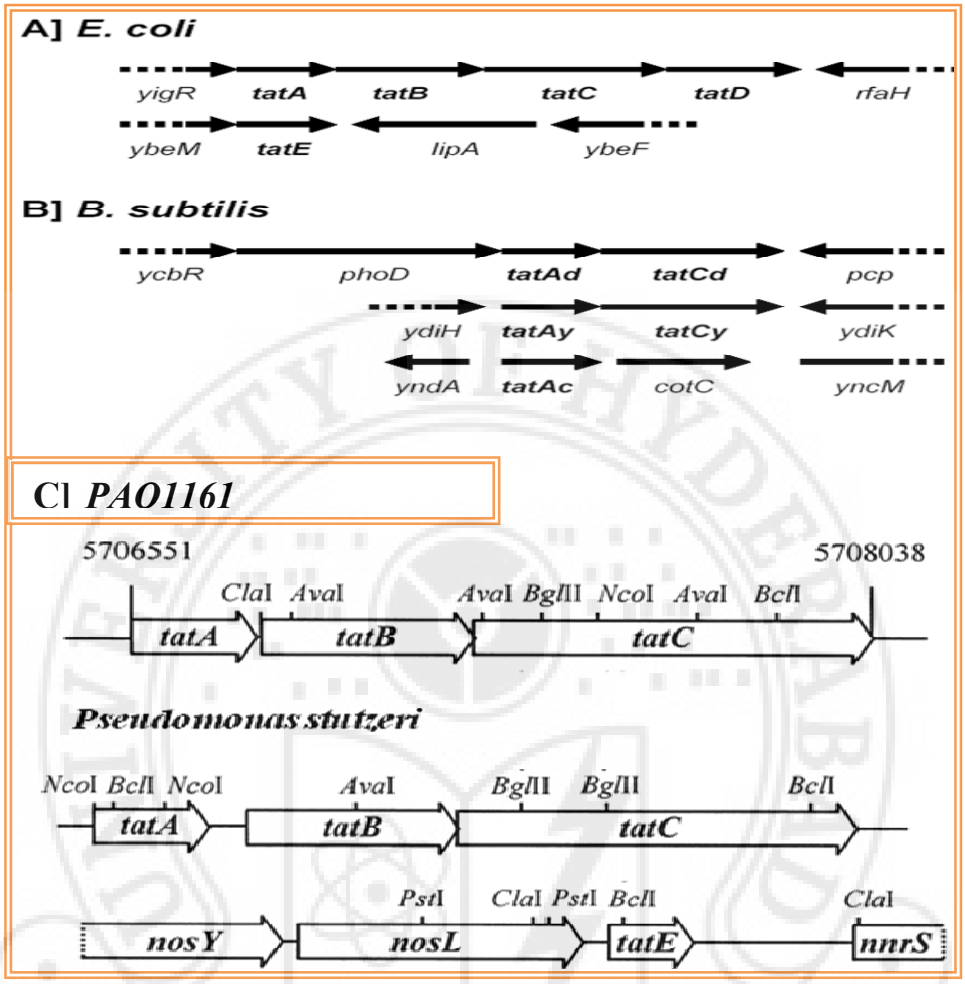
**Table. 1. 1. Components of the TAT pathway**

| Protein | Predicted size (kDa) | Characteristic | Reference |
|---|---|---|---|
| TatA | 9.6 or 11.3 [(a)] | 60% Homologous to TatE; its expression is higher than other *tat* genes | Gouffi et al., 2004; Porecelli et al., 2002; Sargent et al., 1998. |
| TatB | 18.4 [(a)] | Complexes with TatC and prevents its degradation | Sargent et al., 1999 |
| TatC | 28.9 [(a)] | Likely to be a signal peptide binding component | Allen et al., 2002; Behrendt et al., 2004; Buchanan et al., 2002. |
| TatD | 29.5 [(a)] | No effect on protein translocation; presents DNase activity | Wexler et al., 2000. |
| TatE | 6.9 [(b)] | Can partially substitute TatA | Berks et al., 2000. |

(a) Two possible translation initiation sites exist for *tatA* (Sargent et al., 1998).
(b) DNA sequence retrieved from Genbank accession no. NP_308692. Molecular mass calculated by the program PROTPARAM (http://www.expasy.ch/ tools/ protparam.html).

The Tat components are required for the twin arginine translocation pathway to transport prefolded proteins across the biological membrane and organization of *tat* genes in *E. coli, Bacillus subtilis and Pseudomonas* sp. were depicted in Fig. 1. 11. (Weiner et al., 1998; Sargent et al., 1998; Bogsch et al., 1998; Sargent et al., 1999; Jongbloed et al., 2000; Jan Maarten et al., 2002). The first three genes belong to the same operon, while *tatE* is located elsewhere on the bacterial chromosome (Bogsch et al., 1998; Sargent et al., 1998b; Weiner et al., 1998).

**Fig. 1. 11**

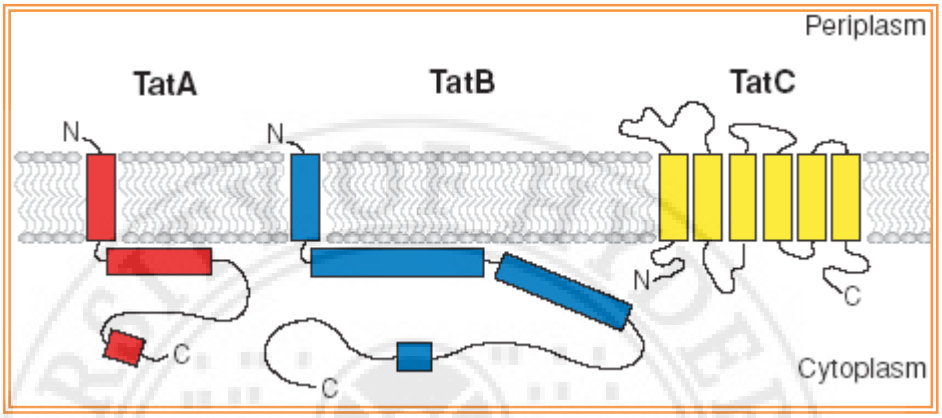



**Chromosomal organisation of the *tat* genes in *E. coli, B. subtilis* and *Pseudomonas sp*. A]** The *tatABCD* genes of *E. coli* are organised in an operon with the *tatE* gene located elsewhere on the chromosome. **B]** The genes for TatAd and TatCd that make up the TatAdCd translocase are located downstream from the *phoD* gene encoding the TatAdCd-dependent phosphodiesterase D. The genes for the second Tat translocase, *tatAy* and *tatCy*, are similarly organised in an operon elsewhere on the chromosome. The *tatAc* gene is monocistronic and does not have a *tatC* counterpart. This figure is adapted from Jongbloed et al., 2000). **C]** Physical maps of the DNA regions carrying *tat* genes in *Pseudomonas aeruginosa* and *Pseudomonas stutzeri.* Transcriptional directions are indicated by the open arrows (Jan Maarten et al., 2002 *Current Opinion in Microbiology*).

The *tatA* and *tatB* encode small proteins of 9.6 and 18.4 kDa, respectively, each of which contains a single transmembrane (TM) span (Sargent et al., 1998; Bolhuis et al.,

2001; Porcelli et al., 2002). TatC is a 28.9-kDa protein that was predicted to contain 6 trans membrane (TM) spans (Bogsch et al., 1998) (Fig. 1. 12).

**Fig. 1. 12**




The predicted structure and topology of the *E. coli* Tat components. Predicted helical regions are shown as boxes (Philip et al., 2006. *Annual Reviews of Microbiology*)

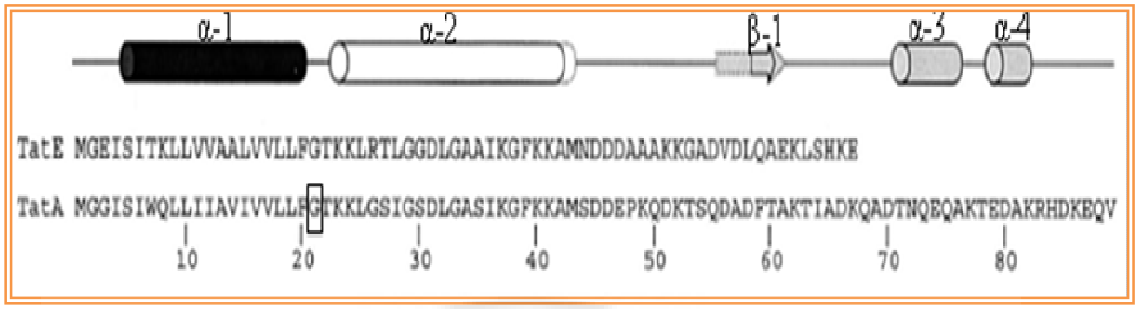
Deletion of either *tatB* or *tatC* leads to a complete loss of Tat-dependent protein export (Sargent et al., 1998; Bogsch et al., 1998), whereas the *ΔtatA* mutant can support a minimal level of export activity. This is due to the presence of fourth *tat* gene in *E. coli*, *tatE*, which encodes a TatA paralogue that is expressed at very low levels. Disruption of both *tatA* and *tatE* blocks export by the Tat pathway, but overexpression of *tatE* complements the *ΔtatA* mutant (Sargent et al., 1998; Sargent et al., 1999), confirming a similar basic function. However, the evolutionary rationale for maintaining both genes in *E. coli* is unclear since the TatE protein is expressed at such very low levels, and many Gram-negative bacteria possess only one *tatA*-type gene (Yen et al., 2002). As shown in the Fig. 1. 11 the *tat* operon encodes a fourth gene, designated *tatD*, but studies have shown this gene to encode a soluble protein with DNase activity that plays no apparent role in Tat-dependent translocation (Wexler et al., 2000).

**TatA**

The TatA protein is the most abundant of the Tat components and is estimated to be present at around 20 times more than TatB and TatC (Jack et al., 2001; Sargent et al., 2001). The TatA protein is predicted to have a structure similar to TatB, with an N-terminal trans membrane (TM) domain followed by a short hinge region leading to an amphipathic helical region and an unstructured soluble C-terminal tail (Fig. 1. 13). The *E. coli tatA* gene encodes a 9.6-kDa polypeptide of 89 amino acid residues. Structural predictions (Fig. 1. 13) suggest an N-terminal hydrophobic α-helix to residue 20, followed by a short hinge region and a longer amphipathic α-helix extending to amino acid 42. TatA is thought to function at a late stage in translocation and likely forms the major component of the Tat pore itself (Alami et al., 2003; Gohlke et al., 2005).

The TM domain of TatA is critical for export and is important for interactions with TatB (Lee et al., 2002; Dabney et al., 2003; Barrett and Robinson, 2005). The *tatE* gene encodes a protein of 67 residues with a predicted mass of 7.0 kDa. It shares almost 60% amino acid identity with TatA along its entire length and can functionally substitute for TatA. In keeping with the high percentage of amino acid identity, the positions of the predicted hydrophobic and amphipathic α-helical regions in TatE are indistinguishable from those in TatA (Fig. 1. 13).

**Fig. 1. 13**




Structural predictions for TatA and TatE. Secondary structural elements common to both proteins are shown with a solid line, and those unique to TatE are shown with a dotted line. The transmembrane α-helical domains of TatA and TatE are black, and the strongly amphipathic α-helical domains are white. Further predicted helical regions are in grey. (Philip et al., 2006). The invariant glycine residue (Gly21) is boxed (Porcelli et al., 2002 *Biochemistry*).

Philip et al., identified that although the C-terminal domain of TatA protein is not strictly necessary for its activity, but the transmembrane and adjacent helical regions play a critical role in its function (Philip et al., 2006). Both proteins have C-terminal regions that are predicted to be mainly random coil. Because of the high sequence and structural similarity between the two proteins, coupled with the previous studies that TatE is poorly expressed but functionally equivalent to TatA (Sargent et al., 1998; Sargent et al., 1999; Jack et al., 2001). In *E. coli* TatA is present at an approximately 20-fold molar excess over the other essential Tat components, TatB and TatC (Sargent et al., 2001; Jack et al., 2001). Chemical cross-linking studies have shown that TatA forms at least tetrameric homooligomers in the cytoplasmic membrane (Leeuw et al., 2001), while purification experiments suggest that TatA associates with TatB and TatC in at least two types of large (approximately 600 kDa) macromolecular complexes (Sargent et al., 2001; Bolhuis et al., 2001).

***The oligomeric state of TatA in Tat translocation***

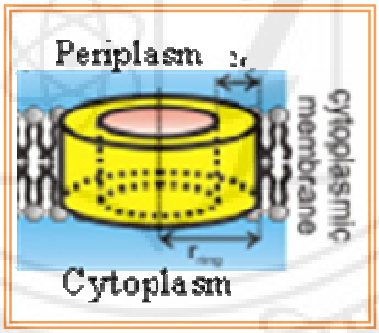
The TatA protein of *E. coli* is an 89-amino acid membrane protein. The secondary structural elements of this protein were predicted (Fig. 1. 13) using the program PSIPRED (Jones, 1999). TatA is strongly predicted to contain a transmembrane α-helix at the amino terminus (helix α-1) followed by a polypeptide segment containing a basic amphipathic α-helix (helix α-2). These predicted helical regions exhibit a high level of sequence conservation between species. However, only a glycine residue between the predicted helices is invariant (Fig. 1. 13) (Settles et al., 1997; Berks et al., 2000).

Detergent-solubilized TatA is found as large, homooligomeric complexes of variable size (Porcelli et al., 2002; Oates et al., 2005; Gohlke et al., 2005; McDevitt et al., 2006). Low resolution structures of purified TatA show doughnut-shaped particles with an internal cavity large enough to accommodate substrates (Gohlke et al., 2005). This and other evidence (Cline and Mori, 2001; Mori and Cline, 2002) suggest that TatA forms the translocating element of the Tat system. The number of chemical cross-links that can be formed between TatA molecules increases under transport conditions (Dabney et al., 2006). These observations have led to the suggestion that substrate interaction with TatBC triggers TatA polymerization (Dabney et al., 2006; Cline and Theg, 2007).

Since cross-linking studies are unable to define the stoichiometry and stoichiometry distribution of the TatA complexes that are present in the membrane, Leake et al., from University of Oxford have used "*in vivo* single-molecule imaging" technique to visualize individual yellow fluorescent protein-labeled TatA (TatA-YFP) complexes expressed at native levels in living *E. coli* cells by using fluorescence microscopy to determine TatA

Stoichiometry (Leake et al., 2008). Their experimental studies also suggests that the TatA-YFP complexes are made up of tetrameric units (4.3 ± 0.6 molecules) of TatA molecules and forms a ring model of translocating pore (Fig. 1. 14). Since it was previously shown that TatA interaction with substrate-bound TatBC complexes requires the transmembrane proton motive force (pmf) (Mori and Cline, 2002; Alami et al., 2003) and does substrate induced structural reorganization of TatA (Dabney et al., 2006), Leake et al., extended their studies and determined the role of pmf on stiochiometry of the TatA complex and its structural features in the membrane by treating the cells with protonophore carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP) which will abolishe the transmembrane pmf in the membrane (Leake et al., 2008) and determined that the oligomeric state of TatA is unaffected by the removal of the pmf shows that pmf is not required to maintain the polymerized state of TatA.

**Fig. 1. 14**




**Schematic illustration of the ring model of TatA complex** (**Leake** et al., 2008 *Proceeding of the Natural Academy of Sciences*)

In addition, the substrate-bound TatBC complex triggers the TatA and TatA polymerizes around the substrate molecule to produce a pore that matches the size of the substrate and this observeration is in accordance with different-sized TatA oligomers were

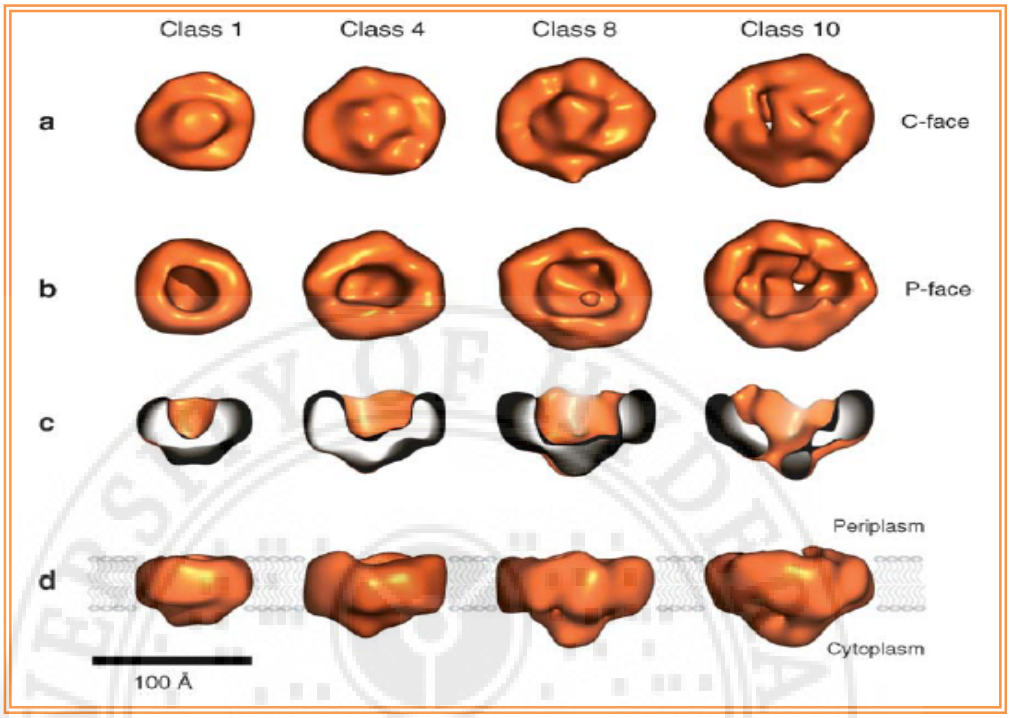
determined in membrane using YFP labeled TatA molecules by applying *in vivo* single molecule imaging analysis tenchinque using fluorescence microscopy (Leake et al., 2008).

***Topological organization of TatA***

Previous observations based on electron microscopy studies revealed that TatA has domains on either side of the membrane (Sargent et al., 2001). To test this hypothesis, Porcelli et al., attempted to determine the topological organization of TatA using protease accessibility measurements by isolating spheroplasts and the inner membrane vesicles which were prepared from cells expressing TatA. These membrane preparations were subjected to proteinase K treatment, and the fate of the TatA molecules present was assessed by immunoblotting.

The TatA in spheroplasts was insensitive to proteinase K treatment and suggests that TatA is not accessible from the periplasmic side of the membrane. Control experiments demonstrate that permeabilization of the spheroplast membrane with the detergent Triton X-100 renders TatA susceptible to proteolysis by proteinase K. These observations suggests that TatA is exposed at the cytoplasmic face of the membrane (Porcelli et al., 2002) and obeying the Positive-Inside rule (von Heijne, 1992). Recently Philip et al., obtained a higher-resolution three-dimensional image of a Tat complex using random canonical tilt electron microscopy (Fig. 1. 15) which also obeys the positive-Inside rule (Philip et al., 2006).

**Fig. 1. 15**

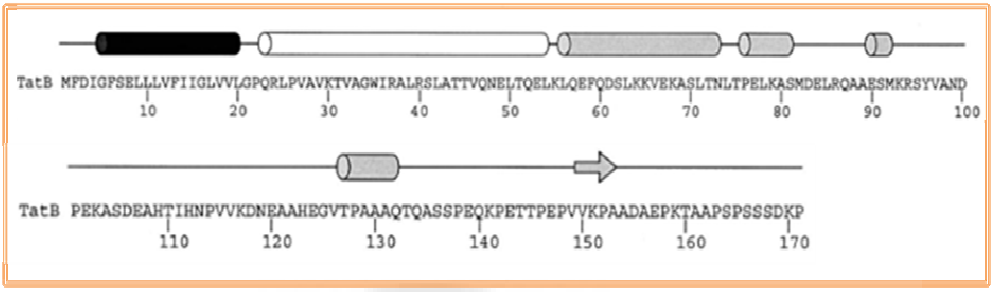



Three-dimensional density maps of TatA complexes. (*a*) TatA complexes viewed from the closed end of the channel proposed to be at the cytoplasmic side of the membrane (C-face). (*b*) TatA complexes viewed from the open end of the channel proposed to be at the periplasmic side of the membrane (P-face). (*c*) Side views of TatA. The front half of each molecule has been cut away to reveal internal features. (*d)* Views of TatA parallel to the membrane plane. The proposed position of the lipid bilayer is indicated in gray. (Scale bar, 100 ¢ª A.) (Philip et al., 2006. *Annual Reviews of Microbiology*)

### TatB

The *E. coli tatB* gene encodes a protein of 171 amino acids with a predicted molecular mass of 18.4 kDa. Structural prediction analysis of TatB (Fig. 1. 16) indicates an N-terminal hydrophobic α-helix, similar to those of TatA and TatE, extending to residue 20. This is followed by a helical region extending to amino acid 81, of which at least the first section is predicted to be strongly amphipathic. The C-terminal portion is predicted to be mainly random coil.

**Fig. 1. 16**




Structural prediction for TatB. α-Helical regions are represented as cylinders, and β-sheets are represented as arrows. The transmembrane α-helical domains of TatB are black, and the strongly amphipathic α-helical domains are white. (Philip et al., 2006 *Journal of Bacteriology*).

TatB shares some sequence homology with TatA/E (approximately 25% identity at the amino acid level) but is considerably longer. TatB is an essential component of the *E. coli* Tat machinery. Thus, an in-frame deletion of the *tatB* gene leads to a complete block in the twin arginine translocation protein export (Philip et al., 2006). But this is exceptional in the case of *Bacillus subtilis* where it doesn't contain any *tatB* genes (Fig. 1. 11) and, *tatA* acts as bifunctional which is able to carry out the function of both *E. coli* TatA and TatB (Jongbloed et al., 2006; Barnett et al., 2008).

TatB was shown *in vitro* to contact the entire length of the signal peptide and also the mature protein more than 20 residues away from the signal peptide cleavage site (Alami et al., 2003). Such interactions were seen only when TatC was present, suggesting that substrate targeting involves a series of sequential interactions, with TatC forming the primary recognition site before the substrate is transferred to TatB. In turn TatB could be considered a mediator between TatC and TatA, contacting the substrate after initial recognition by TatC and then potentially involved in transfer to a complex consisting mainly of the TatA protein proceeding to translocation (Sargent et al., 2001; Alami et al., 2003).

**TatC**

TatC is an essential, largest and most highly conserved component of the Tat system in bacteria and chloroplasts (Bogsch et al., 1998). The *E. coli* TatC is a 258 amino-acid polytopic membrane protein with six Trans Membrane Domains (TMD) with the amino and carboxyl termini located at the cytoplasmic face of the membrane (Behrendt et al., 2004; Ki et al., 2004). Chloroplast TatC (cpTatC) has been demonstrated to bind the twin-arginine motif of Tat signal sequences, and the signal sequence can remain bound to TatC during transport without affecting the functionality of the Tat system (Gerard et al., 2006). Similar to the plant system, bacterial TatC binds the region of the twin-arginine motif of Tat substrates (Alami et al., 2003). Important residues have been identified in TatC, but so far the exact functions of these positions are unknown (Allen et al., 2002; Buchanan et al., 2002; Barrett et al., 2005).

*E. coli* TatC has several weakly conserved protonatable residues in its TMDs. At the position E170 also glutamine residues can be found at the E227 position, tryptophan and phenyl alanine residues are present in orthologs of some phyla. In *E. coli*, the TatC E170A and E227A mutants did not turn into the functional Tat system (Buchanan et al., 2002). The mutation of a more conserved glutamate residue (E103) in cytoplasmic loop of TatC was reported to inactivate the Tat system (Buchanan et al., 2002). However, in other assay systems, this mutation did not affect transport (Barrett et al., 2005). One conserved aspartate (D211) has been reported to be essential for Tat transport of TMAO reductase (Buchanan et al., 2002). This residue is not positioned within a TMD and its mutation does not affect SDS-sensitivity. Of all of the components of the Tat pathway it is TatC that

shows the highest level of amino acid conservation. Twenty-one amino acids, several of which are polar residues, are strictly conserved amongst the eubacterial TatC proteins and seven of these are also conserved amongst the eukaryotic homologues (Buchanan et al., 2002; Allen et al., 2002). The majority of these conserved residues fall within predicted cytoplasmic loops of the protein. Recent site-directed mutagenesis experiments have confirmed an essential role for some of these residues in the operation of the Tat pathway (Allen et al., 2002; Buchanan et al., 2002).

The cytoplasmic side of TatC seems to be particularly important for function. Several residues within cytoplasmic loops are completely or functionally conserved across prokaryotes, chloroplasts, and plant mitochondria, and site-directed mutagenesis of a number of these residues has revealed that they are important or essential for TatC function (Buchanan et al., 2002; Allen et al., 2002; Barrett and Robinson, 2005). In vitro biochemical studies have revealed that TatC serves as the initial docking site for Tat signal peptides. Signal peptides from bacteria and chloroplasts interact with TatC either alone or in a complex with TatB (Cline and Mori, 2001; De Leeuwet al., 2002; Alami et al., 2003). It is likely that more than one TatC monomer is required for each targeting event, since two mutant versions of TatC that alone blocked Tat-dependent transport support transport when expressed together (Buchanan et al., 2002). Based on cross-linking of *in vitro* translocated proteins into inverted membrane vesicles, Alami et al. discovered that TatC in particular exhibits extensive contact with the signal peptide and recognizes the RR motif (Cline and Mori, 2001; Alami et al., 2003).

### 1.2.4.2.2. The Translocation Pore

The evolution of specific protein translocators solved an essential problem for cells, allowing proteins to be moved across or inserted into the lipid bilayers. The Tat translocation pore of both the *E. coli* cytoplasmic membrane and the chloroplast thylakoid membrane comprises three integral membrane proteins, TatA, TatB, and TatC (in chloroplasts also called Tha4, Hcf106, and cpTatC, respectively). In *E. coli* and some enterobacteria, an additional protein highly identical to TatA is encoded by the *tatE* gene (Wu et al., 2000) which is most probably the result of a cryptic gene duplication (Jack et al., 2001). TatA and TatB both are anchored in the membrane by one N-terminal transmembrane helix. Structural analyses predict at least one subsequent amphiphilic helix for each protein. The amphiphilic helix of TatA being distinctly positive interacts with anionic phospholipids and forms upon interaction with liposomes (Porcelli et al., 2002).

An invariant Glycine seems to be part of a flexibility-conveying hinge region located in between the transmembrane and amphiphilic helices of TatA and TatB (Barrett et al., 2003). A transmembrane Glu residue that is found conserved among the TatA and TatB proteins of plastids and that is lacking from the bacterial TatA, seems to play a role in the assembly of TatA into the translocase (Dabney-Smith et al., 2003). The relative stoichiometric ratio between TatA:TatB:TatC in the *E. coli* cytoplasmic membrane has been determined to be approximately 20-30:1:0.4 (Berks et al., 2003). Complexes of about 600-700 kDa containing varying amounts of TatA, TatB, and TatC were isolated from solubilized cytoplasmic membranes of a variety of bacterial organisms as well as from the

thylakoids of plant chloroplasts (Berghofer and Klosgen, 1999; Bolhuis et al., 2001; Cline and Mori, 2001; de Leeuw et al., 2002; Oates et al., 2003; Sargent et al., 2001).

It has been predicted that the Tat pore may be able to attain a diameter of up to 50–60 $A^0$ when fully open which helps to explain how the Tat system can accommodate a wide range of folded, native-like proteins with molecular weights as large as 120 kDa (Robinson and Bolhuis, 2001; Sargent et al., 2001; Berks et al., 2000). The isolation of different Tat protein complexes, together with experimental evidence from both the bacterial and cpTat pathways, has led to the idea of a modular and highly dynamic system whereby the Tat proteins exist as separate complexes in the resting state and come together to form a complete translocation pore upon substrate binding (Alami et al., 2003) and complexes of varying sizes may form pores that match the different sizes of folded Tat substrate proteins (Lee et al., 2006).

The first structural view of the Tat translocon was obtained by Sargent et al. (Sargent et al., 2001) using negative stain electron microscopy. Pore-like structures derived from a purified TatAB complex that contains an approximate 20 times molar excess of TatA to TatB and no detectible TatC were observed. The external diameter of these structures was 155 to 160 $A^0$, with an internal diameter of around 65 $A^0$ containing one to two density features, possibly forming a gate to the pore. Single-particle electron microscopy was also used to obtain low-resolution structures of TatABC complexes from three different bacteria expressed in *E. coli*. Similar structures were observed for the TatABC proteins from the different species, indicating structural conservation.

### 1.2.4.2.3. Tat Substrates

The most prominent feature of the Tat pathway is its capability to transport fully folded proteins and the protein which has consensus invariant twin arginine motif (S/T**RR**xFLK) in their signal peptides (Berk et al., 2005). Prior to export, the majority of bacterial Tat substrates undergo a complex cytosolic incorporation of cofactors like molybdopterin, FAD, NADP, iron-sulphur and iron-nickel clusters, copper and others (Berks et al., 2003). Cofactor insertion is, however, not a prerequisite for being exported by the Tat pathway. Several cofactor-less hydrolases were shown to use the Tat pathway. Rapid folding kinetics inherent in these enzymes might be the reason why they evolved as passengers of the Tat pathway (Jongbloed et al., 2000; Angelini et al., 2001; Voulhoux et al., 2001).

### 1.2.4.2.4. Tat mechanism

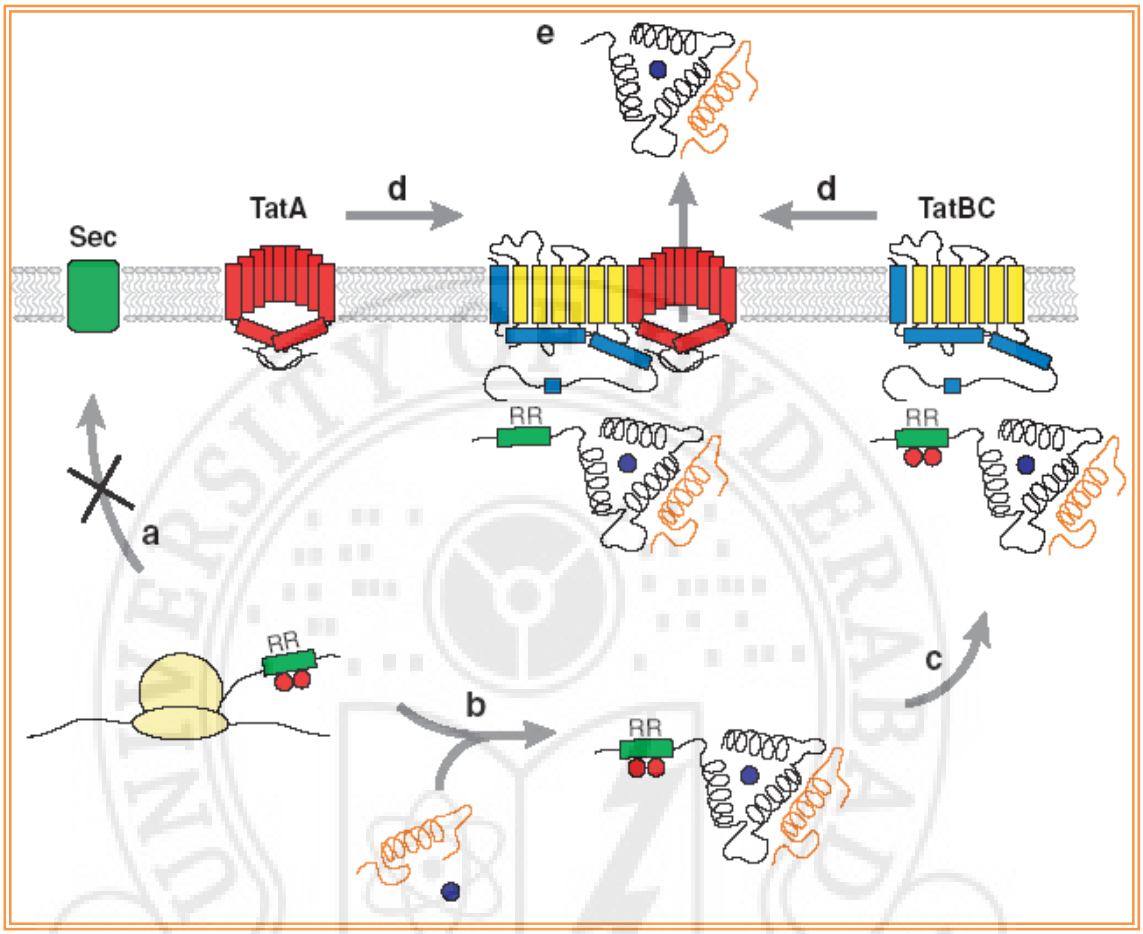
The Tat pathway is an alternate secretory pathway present in chloroplasts and many prokaryotic organisms. It is important for a wide range of biological processes, including cell division, bacterial pathogenesis, anaerobic respiration, degradation and acquisition of organic and inorganic compounds (Santini et al., 1998; Ochsner et al., 2002; Pop et al., 2002; Ding et al., 2003; Ize et al., 2003). Unlike the Sec pathway, Tat is neither ubiquitous nor essential for viability in most organisms tested (Jongbloed et al., 2000; Wu et al., 2000; Voulhoux et al., 2001; Ding et al., 2003).

Proteins using Tat pathway contain a typical highly invariant twin-arginine motif in their signal peptide (Chaddock et al., 1995; Berks, 1996) and, therefore, it was designated

as the twin-arginine translocation (Tat) pathway. In contrast to the Sec pathway, the Tat system has the unique ability to transport fully folded proteins and does not depend on the presence of nucleoside triphosphates (Robinson and Bolhuis, 2001). This pathway is involved in the export of proteins that either have to fold before translocation, such as certain co-factor containing proteins, or just fold too quickly. These folded proteins are incompatible with the Sec machinery and can only be exported via the Tat system. Most bacteria and chloroplasts contain three major components that are required for the Tat pathway. These are TatA (Tha4 in chloroplasts), TatB (Hcf106 in chloroplasts) and TatC (Bogsch et al., 1998; Sargent et al., 1998; Weiner et al., 1998; Walker et al., 1999).

As and when the pre-proteins of Tat substrates emerges from the ribosome the preprotein must avoid targeting to other pathways such as Sec, which is possible by the characteristics of the signal peptide and mature protein (Cristobal et al., 1999) or the binding of Tat-specific chaperones (Jack et al., 2004; Turner et al., 2004). After folding in the cytoplasm, any cofactors or additional subunits are added to the pre-proteins prior to targeting to the TatBC receptor complex (Cline and Mori, 2001; De Leeuwet al., 2002; Alami et al., 2003). The proton motive force drives the formation of an active translocase and the substrate is transported through a pore consisting mainly of TatA (Mori and Cline, 2002; Alami et al., 2003; Gohlke et al., 2005). Finally the folded Tat substrates will be reached to the trans side of the membrane or anchored to the membrane depending upon removal of the signal peptide by the signal peptidases. The model Tat transport cycle in *E. coli* was represented in Fig. 1. 17.

**Fig. 1. 17**




**A schematic model of Tat targeting and transport in *Escherichia coli*. (Philip et al., 2006.** *Annual Reviews of Microbiology*)

### 1.3. Definition of the problem

Mulbry and his associates have shown existence of a 29 amino acid long signal peptide in the OPH purified from *Flavobacterium* sp. (Mulbry and Karns, 1989). Similar situation is shown to exist in *B. diminuta* (Mulbry et al., 1986). On careful perusal of its structure we have noticed existence of a sequence that resembles a typical twin arginine transport (Tat) motif (MQT**RR**VVLK). The Tat motif is seen only in membrane and extracellular proteins targeting or transporting across the membrane in a prefolded form.

These are mainly proteins requiring large cofactors (Berks, 1996; Palmer et al., 2005). The OPH is an esterase depending on metal ions for its activity. In fact existence of a Tat motif in a protein that uses small molecules like metal ions as cofactor is rather unusual. However the presence of Tat motifs in most of the reported OPH sequences signifies its functional relevance. The present study is designed to investigate the requirement of Tat motif for membrane targeting of OPH in *B. diminuta.*

# CHAPTER - II

## To study the topology of OPH in the inner membrane of *B. diminuta*

## 2. 1. Materials

### 2.1.1. Antibiotics

| Name of the antibiotic | Name of the Supplier |
|---|---|
| Ampicillin sodium salt | Amersham Pharmacia Biotech, UK |
| Chloramphenicol | Amersham Pharmacia Biotech, UK |
| Kanamycin Sulfate | Amersham Pharmacia Biotech, UK |
| PolymixinB | SRL Biochems |
| Streptomycine | Amersham Pharmacia Biotech, UK |
| Tetracycline hydrochloride | Amersham Pharmacia Biotech, UK |

### 2.1.2 Biochemicals

| Name of the Chemical | Name of the Supplier |
|---|---|
| Absolute alcohol | Qualigens |
| Acetic acid (Glacial) | Qualigens |
| Acrylamide | Amersham Pharmacia Biotech,UK |
| Agarose | Sigma Aldrich Chemicals |
| Ammonium persulphate | Amersham Pharmacia Biotech,UK |
| Boric acid | Qualigens |
| Bovine serum albumin | Amersham Pharmacia Biotech, UK |
| Bromophenol blue | Amersham Pharmacia Biotech,UK |
| Butanol | Qualigens |

| | |
|---|---|
| Calcium chloride | Qualigens |
| Chloroform | Qualigens |
| Cobalt chloride | SRL |
| Coomassie Brilliant blue R | Amersham Pharmacia Biotech,UK |
| Copper sulfate | Sigma Aldrich Chemicals |
| D-Glucose-6-phosphate | Sigma Aldrich Chemicals |
| Deoxynucleotide triphosphates | Amersham Pharmacia Biotech,UK |
| Diethyl pyrocarbonate (DEPC) | Sigma Aldrich Chemicals |
| Dipotassium hydrogen orthophosphate | Merck |
| DNA molecular weight marker | MBI Fermentas |
| Ethidium bromide | Amersham Pharmacia Biotech,UK |
| Ethylenediaminotetraacetic acid disodium salt (EDTA) | Amersham Pharmacia Biotech,UK |
| FM4-64 (N-(3-triethylammoniumpropyl)-4-(pdiethylaminophenylhexatrienyl)-pyridinium 2Br) | Sigma Aldrich Chemicals |
| Folin-Ciocalteau reagent | Qualigens |
| Glucose | Qualigens |
| Glycerol | Qualigens |
| Glycine | Merck |
| Glycylglycine | Sigma Aldrich Chemicals |
| Hydrochloric acid | Qualigens |

| | |
|---|---|
| Isopropyl thiogalactopyranoside (IPTG) | Amersham Pharmacia Biotech, UK |
| Lysozyme | Amersham Pharmacia Biotech, UK |
| Magnesium chloride | Qualigens |
| Magnesium sulfate | Qualigens |
| Methanol | Qualigens |
| Methyl parathion | Sigma Aldrich Chemicals |
| Methyl viologen | Sigma Aldrich Chemicals |
| NADP (Nicotinamide Adenine Dinucleotide Phosphate) | Sigma Aldrich Chemicals |
| N-1-nathylene-diamine dihydrochloride | Sigma Aldrich Chemicals |
| N,N'-Dimethylformamide | Amersham Pharmacia Biotech, UK |
| N,N'-Methylene bis acrylamide | Amersham Pharmacia Biotech, UK |
| ONPG (*O*-nitrophenyl-β-D-galactopyranoside) | Sigma Aldrich Chemicals |
| *p*-nitrophenylphosphate | Sigma Aldrich Chemicals |
| Peptone | Himedia |
| Phenol (water saturated) | GENEI |
| Phosphoric acid | Merck |
| Poly-L-Lysine solution | Qualigens |
| Potassium acetate | Qualigens |
| Potassium chloride | Qualigens |
| Potassium dihydrogen Phosphate | Merck |

| | |
|---|---|
| Potassium nitrate | Merck |
| PSMF | Sigma Aldrich Chemicals |
| Sodium acetate | Qualigens |
| Sodium chloride | Qualigens |
| Sodium citrate | Qualigens |
| Sodium dihydrogen orthophosphate | Qualigens |
| Sodium dodecyl sulfate | Amersham Pharmacia Biotech, UK |
| Sodium hydrogen orthophosphate | SRL |
| Sodium hydrosulfite | Sigma Aldrich Chemicals |
| Sodium hydroxide | Sigma Aldrich Chemicals |
| Sucrose | SRL |
| Tetra ethyl methylene diamine (TEMED) | Amersham Pharmacia Biotech, UK |
| Tris-base | Amersham Pharmacia Biotech, UK |
| Tween-20 | GENEI |
| X-gal | Amersham Pharmacia Biotech, UK |
| Xylene cyanol | Amersham Pharmacia Biotech, UK |
| Yeast extract | Himedia |
| Zinc chloride | SRL |
| β-mercatoethanol | Amersham Pharmacia Biotech, UK |

### 2.1.3. Restriction and other enzymes

| Name of the enzyme | Name of the Supplier |
|---|---|
| *Bam* HI | Amersham Pharmacia Biotech, UK |
| *Bgl* II | Amersham Pharmacia Biotech, UK |
| *Dpn*I | MBI Fermentas |
| *Eco* RI | Amersham Pharmacia Biotech, UK |
| *Hind*III | MBI Fermentas |
| *Nde*I | Amersham Pharmacia Biotech, UK |
| *Pst*I | MBI Fermentas |
| *Sac*I | MBI Fermentas |
| *Sal* I | Amersham Pharmacia Biotech, UK |
| *Xho*I | Amersham Pharmacia Biotech, UK |
| Proteinase K | Amersham Pharmacia Biotech, UK |
| RNase A | Amersham Pharmacia Biotech, UK |
| $T_4$ DNA Ligase | Amersham Pharmacia Biotech, UK |
| *Taq* DNA polymerase | MBI Fermentas |

## 2.2. Bacterial strains

| Name of the strain | Genotype / phenotype | Source |
|---|---|---|
| *E. coli* DH5α | *supE44 ΔlacU169 (Δ80 lacZ ΔM15) hsdR17 recA1 endA1 gyrA96 thi1 relA1* | Hanahan, 1983. |

| | | |
|---|---|---|
| *E. coli* BL21 | *hsdS gal(ΔcIts857 ind1 Sam7 nin5) lacUV5 T7 gene1* | Studier and Moffat, 1986. |
| *E. coli* S17-1 | *thi pro hsdR hsdM recA RP4 2-Tc::Mu-Km$^r$::Tn7 (Tp$^r$, Sp$^r$, Sm$^r$)* | Simon et al., 1983. |
| *E. coli* MC4100 | *F$^-$ ΔlacU169 ara D139 rpsL150 relA1 ptsF rbs flbB5301* | Casadaban and Cohen, 1979. |
| *E.coli* MC4100 (Δ*tatA/E*) | *F$^-$ ΔlacU169 ara D139 rpsL150 relA1 ptsF rbs flbB5301 Δtat A/E* | Sergeant et al., 1998. |
| *E. coli* MC4100 (Δ*tatB*) | *F$^-$ ΔlacU169 ara D139 rpsL150 relA1 ptsF rbs flbB5301 Δtat B* | Sargent et al., 1999 |
| *E. coli* MC4100 (Δ*tatC*) | *F$^-$ ΔlacU169 ara D139 rpsL150 relA1 ptsF rbs flbB5301 Δtat C* | Bogsch et al., 1998. |
| *B. diminuta* | Sm$^r$, PmB$^r$, *opd*$^+$ | Serdar et al., 1982. |
| *B. diminuta* (*opd::tet*) | Sm$^r$, Tc$^r$, PmB$^r$, *opd*$^-$ | This work |

**2. 3. Plasmids**

| Plasmid | Relevant Genotype | Reference |
|---|---|---|
| pMMB206 | Cm$^r$, broad host range, low copy number, expression vector | Morales et al., 1991. |
| pSUP202 | Mob$^+$, Amp$^r$, Cm$^r$, Tc$^r$ and mobilizable suicidal vector | Simon et al., 1983. |
| pDHB5700 | Cm$^r$, expression vector | De Gier et al., 1998. |
| pBQGFP | Amp$^r$, *gfp* cloned in pUC18 | Gift from Dr. J.R. Wild |

| | | |
|---|---|---|
| pUC4KIXX | $Amp^r$, $Km^r$, an omega fragment containing kanamycin cassette flanked with transcription terminator sequences | Mazodier et al., 1985. |
| pGEMT-Easy | $Amp^r$, TA cloning vector | PROMEGA, USA. |
| pMW18 | $Amp^r$, contains *pretorA*-23K | Bogsch et al., 1998. |
| pOPS400 | $Amp^r$, Signal peptide coding region of *opd* gene cloned in pMW18 as *Pst*I-*Eco*RI fragment to generate inframe fusion with 23K protein coding sequence | This work |
| pTSL400 | $Cm^r$, *opd*-23K fusion in expression vector pDHB5700 as *Sal*I and *Sac*I fragment, codes for preOPH-23K protein | This work |
| pHYS400 | $Amp^r$, complete *opd* gene encoding preOPH cloned in pET23b as *Nde*I - *Xho*I fragment | Siddavattam et al., 2006. |
| pHNS400 | $Amp^r$, *opd* sequence encoding mature OPH cloned in pET23b as *Nde*I - *Xho*I fragment | Pandey et al., 2009. |
| pSM5 | $Cm^r$, *opd* gene coding preOPH6His cloned as an *Eco*RI-*Bam*HI fragment in pMMB206 under the control of p*tac* promoter | Siddavattam et al., 2003. |
| pHLNS400 | $Cm^r$, *opd* gene encoding mature OPH6His cloned as an *Eco*RI- *Bam*HI fragment in pMMB206 under the control of p*tac* promoter | This work |
| pHIR400 | $Amp^r$, contains *opd* variant *opd'* having nucleotide changes at every third base of the codon into a non-complementary base in the inverted repeat (IR) sequence of *opd* gene. The *opd'* is cloned as *Nde*I-*Xho*I fragment in pET23b | This work |
| pSM7 | $Tet^r$, upstream region of *opd* gene cloned in pMP220 as *Bam*HI-*Pst*I fragment | Manavathi et al., 2005. |
| | | |

| | | |
|---|---|---|
| pSM8 | $Tet^r$, Upstream region of *opd* gene cloned in pMP220 as *Bam*HI-*Sph*I fragment | Manavathi et al., 2005. |
| pSM9 | $Tet^r$, Upstream region of *opd* gene cloned in pMP220 as *Bgl*II-*Sph*I fragment | Manavathi et al., 2005. |
| pSM10 | $Tet^r$, Upstream region of *opd* gene cloned in pMP220 as *Bam*HI-*Xba*I fragment | This work |
| pSM11 | $Tet^r$, $Km^r$, omega fragment (kanamycin cassette having transcription termination sequences on either side) was cloned at *Bgl*II site found between putative σ70 and σ54 promoters | This work |
| pCMS1::*tet* | pCMS1 containing an *opd::tet* insertion | This work |
| pPHRA400 | $Cm^r$, derivative of pSM5 encoding OPH R5A | This work |
| pPHRK400 | $Cm^r$, derivative of pSM5 encoding OPH R5K | This work |
| pPHQQ400 | $Cm^r$, derivative of pSM5 encoding OPH R4Q,R5Q | This work |
| pPHKK400 | $Cm^r$, derivative of pSM5 encoding OPH R4K,R5K | This work |
| pSM3 | $Amp^r$ $Tc^r$, *tet* gene inserted into *opd* coding region (*opd*::*tet*) | Siddavattam et al., 2003. |
| pSUP*opd::tet* | $Cm^r$ $Tc^r$, *opd*::*tet* cloned in pSUP202 as a *Pst*I fragment | This work |
| pGEM-GFP | $Amp^r$, *gfp* amplicon generated from pBQGFP was cloned in pGEMT-Easy TA cloning vector | This work |
| pYSGFP400 | $Cm^r$, *opd'-gfp* derived by cloning *gfp* as a *Sal*I-*Hin*dIII fragment in pSM5 codes for preOPH-GFP protein | This work |
| pYSGFPD400 | $Cm^r$, *opd'-gfp-ssrA* derived by cloning *gfp-ssrA* as a *Sal*I-*Hin*dIII fragment in pSM5 codes for preOPH-GFP with C-terminal SsrA signal | This work |

| | | |
|---|---|---|
| pNSGFP400 | $Cm^r$, *gfp,* derived by cloning *gfp* as a *Sal*I-*Hind*III fragment in pHLNS400 codes for mOPH-GFP protein | This work |
| pNSGFPD400 | $Cm^r$, *gfp-ssrA,* derived by cloning *gfp-ssrA* as a *Sal*I-*Hind*III fragment in pHLNS400 codes for mOPH-GFPSsrA | This work |

**2. 4. Preparation of media**

The following media was used to propagate bacteria. The media was sterilized by autoclaving for 20 min at 15 lb/Sq. When required appropriate amounts of antibiotics of Ampicillin (100 µg/ml), Tetracyclin (25 µg/ ml), Chloramphenicol (30 µg/ ml), Streptomycin (20 µg/ml), PolymixinB (10 µg/ml) and Kanamycin (25 µg/ml) were added after cooling the media to 45$^oC$.

**2.4.1. Luria Bertani (LB) broth**

The LB medium was prepared by dissolving 10 gm of peptone, 5 gm of yeast extract and 10 gm of NaCl in 500 ml of distilled water. The contents were stirred and finally made up to 1 liter with distilled 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.

**2.5. REPARATION OF STOCK, WORKING SOLUTIONS AND BUFFERS**

**2.5.1. Ampicillin**

Ampicillin stock solution was prepared by dissolving 100 mg of ampicillin in 1 ml of double distilled water and then filter sterilized. The sterilized stock solution was stored in

aliquots at -20°C. When required 10 µl of ampicillin stock solution is added to the 10 ml of medium to get the working concentration of 100 µg/ml.

**2.5.2. Chloramphenicol**

Chloramphenicol stock solution was prepared by dissolving 30 mg of chloramphenicol in 1 ml of 100% ethanol. The stock solution was stored in aliquots at -20°C. When required 10 µl of chloramphenicol stock solution is added to the 10 ml of medium to get the working concentration of 30 µg/ml.

**2.5.3. Tetracycline**

Tetracycline stock solution was prepared by dissolving 25 mg of tetracycline hydrochloride in 1 ml of 70% ethanol/water and stored in aliquots at -20°C. When required 10 µl of tetracycline stock solution is added to the 10 ml of medium to get the working concentration of 25 µg/ml.

**2.5.4. Kanamycin**

Kanamycin stock solution was prepared by dissolving 25 mg of kanamycin sulfate in 1 ml of double distilled water. The stock solution was stored in aliquots at -20°C until further use. When required 10 µl of kanamycin stock solution is added to the 10 ml of medium to get the working concentration of 25 µg/ml.

**2.5.5. PolymixinB**

PolymixinB stock solution was prepared by dissolving 10 mg of polymixinB in 1 ml of double distilled water. The stock solution was stored in aliquots at -20°C. When required 10 µl

of polymixinB stock solution is added to the 10 ml of medium to get the working concentration of 20 µg/ml.

**2.5.6. Streptomycin**

Streptomycin stock solution was prepared by dissolving 20 mg of streptomycin in 1 ml of double distilled water. The stock solution was stored in aliquots at -20°C. When required 10 µl of streptomycin stock solution is added to the 10 ml of medium to get the working concentration of 20 µg/ml.

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

Stock solution of methyl parathion was prepared by dissolving 263.23 in 1 ml of methanol to get 1 M concentration and the stock solution was stored at -20°C until further use. Whenever needed for parathion hydrolase enzyme assay 100 µM of methyl parathion was used from the stock solution.

**2.5.8. ONPG solution (0.4%)**

20 mg of *O*-nitrophenyl-β-D-galactopyranoside (ONPG) was dissolved in few ml of 0.1 M phosphate buffer (pH 7.0) and the volume was made up to 5ml using phosphate buffer.

**2.5.9. Z-buffer**

16.1 gm of $Na_2HPO_4$. $7H_2O$, 5.5 gm of $NaH_2PO_4.H_2O$, 0.75 gm of KCl, 0.264 gm of $MgSO_4.7H_2O$ and 7ml of β-mercaptoethanol were dissolved in 500 ml of distilled water and pH

of the solution was adjusted to 7.0. Finally the volume of the buffer was made up to 1000 ml with distilled water and stored at 4°C until further use.

**2.5.10. SET buffer**

17.1 gm of sucrose, 0.186 gm of EDTA and 0.605 gm of Tris were dissolved in 50 ml of double distilled water and pH of the solution was adjusted to 8.0 using 1N HCl. Finally the volume of the buffer was made up to 100 ml using Millipore pore water.

**2.5.11. IPTG**

1 M IPTG stock solution was prepared by dissolving 238.3 mg IPTG in 1 ml of double distilled water and filter sterilized using 0.2 µm Sartorius filter. The stock solution was stored in aliquots at -20°C. When required the stock solution was thawed on ice for few minutes and adequate amount is added to the medium to get 1 mM working concentrations.

**2.5.12. X-Gal**

2% of X-gal stock solution was prepared by dissolving 40 mg of X-gal in 1 ml of N,N'-dimethylformamide. When necessary 1µl of stock solution is added to 1 ml of medium.

**2.6. PREPARATION OF SOLUTIONS AND BUFFERS FOR SDS-PAGE**

**2.6.1. 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 double distilled water and the solution was made up to 100 ml with double distilled water. The contents were then filtered and the solution was stored at $4^{o}$C.

### 2.6.2. Staking gel buffer

3.93 gm 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 double distilled water.

### 2.6.3. Running gel buffer

59.93gm 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 was made up to 500 ml using double distilled water.

### 2.6.4. Protein molecular weight markers

Low molecular weight 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.

### 2.6.5. Protein sample loading buffer

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 $\beta$-mercaptoethanol, 0.001 gm of bromophenol blue were taken into a 10 ml reagent bottle, mixed well before volume of the content was made up to 10 ml with double distilled water and solution was stored at 4°C.

### 2.6.6. Tank buffer for SDS-PAGE (pH 8.5)

6.005 gm of Tris, 28.8 gm of glycine and 1 gm of SDS was dissolved in 500 ml of distilled water and pH of the solution was adjusted to 8.5 with 2N HCl. Finally the volume of the buffer was made up to 1000 ml with double distilled water.

### 2.6.7. Staining solution

0.2 gm of coomassie brilliant blue was dissolved in 50 ml of methanol and mixed with 10 ml of acetic acid, and finally the volume was made up to 100 ml using double distilled water. The staining solution made in this manner was stored at room temperature until further use.

### 2.6.8. Destaining solution

Destaining solution was prepared by adding 30 ml of methanol and 10 ml of glacial acetic acid to 100 ml distilled water. Required amount of destaining solution was freshly prepared and used for destaining protein gels.

## 2.7. MARKER ENZYME ASSAYS: PREPARATION OF REAGENTS

### 2.7.1. Reagents for Acid phosphatase assay

#### 2.7.1.1. p-nitrophenylphosphate (25mM)

*p*-nitrophenylphosphate (PNPP) solution was prepared by dissolving 0.092 gm of PNPP in 10 ml deionized water and stored at 2-8$^{o}$C until further use.

#### 2.7.1.2. Glycine hydrochloride (250mM) (pH 2.5)

2.78 gm of Glycine hydrochloride was dissolved in 70 ml of deionized water and the pH of the solution was adjusted to 2.5 with concentrated HCl and the final volume was made up to 100 ml with deionized water. The solution was kept at room temperature for further use.

### 2.7.2. Reagents for Glucose-6-phosphate dehydrogenase assay

#### 2.7.2.1. Glycylglycine buffer (250mM, pH 7.4)

Glycylglycine buffer was prepared by dissolving 3.03 gm of Glycylglycine free base in 70ml of deionized water and the pH of the solution was adjusted to 7.4 with 1M NaOH and finally the volume was made up to 100 ml with deionized water. The solution was stored at 25$^{o}$C until further use.

#### 2.7.2.2. D-Glucose-6-Phosphate Solution (60mM)

D-Glucose-6-Phosphate solution was prepared by dissolving 0.04gm of D-Glucose-6-Phosphate in 2ml of distilled water and stored at -20$^{o}$C until further use.

#### 2.7.2.3. NADP solution (20mM)

NADP solution was prepared by dissolving 297 mg of Nicotinamide Adenine Dinucleotide Phosphate in 20ml of deionized water and the solution was kept at -20$^{o}$C until further use.

#### 2.7.2.4. Magnesium Chloride (300mM)

0.121gm of Magnesium Chloride was dissolved in 2 ml of deionized water and stored at 4$^{o}$C until further use.

### 2.7.3. Reagents for Nitrate reductase

#### 2.7.3.1. Methyl viologen

Methyl viologen solution was prepared by dissolving 50 mg of sodium hydrosulfite in 10 ml of 0.01 M NaOH and 4% sulfanilamide containing 25% HCl. The contents were mixed

with 0.08% N-1-napthhylethylene-diamine dihydrochloride. The solution was kept at 4$^{o}$C until further use.

#### 2.7.3.2. 0.1M potassium nitrate

10.11gm of potassium nitrate was dissolved in few ml of double distilled water and the volume of the contents were made up to 1000 ml with double distilled water to get 0.1 M of potassium nitrate.

#### 2.7.3.3. 0.1M phosphate buffer (pH 7.2)

17.4gm of $K_2HPO_4$ and 13.6gm of $KH_2PO_4$ were dissolved in 700 ml of double distilled water and the pH of the buffer was adjusted to 7.2 with acetic acid, and final volume was made up to 1000 ml with double distilled water.

## 2.8. PREPARATION OF BUFFERS FOR WESTERN BLOTTING

### 2.8.1. Protein transfer buffer

3.03 gm of Trizma-base and 14.4 gm 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 4$^{o}$C until further use.

### 2.8.2. TBS-T Buffer

2.42 gm of Tris and 8 gm 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 Millipore water.

### 2.8.3. Blocking Buffer

Blocking buffer was prepared by adding 10 gm 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.

## 2.9. DNA MANIPULATIONS: PREPARATION OF SOLUTIONS AND BUFFERS

### 2.9.1. Tris Borate EDTA (TBE) buffer

A stock solution of 10 X TBE buffer was prepared by dissolving 108 gm of Tris, 55 gm boric acid, 40 ml of 0.5 M EDTA in 900 ml of distilled water and final volume was made up to 1000 ml with double distilled water. When necessesory appropriate volume of stock (10X TBE) was taken and diluted to get 1X TBE.

### 2.9.2. TE buffer

TE buffer was prepared by dissolving 121 mg of Tris and 37.2 mg of EDTA in 80 ml of distilled water. The pH of the buffer was adjusted to 8.0 and finally the volume was made up to 100 ml with double distilled water.

### 2.9.3. Sample Loading buffer (6X) for Agarose gel electrophoresis

5 mg of bromophenol blue, 25 mg of xylene cyanol and 4 gm of sucrose were dissolved in few ml of distilled water and finally the volume was made up to 10 ml with distilled water. The contents were stored at 4°C until further use.

### 2.9.4. 10% SDS

10gm of sodium dodecyl sulfate was dissolved in few ml of distilled water and the final volume was made up to 100 ml of double distilled water.

**2.9.5. 20XSSC**

Sodium citrate and sodium chloride buffer was prepared by dissolving 88.2gm of sodium citrate (0.3M) and 175.3gm of sodium chloride (3M) in 500 ml of distilled water and the solution was made up to 1000 ml with double distilled water.

**2.9.6. Alkaline transfer buffer**

Alkaline transfer buffer (0.4N NaOH and 0.1M NaCl) was prepared by dissolving 16gm of sodium hydroxide pellets and 58.43gm of sodium chloride in few ml of double distilled water and the solution was made up to 1000ml.

**2.9.7. Denaturation solution**

Denaturation solution (1.5M NaCl and 0.5M NaOH) was prepared by dissolving 87.65gm of sodium chloride and 20gm of sodium hydroxide pellets in few ml of distilled water and finally the solution was made up to 1000 ml with double distilled water.

**2.9.8. Neutralization buffer**

121.14gm of Tris and 87.6gm of sodium chloride were dissolved in 500 ml of distilled water and the pH of the solution was adjusted to 7.4 with 1N HCl. The final volume of the buffer was made up to 1000 ml with double distilled water to get neutralization buffer containing 1M Tris (pH 7.4) and 1.5M NaCl.

**2.9.9. 10X MOPS Electrophoresis buffer**

41.8 gm of MOPS was dissolved in 700 ml of double distilled water containing 1% DEPC and pH of the solution was adjusted to 7.0 using 2N NaOH. To this 20 ml of 1M sodium acetate and 20 ml of 0.5 M EDTA (pH 8.0) was added before adjusting the final volume to 1000 ml with double distilled water. The buffer was filter sterilized using 0.45 µM Millipore filter and stored in dark room at room temperature.

### 2.9.10. 2X RNA loading buffer

47.5 ml of 95% formamide, 12.25 mg SDS (0.025%), 12.25 mg bromophenol blue (0.025%), 50µl of 0.5 M EDTA, 12.25 mg xylene cyanol (0.025%) and 12.25% ethidium bromide (0.025%) was dissolved in 30 ml of distilled water and the final volume was made up to 50 ml. The solution was stored at 4$^{o}$C until further use.

## 2.10. PROTEIN ANALYSIS

### 2.10.1. Determination of protein concentration

Protein concentration in subcellular fractions was estimated by following the procedures described by Lowry et al., 1951. In a clean test tube, 100 µl of protein sample was taken and mixed with 900µl of water and 5 ml of alkaline solution (2% $Na_2CO_3$ and 0.5 g copper sulphate in 50:1 ratio). Then the contents were thoroughly mixed before allowing to stand at room temperature for 10 min. Immediately after incubation 0.5 ml of Folin-Ciocalteau reagent (diluted with distilled water in 1:1 ratio before use) was added and thoroughly mixed. The reaction contents were incubated at room temperature for 30 min. After incubation blue color developed was measured at 750 nm. A solution containing 5 ml of alkaline solution and 0.5 ml of Folin-Ciocalteau reagent served as blank. Known concentrations of bovine serum albumin were used for preparation of protein standards.

### 2.10.2. SDS-PAGE Gel Electrophoresis

The gel cast was assembled with 1.0 mm spacers and tested with dd$H_2O$ for leaks. Once the cast was assembled and tested for leaks, 12.5 % gels were prepared for separation. These were prepared with 25 % (v/v) separating buffer (0.55 M Tris base, 0.4 % (w/v) SDS pH

adjusted to 8.8 with HCl) 31.25 % (v/v) acrylamide / bisacrylamide solution, 42.6 % (v/v) $ddH_2O$, 0.1 % (v/v) TEMED with 1 % (w/v in $ddH_2O$). Ammonium persulphate (APS) was added to start the crosslinking reaction (1 % v/v). This solution was immediately transferred into the cast with a plastic Pasteur pipette and overlaid with water saturated butanol, to stop interfering of air with the crosslinking reaction. Once polymerization process was complete the butanol was removed using blotting paper and the stacking gel (5 %) prepared by mixing 25 % (v/v) stacking buffer (1.64 M Tris base, 0.4 % (w/v) SDS pH adjusted to 6.8 with HCl), 12.5 % (v/v) acrylamide / *bis*acrylamide, 61.5 % (v/v) $ddH_2O$, 0.1 % (v/v) TEMED and 1 % APS (w/v in $ddH_2O$) was immediately overlaid on the separating gel. Then the gel comb was carefully inserted without trapping bubbles in the wells. The entire assembly was kept for 15 min (confirmed by the remaining gel setting in the tube) to facilitate polymerization of stacking gel. The gel assembly was then taken out of the casting apparatus and put into the gel tank filled with running buffer (0.025 M Tris base, 0.192 M Glycine, 0.1 % SDS). The comb was carefully removed before loading appropriate amount of protein samples. Electrophoresis was performed at (200) constant voltage for 45 min. After completion of electrophoresis the gel was stained for 1hr by submerging the gel in staining solution. The stained gels were then destained till the protein bands were seen on the gel. A parallel gel run under identical conditions was used for western blotting

**2.10.3. Western blotting**

Western blotting was performed using ECL+Plus western blotting detection system supplied by Amersham Pharmacia Biotech, UK. After performing SDS-PAGE, the unwanted and unused areas of the gel was removed to keep the size of the gel as minimum as possible.

Left-hand corner on the top of the gel was cut to mark the orientation of the gel. Four pieces of Whatman filter paper soaked in towbin buffer (0.3M Tris-Cl pH 10.4; 20% methanol) were placed on anode graphite plate. On top of these filter papers, 2 pieces of Whatman filter paper soaked in towbin buffer were placed. The polyvinylidene fluoride (PVDF) membrane (Highbond-P) was kept in methanol for 5 sec to prewet and immediately rinsed with towbin buffer and placed on to the anode plate. Then the SDS-PAGE gel submerged in towbin buffer was placed on top of the membrane. On top of the gel, 2 pieces of filter papers soaked in towbin buffer were placed. Transfer of proteins onto PVDF membrane was carried out at (40 volts) constant voltage at 4°C. The process of protein transfer was continued for 12 hrs.

Once the transfer process was finished, membrane was stained with Panceau reagent to check the rate of proteins transfer on to PVDF membrane followed by removing the Panceau reagent by rinsing 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 6xHis peptide was 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 excess / unbound primary antibody. After primary antibody incubation, the membrane was incubated with secondary antibody (anti-mouse IgG supplied by ECL+Plus kit, Amersham Pharmacia Biotech, UK) in a ratio of 1:5000 in 1X TBS-T buffer containing 10% blocking reagent for 45 min at room temperature with constant

shaking. Then the membrane was washed three times with 1X TBS-T as mentioned above and protein signals were detected by following the manufacturers instructions.

### 2.10.4. Detection

Detection was performed using ECL+Plus kit supplied by 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 depending the signal intensity. After incubation membrane was dried on tissue papers to remove excess detection reagent and then wrapped 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.

## 2.11. DNA MANIPULATIONS

### 2.11.1. Isolation of plasmids

Plasmid vectors were purified using QIAgen mini preparation kit especially when used for cloning. A single bacterial colony carrying plasmid was inoculated into a 5 ml of LB medium containing appropriate antibiotic and was incubated overnight at 37°C with vigorous shaking (~150 rpm). The overnight culture was centrifuged at 8000 rpm for 10 minutes and supernatant was discarded. The bacterial cell pellet was resuspended in 250 µl of buffer TE containing RNase (2µg/ml). The cell suspension was then mixed with 250 µl of lytic mixture (0.2N NaOH and 1% SDS) and inverted 4-6 times to mix the contents thoroughly. After lysis of the cells, 350 µl of neutralization buffer (3M Sodium acetate, pH 4.8) was added and the contents were mixed by vortexing the tubes for 4-6 times. Then contents were centrifuged at 13000 rpm for 10

minutes to pellet down the cell debris. After centrifugation the clear supernatant was directly transferred to a QIA preparation column placed in a collecting tube and centrifuged for 1 minute at 13000 rpm. The column was then washed with 0.75 ml of wash buffer PE (70% Alchahol). To remove the residual wash buffer, column fitted with the collection tube was centrifuged at 13000 rpm for 1 minute. Finally plasmid DNA was eluted from the column by adding 50 µl elution buffer (10 mM Tris-HCl, pH 8.5) or Millipore water to the center of QIA preparation column followed by brief centrifugation at 13000 rpm. The plasmid DNA was stored at –20°C until further use.

**2.11.2. Isolation of plasmid DNA by Alkaline Lysis method**

Mini preparations of plasmid DNA was carried out by following the standard procedures (Sambrook et al., 1989) 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 (150 rpm). 1 ml of overnight culture was centrifuged at 6000 rpm for 10 minutes and supernatant was discarded. Bacterial cell pellet was resuspended in 100 µl of ice-cold solution I (50mM glucose, 25mM Tris-Cl, pH 8.0; 10mM EDTA, pH 8.0) by vigorous vortexing. To the above bacterial suspension a 200 µl of freshly prepared lysis solution (0.2N NaOH, 1% SDS) was added and the contents were mixed by inverting the tube 5-6 times. Then 150 µl of ice-cold neutralization buffer (3M Sodium acetate, pH 4.8) was added to the above bacterial lysate and mixed by inverting the tube for 4-5 times. Then tube was kept on ice for 3-5 minutes. A white precipitate formed in the above mixture was removed by spinning the contents at 13000 rpm for 10 minutes. Then the supernatant was transferred into a fresh tube and extracted with equal volumes of phenol:chloroform. The

contents were centrifuged at 13000 rpm for 10 minutes to separate aqueous and organic layers. The volume of the aqueous phase was measured before transferring to a fresh tube. A 1/10$^{th}$ volume of 3 M sodium acetate and 2 volumes of ice-cold absolute alcohol were added to the aqueous phase and the tubes were kept at –20°C for 30 minutes. The plasmid DNA was then precipitated by centrifuging at 13000 rpm for 10 minutes at 4$^{o}$C. The DNA pellet was further washed with 70% ethanol to remove traces of salts associated with plasmid. Subsequently the plasmid DNA was dried before redissolving it in 50 µl of TE (pH 8.0).

**2.11.3. Purification of DNA fragments from agarose gel**

The DNA fragments were extracted from agarose gel by using QIAgen gel extraction kit. After performing agarose gel electrophoresis, appropriate portion of the gel containing desired DNA fragment was sliced, weighed and carefully taken into a microcentrifuge tube. Then 3 volumes (W/V) of buffer QG was added and incubated at 50°C for 10 min to dissolve the gel. After the gel slice is completely dissolved 1 volume (w/v) of isopropanol was added. Meanwhile the QIA quick column was placed in a collection tube and the contents were carefully transferred into the column. Immediately the column fitted in collection tube was placed in a microfuge and the contents were centrifuged at 13000 rpm for 1 minute. The flow-through was discarded and the column was again washed with 0.75 ml of buffer PE (70% alcohol). Finally the column was subjected to a brief spin to remove residual amount of ethanol found in buffer PE. The QIA quick column was then placed in a 1.5 ml eppendorf tube and the DNA was eluted in 50 µl of buffer EB (10mM Tris-Cl, pH 8.5) or Millipore water.

### 2.11.4. Isolation of genomic DNA

Genomic DNA from the bacterial cells was isolated by using QIAGEN DNeasy tissue kit. Cells were collected from 5 ml of the overnight culture by centrifuging for 10 minutes at 7500 rpm. Cell pellet was then resuspended in 180 µl of ATL buffer (tissue lysis buffer). Further the cells were mixed with 20 µl of Proteinase K (100mg/ml) before incubating the contents at 55$^{o}$C for 1 hour. The incubation was continued till a clear cell suspension is seen in the tubes ensuring the complete lysis. After complete lysis 4 µl of RNase (100µg/ml) was added and incubated at room temperature for 2 minutes. After RNase treatment, a 200 µl of buffer AL (lysis buffer) was added and incubated at 70$^{o}$C for a further period of 10 minutes. After incubation the contents were brought to the room temperature and 200 µl of 95% ethanol was added before mixing the contents thoroughly. The contents were loaded on to a DNeasy mini column and centrifuged at 8000 rpm for 1 minute to facilitate passing of contents through matrix. The flow through was discarded and washed with 500 µl of buffer AW1 (wash buffer 1) followed by 500 µl of buffer AW2 (wash buffer 2). After washing the DNeasy mini column was kept in a clean 1.5 ml eppendorf tube and the genomic DNA was eluted with 200 µl of buffer AE (elution buffer).

### 2.11.5. PCR amplification

Amplification reactions were performed in 25 µl volume containing 1.5 mM $MgCl_2$, 200 µM each dATP, dCTP, dGTP and dTTP mix (Amersham), 10 pico mol of forward and reverse primers, 1.0 Unit of *Taq* polymerase, 10-20 ng of plasmid / genomic DNA used as a template. Amplifications were carried out in the thermal cycler by following three stage PCR programme. The stage I involves initial 5 cycles for 1 min at 94$^{o}$C, 1 min at 45$^{o}$C and 1 min at 72$^{o}$C. The

stage II is set for 30 cycles, each having 1 min at 94°C, 1 min at 55°C and 1 min at 72°C. A final extension step at 72°C for 10 min was included as final stage. Amplification products were analyzed on 0.8% agarose gel electrophoresis.

**2.11.6. Isolation of total RNA**

Single colonies of *E. coli* cells containing expression plasmids were independently inoculated in LB broth for overnight cultures at 37°C and 1% of this overnight culture was used as inoculum in 10 ml LB broth to prepare mid log phase cultures. Then the cells were harvested at 6000 rpm for 5 minutes at 4°C and supernatant was discarded. The total RNA was isolated from these cells by following Qiagen RNeasy mini prep manufacturer's protocol. The cell pellet was then suspended in TE buffer (pH 8.0) containing Lysozyme (400µg/ml) and incubated at room temperature for 20 minutes. After incubation 350 µl of buffer RLT (lysis buffer) was added to the above cell suspension and mixed the contents thoroughly by vortexing vigorously. The insoluble material was removed by centrifuging the suspension for 5 minutes at 13000 rpm at room temperature and clear supernatant was taken for further steps. To this supernatant about 250 µl of absolute alchahol was added and mixed thoroughly before applying it an RNeasy mini column and centrifuged at 13000 rpm for 1 min at room temperature. The flow through was discarded and RW1 (supplied by Qiagen) wash buffer was applied to the column and centrifuged at 13000 rpm for 1 minute at room temperature. About 500 µl of RPE (supplied by Qiagen) buffer was applied to column and centrifuged again at 13000 rpm for 1 min at room temperature. The column was washed once again with RPE buffer as mentioned above and the column was recentrifuged once again for a minute at maximum speed to remove the traces of RPE buffer.

Finally 50µl of elution buffer was directly applied to the column and total RNA was eluted by centrifuging at 13000 rpm for 1 minute at room temperature.

### 2.11.7. Real-time PCR

The mRNAs of the *opd* specific variants were quantified by performing Real Time PCR. The total RNA was isolated from *E. coli* cells containing expression plasmids pHYS400, pHNS400 and pHIR400 were used as templates, while 16S rRNA was used as an internal control. One µg of total RNA was reverse transcribed using random hexamers and MMLV-reverse transcriptase (Invitrogen) and the corresponding cDNAs were used as templates in a 25 µl PCR reaction using power SYBRGreen master mix (Applied Biosystems). The PCR amplifications and fluorescence detection were performed using real-time PCR machine (ABI 7500, Applied Biosystems) by following the manufacturer's universal thermal cycling conditions. The *opd, opd'* variants and 16S rDNA were amplified in a separate reaction using the same pool of cDNA template. The transcript abundance of *opd and* its variants were normalized to that of 16S rDNA. The *opd* and *opd'* variants mRNA relative abundance was calculated using the following formulae.

Relative expression of gene = 2 x^y (-ΔCt value of target gene) – (-ΔCt value of internal control)

$$= \text{'X' of } 2^{-\Delta Ct}$$

### 2.11.8. Purification of PCR products

After performing the polymerase chain reaction (PCR), the total PCR reaction mixture (25 µl) was taken into an eppendorf tube and the PCR product was purified following the manufactures protocol using QIAgen PCR purification kit. The 25 µl PCR reaction mixture was

mixed with 125 µl of buffer PB (binding buffer) and passed through Qia quick spin column fitted in a collecting tube. The flow through was discarded and the QIA quick spin column was washed with 0.75 ml of PE (wash buffer). The column was then placed in a new eppendorf tube and subjected for brief centrifugation to remove the traces of ethanol. The column was then placed in a 1.5 ml eppendorf tube and the DNA was eluted by adding 50 µl of buffer EB (10 mM Tris-Cl, pH 8.5).

**2.11.9. Agarose Gel Electrophoresis**

Agarose gel electrophoresis was performed by the following standard procedures (Sambrook et al., 1989). Agarose (0.8 gm) was taken in 100 ml of 1X TBE buffer and the contents were boiled till the agarose was completely dissolved. The solution was then cooled to 50-55°C and poured on a clean sterile gel tray fitted with combs and 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 samples were mixed with 4µl 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 in staining solution (0.5 µg/ml ethidium bromide in water) for 15 minutes and was destained in distilled water for 15 minutes. Then the gel was visualized under UV transilluminator and the electrophoretic mobility of DNA was recorded by taking the image of the gel using gel documentation system.

### 2.11.10. Cloning of PCR amplicon into pGEM-T easy vector

In order to clone the PCR product into pGEM-T easy vector, 100 ng of PCR product was taken in a clean eppendorf tube and mixed with 50 ng of pGEM-T easy vector (Purchased from Promega, USA). Ligation was carried out by adding appropriate amounts of 2X T4 ligase buffer and 0.1 U T4 DNA ligase. The ligation was performed for 16 hours at 16$^{o}$C. After completion of ligation, the ligation mixture was transformed into *E. coli* DH5α competent cells as described elsewhere and the transformants were selected by plating the cells on LB plate containing 100 µg/ ml of ampicillin, IPTG (1 mM), X-gal (2%). The colonies containing recombinant plasmids were selected based on blue white discrimination.

### 2.12. Preparation of competent cells

Competent cells were prepared following the procedures of Mandel and Higa with slight modifications (Mandel and Higa, 1970). *E. coli* cells to be made competent were grown in 100 ml of LB broth at 37°C. The culture was grown 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 pellet was then suspended in 50 ml of ice cold 0.1M $CaCl_2$ and incubated on ice for 30 minutes. After incubation the cells were collected and gently resuspended in 10 ml of ice cold 100mM $CaCl_2$ containing 15% glycerol and, frozen in liquid nitrogen and stored in aliquots of 50 µl at -80°C until further use.

### 2.12.1. Transformation

The frozen competent cells were thawed by placing them on ice bath for few minutes. 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 seconds and immediately brought to 37°C incubator for 45 minutes after adding 1 ml of LB medium. The cells were collected from the tube by brief centrifugation and resuspended in 200 µl of LB broth before plating them on LB agar plates containing appropriate antibiotics. The plates were incubated at 37°C for 12 hr for colonies to appear.

### 2.13. Conjugation

All the *opd-lacZ* fusions and expression plasmids were mobilized into *Brevundimonas diminuta* by non-quantitative biparental plate mating method (Simon et al., 1983). All the *E coli* S17-1 strains carrying the *opd-lacZ* fusions and expression plasmids were taken as donors and *B. diminuta* as recipient strains. Donor and recipient strains were grown to mid log phase were taken in the ratio of 1:3 and plated on LB plate. The cells were allowed to grow at $30^0$C for 72hrs. After incubation the mating mixture was scraped into an eppendroff tube containing 1 ml of 0.9% steriled NaCl solution and suspended gently. The cell suspension was then serially diluted and the exconjugants were selected on LB agar plates containing polymixinB (10 µg/ml) and either tetracycline (10 µg/ml) or chloramphenicol (10 µg/ml) depending on the resistance marker available on plasmid. Selection on polymixinB eliminates *E. coli* donor strain and allows to grow only exconjugants of *B. diminuta* by eliminating host cells. Presence of expression plasmid was again confirmed by performing PCR using gene specific primers.

## 2.14. Generation of *opd* knockout in *Brevundimonas diminuta*

### 2.14.1. Construction of pSUP*opd::tet*

The plasmid pSM3 (Siddavattam et al., 2003) containing *opd::tet* cassette was digested with *Pst*I to release the 2.1 kb *opd::tet* fragment and the fragment was gel extracted using Qiagen gel extraction kit. The vector pSUP202 (Simon et al., 1983) containing tetracyline, chloramphenicol and ampicillin selection markers, and *oriT* (required for origin of transfer) and *mob* (involves in mobilization of plasmid) regions was used to construct pSUP*opd::tet* suicidal plasmid. The gel extracted *Pst*I fragment was then ligated into pSUP202 vector digested with similar enzyme. The insert and vectors were taken in a 1:3 ratio and the ligation mixture was transformed into *E. coli* DH5α cells and colonies were selected on LB agar plates containing tetracycline and chloramphenicol. Further the recombinant plasmids were isolated and restriction digested with the *Pst*I to show the presence of 2.1 kb insert containing *opd::tet* cassette. The size of the insert was confirmed by running 0.8% agarose gel electrophoresis and thus generated construct was designated as pSUP*opd::tet*.

### 2.14.2. Mobilization of pSUP*opd::tet* into *B. diminuta*

The plasmid pSUP*opd::tet* was transformed into *E. coli S17-1* which has RP4 mediated conjugal system required for the mobilization and transfer of the plasmids from donor to the recipient strains (Simon et al., 1983). The transformation was performed by following the method described elsewhere in this section. The transformants were selected on LB agar plates containing streptomycine and tetracycline. *E. coli S17-1* is resistant to streptomycine and the presence of pSUP*opd::tet* plasmid confirms resistance to tetracyline. The *E. coli S17-1* (pSUP*opd::tet*) was then used as a donor to mobilize pSUP*opd::tet* into *B. diminuta* using it as recipient strain. The cultures were grown to log phase and conjugation was performed by the

method described elsewhere in methodology section. The exconjugants were selected on LB agar plates containing either on polymixinB, tetracycline or polymixinB, chloramphenicol. The colonies showed resistance to polymixinB, tetracycline and sensitivity to polymixinB, chloramphenicol were selected for further screening.

### 2.14.3. Screening of *opd* knockout in *B. diminuta*

#### 2.14.3.1. Colony PCR screening

The exconjugants showed resistance to polymixinB, tetracycline were selected and replica plated on LB agar plates containing polymixinB, chloramphenicol. The colonies showed sensitivity to polymixinB and chloramphenicol and resistance to polymixinB and tetracycline were selected for colony PCR screening using *opd* specific primers. The strains of wild type *B. diminuta* were used as a positive control. PCR reaction was performed by following three stage programme described elsewhere in methodology and the presence of 2.2 kb *opd::tet* 1.2 kb *opd* gene were confirmed by running 0.8% agarose gel electrophoresis. The purity of the exconjugants of *B. diminuta* cells was confirmed by determing the restriction profile of 16S rRNA gene amplified using universal 16S rRNA specific primers. Further the positive colonies were selected and inoculated in LB broth to isolate plasmid pCMS1. Presence of *tet* gene in *opd* sequence was confirmed by performing southern hybridization.

#### 2.14.3.2. Southern hybridization

The selected exconjugants of *B. diminuta* were inoculated in fresh LB broth and allowed them to grow for overnight at 30$^{o}$C. The indigenous plasmids pCMS*opd::tet* from *B. diminuta* (*opd::tet*) were isolated and digested with *Pst*I and *Bam*HI. The restriction fragments were then

separated on 0.8% agarose gel and transferred on to Hybond-$N^+$ nylon (Amersham) membrane by capillary transfer. Membrane was hybridized with radiolabeled probes prepared from *opd* and *tet* genes independently and signals were detected after high stringency washes.

***Preparation of DNA probes for hybridization***

About 25 to 50 ng of probe DNA amplified using either *opd* or *tet* gene specific primers were taken independently in a clean eppendorf tube and incubated for 5 minutes in a boiling water bath. To this 3µl of $^{32}P$ labelled dCTP was added. The labeling reaction was carried out using GENEI labeling kit and strictly following manufacturers protocols. The reaction mixture contained 2.5µl of 10X assay buffer, 1µl of random primer (100 ng/µl), 2.5µl of 20 mM DTT, 2µl of dATP, dGTP and dTTP mixture. The reaction mixture was added with 3µl of double distilled water and 1µl of Kleno fragment (3 units/µl). Finally the reaction was incubated at room temperature for 2 hours. Then the $^{32}P$ labeled double stranded DNA was denatured by incubating at $100^0C$ for 5 minutes and subsequently used it for hybridization process.

***Hybridization***

The DNA in the gel was depurinated for 10 min at room temperature with 250mM HCl. The gel was rinsed with double distilled water and denatured at room temperature with gentle agitation by denaturation solution (1.5M NaCl, 0.5M NaOH). After denaturation the gel was neutralized with 0.5M NaCl, 0.5M Tris-HCl (pH 7.5) for 30 min. The capillary blotting apparatus was assembled using 10X SSC as the transfer buffer and the DNA was transferred overnight onto Hybond N+. After transfer the membrane was UV cross-linked and treated with Prehybridization buffer (0.5M phosphate buffer pH 7.2, 7% SDS and 1mM EDTA pH 8.0) for 4

hours at 65$^{o}$C. The *opd* and *tet* fragments were α-labelled with radioactive dCTP using the random priming labeling kit (Genei Pvt ltd) by following manufacturer protocol. The probe was denatured and added to the hybridization buffer before transferring into the plastic bag containing equilibrated nitrocellulose filter. The hybridization reaction was performed for 18 hours at 65$^{o}$C. Following hybridization the filter was washed by following high stringent conditions.

***Washing and detection***

The membrane filters were removed from hybridization solution and immediately immersed the filters in a large volume of (300 to 500 ml) wash solution 1 (2X SSC + 0.1% SDS) at room temperature for 5 minutes. The washing step was continued for three times with gentle agitation. Then the membrane filters were transferred to 300 to 500 ml of wash solution 2 (1X SSC + 0.1% SDS) and continued the washing for 1.30 hrs at 68$^{0}$C. After second wash the filters were dried at room temperature using 3MM Whatman filter papers and wrapped them in a clean saran wrap paper. The filters were then analyzed by exposing them to X-ray film for 12 to 16 hrs at -70$^{0}$C with an intensifying screen to obtain autoradiographic image. After exposure, the autoradiogram was developed by following the standard procedure. The film was immersed 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.

## 2.15. CONSTRUCTION OF EXPRESSION PLASMIDS

### 2.15.1. Construction of expression plasmid with and without signal coding region of *opd* gene

Expression vectors pET23b (Novagen) and pMMB206 (Morales et al., 1991) were used to express cloned genes in *E. coli* and *B. diminuta* respectively. While cloning in pET23b an artificial *Nde*I site was created through PCR mutagenesis by using a forward primer with *Nde*I site at an appropriate portion of the gene. The genes that were expressed in *E. coli* were then amplified along with the vector ribosomal binding site and terminator sequence as *Eco*RI/*Bam*HI fragment and cloned in pMMB206 under the control of *ptac* promoter after digesting both vector and insert with *Eco*RI/*Bam*HI. The resultant expression plasmids were then introduced into *E. coli* or *B. diminuta* either by transformation or by conjugation protocol described in the elsewhere sections of this chapter.

***DNA ligation***

Before ligation reaction was set, the concentration of vector and insert was measured and were then taken in a sterile eppendorf tube in 1:3 (vector : insert) reaction. One unit of $T_4$ and 2µl of 10X ligation buffer was added to the reaction tube. Finally the reaction volume was adjusted to 20µl and ligation reaction was performed at $16^0$C for 18 hours. A portion of ligation mixture was taken to transform *E. coli* competent cells. The colonies obtained on selective plates were used to isolate recombinant plasmid. Presence of insert was established by restriction digestion and subsequence analysis of the generated fragment on 0.8% agarose gel.

### 2.15.2. Secondary structure predictions in *opd* mRNA

The *opd* mRNA expected to generate stem loop structure forming sequence was subjected to nucleotide BLAST and the highly similar sequences were selected to predict the mRNA secondary structure. Among several *opd* sequences four were found to be highly conserved. These four sequences were subjected to bioinformatic analysis using mFOLD programme (http://www.bioinfo.rpi.edu/applications/mfold) to predict the mRNA secondary structures and to calculate ΔG values (at 37$^{o}$C and physiological pH). Further the secondary structure of mRNA sequence of *opd*' was also used to evaluate its potential to form secondary structure and ΔG value.

### 2.15.3. Site directed mutagenesis

A PCR site directed mutagenesis was performed in a 50µl reaction buffer containing 2.5 units of *pfu* DNA polymerase by following the programme consisting of stage I at 94$^{o}$C for 5min and stage II at 94$^{o}$C for 1min, 57$^{o}$C for 1 min and 72$^{o}$C for 3 min and continued for 28 cycles. The final stage was carried out for final extension at 72$^{o}$C for 10min. The generated PCR amplicon was then purified using Qiagen PCR purification kit and the amplicon was treated with *Dpn*I (37$^{o}$C for 1 hr) enzyme to eliminate the methylated parental plasmid DNA template (McClelland, 1992) and to rescue the unmethylated mutagenized plasmid DNA amplicon. After treating with *Dpn*I the amplicon was transformed into *E. coli* DH5α cells by performing transformation following the method described elsewhere. Once after transformation the nicks of the generated amplicon were sealed by the action of T4 DNA ligase of host *E. coli* DH5α cells.

Further the mutagenized plasmid was isolated and confirmed by restriction digestion and sequencing analysis.

### 2.15.4. OPH variants with changes in the Tat motif

Mutants of *opd* encoding OPH variants in which the invariant arginines of the Tat motif were replaced with amino acids having either similar charge or shape were generated by PCR mutagenesis. Plasmid pHYS400 was used as template, the forward primer contained the desired mutation, and the $T_7$ terminator primer was used as the reverse primer. PCR was performed using *Pfu* DNA polymerase, and the amplicons were digested with *Eco*RI and *Bam*HI and cloned in pMMB206. The resultant plasmids and the mutations they carry are listed in Table. 2. 3.

### 2.15.5. Construction of *opd-gfp* fusions

While constructing the *opd-gfp* fusions, *B. diminuta* expression plasmids pSM5 and pHLNS400 were used as source plasmids. These plasmids were digested with *Sal*I and *Hind*III to remove most of the 3' region of *opd* gene and used as vectors to generate *opd-gfp* fusions. The *gfp* gene was amplified as *Sal*I-*Hind*III fragment using pBQGFP as template. The PCR fragment digested with *Sal*I-*Hind*III were ligated to generate expression plasmids coding OPH-GFP fusion proteins in *B. diminuta.* When SsrA signal was added the reverse primer was designed by including codons specifying SsrA signal and the above described process was repeated to generate OPH-GFP fusion proteins with C-terminal SsrA signals.

### 2.15.6. Construction and expression of plasmid encoding preOPH -23K fusion protein

The signal peptide coding sequence of the *opd* gene was PCR amplified using forward and reverse primers having *Pst*I and *Eco*RI restriction sites respectively. The amplicon was then

cloned in pMW18 as *Pst*I and *Eco*RI fragment which facilitates in frame fusion of *opd* signal sequence with the spinach 23KDa oxygen evolving complex protein (23K) encoding cDNA. The generated plasmid was designated as pOPS400. The complete fragment coding preOPH-23K was taken as *Sac*I-*Sal*I fragment and cloned into an expression vector pDHB5700 (De Gier et al., 1998) to generate expression plasmid pTLS400 which codes for 23KD protein with OPH signal sequence. Further the generated expression plasmid pTLS400 was transformed in to *E. coli* MC4100 wild type and *tat* mutants and expression of preOPH-23K fusion protein was done by following standard protocols.

### 2.15.7. Construction of *opd-lacZ* fusions

Transcriptional *opd–lacZ* fusions were generated using the low copy number, mobilizable promoter probe vector pMP220 having promoter less *lacZ* as a reporter system (Spaink et al., 1987). The upstream region of *opd* gene was taken as various restriction fragments and directionally cloned in the multiple cloning site of the vector found upstream of reporter gene. In certain cases, the plasmids generated in our previous studies were used (Manavathi et al., 2005). The omega fragment containing transcriptional terminator sequence on either side of kanamycin gene was obtained from pUC4KIXX (Beck et al., 1982) as *Bam*HI fragment and cloned at *Bgl*II site found in between putative $\sigma^{70}$ and $\sigma^{54}$ promoters. Various *lacZ* fusions generated were then transformed into *E. coli* S17-1 and mobilized in *B. diminuta* cells as described elsewhere in this section.

### 2.15.7.1. β-galactosidase assay

*β*-galactosidase was assayed following the procedure described by Miller, 1972. *B. diminuta* cells containing *opd-lacZ* fusions were inoculated in 10 ml of LB medium and allowed to grow till the OD of the culture reached to 0.5 units at 600 nm. About 500 µl of culture was withdrawn from the 10 ml culture flask and cells were pelleted down by centrifuging at 8000 rpm for 10 minutes. Then supernatant was carefully discarded and bacterial pellet was dissolved in 0.5 ml of Z buffer. The cells were then lysed by adding 2 drops of chloroform and 1 drop 0.1% of SDS. After brief incubation at 30°C, the volume of the above reaction mixture was made up to 1 ml using Z-buffer. The contents were then mixed carefully before adding 200 µl of *O*-nitrophenyl-β-D-galactopyranoside (ONPG) solution (4 mg/ml of ONPG in 0.1 M $Na_2PO_4$ buffer, pH 7.0). After 30 minutes of incubation, 500 µl of 1 M $Na_2CO_3$ solution was added to the tubes to arrest the reaction. Tubes that contain all the above components except cells were taken as negative control. The *β*-galactosidase activity was measured by monitoring the release of *p*-nitrophenol from substrate ONPG. The *β*-galactosidase activity levels were calculated using the formula given below and the specific activity was represented as Miller units.

$$\beta\text{-galactosidase activity} = \frac{A_{420} \times 1000}{t \times V \times OD_{600}}$$

t = time elapsed between ONPG added and addition of $Na_2CO_3$ solution in minutes
V = volume of the culture added to the reaction mixture

$A_{600}$ = OD of the culture at 600 nm
$A_{420}$ = OD of the reaction mixture at 420 nm

## 2.16. METHODS FOR SUBCELLULAR LOCALIZTION STUDIES

### 2.16.1. Expression of OPH6His in *B. diminuta* (*opd::tet*)

Single colonies of *B. diminuta* (*opd::tet*) containing expression plasmids were inoculated in 2 ml of LB medium and grown for overnight with appropriate antibiotics. The overnight culture (1%) was used as inoculum to get 10ml (LB medium) of mid log phase culture. The culture was then allowed to grow until the cell density reached to 0.5 OD units at 600 nm. The log phase culture was then induced with 1mM IPTG and the culture was allowed to grow for a further period of 2 hrs to achieve maximum expression of OPH. The control cultures having expression vector pMMB206 was maintained under identical conditions and used them as negative controls. After induction, the cells were pelleted and the pellet was suspended in sample buffer (62 mM Tris-Cl (pH 6.8); 2% SDS; 10% glycerol; 5% β-mercaptoethanol; 0.01% bromophenol blue). Protein extracts were boiled in water bath for 5 minutes and samples were briefly centrifuged. Then the proteins obtained in clear supernatant were analyzed on 12.5% SDS-PAGE. Identical protocols were followed for induction of OPH and OPH-GFP fusion proteins in *E. coli.*

### 2.16.2. Subcellular fractionation

#### 2.16.2.1. Isolation of periplasmic fraction

The induced cells from 50 ml of culture were harvested by centrifugation at 10000 rpm for 10 min. The cell pellet obtained was then washed twice with saline and was finally resuspended in 1 ml of SET buffer (0.5 M sucrose, 5mM EDTA and 50mM Tris-HCl (pH 8.0) containing Lysozyme (600 µg/ml) and incubated at $30^0$C for 1 hr without agitation. After

incubation the suspension was centrifuged at 15000 rpm for 30 minutes at $4^0$C and the supernatant containing periplasmic fraction was collected and stored at -20$^o$C until further use (Weiner et al., 1998).

**2.16.2.2. Spheroplasts preparation**

A single colony of *B. diminuta / E. coli* containing appropriate expression plasmids were inoculated in 1 ml of LB medium having suitable antibiotic. The overnight culture (10µl) was then used to reinoculate in 10 ml of fresh LB broth having appropriate antibiotics. The culture was then grown at $30^0$C till the density of the culture reaches to 0.5 OD units. After reaching to log phase, the cells containing expression plasmids were induced for 2.30 hrs with 1mM IPTG at 30$^o$C. The induced cells were harvested and pellet obtained was thoroughly washed with saline. The cell pellet was finally resuspended in 1 ml of SET buffer (0.5 M sucrose, 5mM EDTA and 50mM Tris-HCl (pH 8.0) containing Lysozyme (600 µg/ml) and incubated at $30^0$C for 1 hr without agitation. After incubation the suspension was centrifuged at 15000 rpm for 30 minutes at $4^0$C and the supernatant containing periplasmic fraction was collected and stored at -20$^o$C until further use (Weiner et al., 1998). The pelleted spheroplasts were gently resuspended in Tris buffer (pH 8.0) containing 0.75M sucrose and subsequently used either for isolation of membrane / cytoplasmic fractions for treating with Proteinase K.

**2.16.2.3. Isolation of cytoplasmic and membrane fractions**

The spheroplasts generated were subjected to sonication at 32 amplititude with 15 sec pulse on and 45 sec pulse off for 5 cycles at 4$^o$C. The unbroken cells were removed by centrifuging at 10000 rpm for 10 minutes. The supernatant consisting the cell extract was further

centrifuged at 45,000 rpm for 1 hr. The clean supernatant containing cytoplasmic fraction was taken into a clean eppendorf tube and stored at -20$^{o}$C until further use. The pellet containing membrane was resuspended in 1 ml of phosphate buffer (0.1 M, pH 7.4) and recentrifuged at 45,000 rpm for 1 hour to eliminate contamination of cytoplasmic fraction. The pure membrane obtained was stored in -20$^{o}$C until further use. Protein concentrations of the subcellular fractions were determined before obtaining the purity of these fractions by assaying fraction specific marker enzymes.

### 2.16.3. Proteinase K treatment of Spheroplasts

The stock solution of Proteinase K (5 mg/ml) was carefully added to 5 ml of spheroplasts to get a final concentration of 0.5 mg/ml. The contents were gently mixed before incubating the tube on ice for 30 minutes. After incubation, 1 mM PSMF was added and the contents were incubated for a period of 20 minutes. The suspension was then centrifuged at 10000 rpm for 10 minutes at 4$^{o}$C, and the pellet containing Proteinase K treated spheroplasts was collected and resuspended in Tris buffer (pH 8.0) containing 0.75M sucrose. Finally the Proteinase K treated and untreated spheroplasts were used to assay parathion hydrolase activity. The spheroplasts having no Proteinase K served as negative control.

### 2.16.4. Assay of Marker Enzymes

#### 2.16.4.1. Acid phosphatase – Periplasmic marker

Acid phosphatase assay was performed following the modified procedure described elsewhere (Touati et al., 1986). The periplasmic, cytoplasmic and membrane fractions were incubated with 0.5 ml of 25 mM *p*-nitrophenylphosphate in 250 mM glycine hydrochloride (pH 2.5) for 20 minutes at 37$^{o}$C. The reaction was terminated by the addition of 1 ml of 1 N NaOH.

The formation of *p*-nitrophenol was determined at $A_{410}$ nm and the specific activity of acid phosphatase as determined by following formulae, was expressed as µmol of PNP formed per mg of protein per minute by applying molar extinction coefficient of PNP at $A_{410}$ nm (16,500).

$$\text{Acid phosphatase specific activity: } \frac{\text{Sample reading - negative control at } A_{410} \text{ nm X mg of protein}}{\text{16500(Molar extinction of PNP at 410nm)}}$$

= µmol of PNP formed/mg of protein/min

**2.16.4.2. Glucose-6-phosphate dehydrogenase (G6PD) – Cytoplasmic marker**

Glucose-6-phosphate dehydrogenase was used as a cytoplasmic marker enzyme. Subcellular fractions were assayed for G6PD activity by monitoring glucose-6-phosphate-dependent reduction of NADP at 340 nm. The assay was performed at 37°C in Tris buffer (50 mM Tris–HCl pH 7.5) containing 250 mM $NADP^+$. The reaction was initiated by addition of 12.5 mM glucose-6-phosphate (Sargent et al., 1998). The specific activity was calculated using the formulae given below.

$$\textbf{Specific activity } \frac{(\Delta\ A_{340}\text{nm/min Test} - \Delta\ A_{340}\text{nm/min Blank})\ (3)\ (df)}{(6.22)\ (0.1)}$$

Where,

3 = Total volume (in milliliters) of assay

df = Dilution factor

6.22 = Millimolar extinction coefficient of β-NADPH at 340 nm

0.1 = Volume (in milliliters) of enzyme used

**2.16.4.3. Nitrate reductase – Membrane marker**

The assay of nitrate reductase was done by following the method described by Michael and Showe, 1968. The assay was initiated by pre-incubating the subcellular fraction extracts for 5 min at 37$^{o}$C in a 2.4 ml volume of the reaction mixture containing 0.1 M phosphate buffer (pH 7.2), 0.1 M potassium nitrate, and $10^{-4}$ M methyl viologen. Before starting the reaction, 50 mg of sodium hydrosulfite was dissolved by mixing gently in a 10 ml of 0.01 M NaOH solution and the reaction was started by adding 0.1 ml of sodium hydroxide solution to the reaction mixture. After 10 minutes, the reaction was stopped by shaking the tube to oxidize the remaining hydrosulfite and reduced methyl viologen. To determine nitrite, 0.75 ml of sulfanilamide solution prepared by mixing two parts of 4% sulfanilamide in 25% HCl solution with one part of 0.08% N-1-napthhylethylene-diamine dihydrochloride was added to 2.5 ml reaction mixture. After 10 min, the absorbance of the reaction mixture was read at 540nm. One unit of enzyme is defined as the amount of enzyme required to convert 1 µmole of nitrate to nitrite in 10 min. Specific activity is defined as enzyme units per microgram of protein. The specific activity of nitrate reductase was calculated using the following formulae.

**Specific activity** =

$$= \frac{(\text{OD of the Test at } A_{540}\text{nm}) \times df \times 1000}{\text{mg of protein from subcellular fraction} \times \text{time in minutes (10 min)} \times \text{OD of the standard } (A_{540}\text{nm}/\mu\text{mol})}$$

### 2.16.5. Parathion Hydrolase Assay

A spectrophotometric assay described elsewhere was adapted to determine organophosphate hydrolase (OPH) activity. The reaction was carried out in the test tubes containing 100 µmol of methyl parathion in 1 ml of 50 mM Tris glycine buffer (pH 9.0). The reaction was started by adding one of the components such as spheroplasts, or crude extract, periplasmic, cytoplasmic and membrane fractions as source of OPH. The tubes were incubated at 37°C for 60 minutes. An increase in the absorbance at $A_{410}$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 of PNP (16500 $M^{-1}$ $Cm^{-1}$). The specific activity of the enzyme was expressed as micromoles of PNP produced per milligram of OPH protein per minute. The specific activity of the enzyme was calculated using the formula given below.

$$\text{Enzyme activity} = \frac{\text{Absorbance at 410nm}}{\text{16,500 (Extinction coefficient of PNP at 410nm)}}$$

**Specific activity** = Enzyme activity / mg of protein / min

### 2.17. Fluorescence Microscopy

A single colony of *B. diminuta* carrying pYSGFP400 was inoculated in 3 ml of LB broth with appropriate antibiotics. One percent of an overnight culture was then used to inoculate

10 ml of LB broth, grown to an OD600 of 0.5, and induced with 1mM IPTG at 30°C for 2hrs. Cells were harvested at 6000 rpm for 10 min at 4°C and washed three times in 10 percent sterile ice-cold glycerol. Cells were then incubated with FM4-64 (N-(3-triethylammoniumpropyl)-4-(pdiethylaminophenylhexatrienyl)-pyridinium 2Br), a membrane specific dye**,** for 2 min at room temperature and then fixed on a poly-L-lysine coated glass slide for confocal microscopic analysis. Confocal microscopy was performed using a Leica laser scanning confocal microscope (Leica Microsystems S.p.A.,Via Ettore Bugatti 12, Milano, I-20142, Italy) equipped with argon, krypton, helium and neon (ARKN) lasers. The fluorescence of GFP was induced with the 488 nm line of the argon laser and the red fluorescence with 543 nm line of the helium neon laser. Corresponding bright field images were collected through the differential interference contrast channel and detected using the transmission detector of the confocal microscope. Images were processed using the image software supplied by the microscope manufacturer. Contrast and brightness levels were adjusted with Adobe Photoshop 7.0.

**2.18. Fluorescence-activated cell sorting (FACS) analysis of *opd-gfp* expression**

Overnight cultures of *B. diminuta* carrying pYSGFP400, pYSGFPD400, pNSGFP400 or pNSGFPD400 were grown to mid log phase in LB in the presence of appropriate antibiotics and then induced with 1mM IPTG for 2 hrs. Cells were then harvested at 6000 rpm for 10 minutes at 4$^{o}$C. The cells were washed twice in phosphate buffer (pH 7.6) and suspended in 600 µl of the same buffer. The samples were analyzed in a BD Pharmingen FACSort. *B. diminuta* wild type cells served as negative control.

# CHAPTER - III

## To study regulation of *opd* gene expression

# CHAPTER - III

## To study regulation of *opd* gene expression

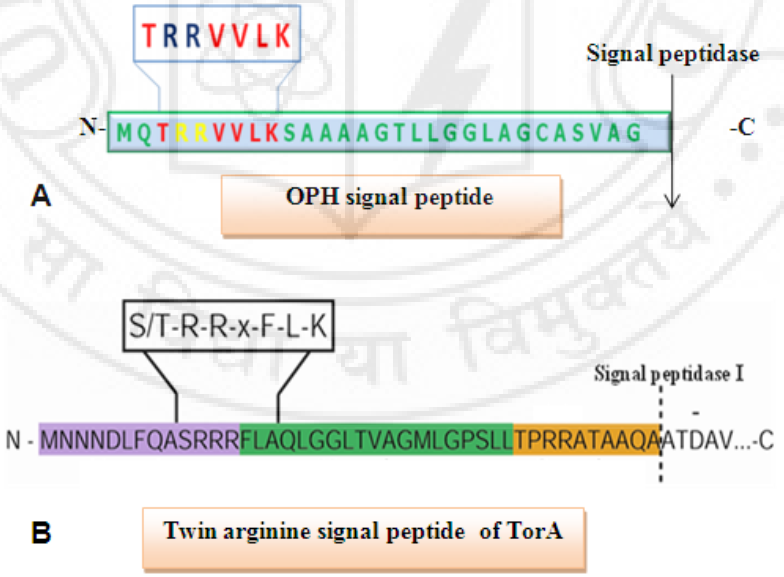
Organophosphate hydrolase (OPH) cleaves the characteristic triester bond found in variety of neurotoxic op-compounds including those present in nerve gas agents such as sarin and soman (Benning et al., 1994; Cho et al., 2004). Organophosphate hydrolases, coded by highly conserved plasmid borne organophosphate degrading (*opd*) genes (Serdar et al., 1982; Serdar and Gibson, 1985; Mulbry and Karns, 1989), have shown structural homology to the proteins of amidohydrolase superfamily. Especially, the tertiary folding of OPH contain a 8 barrel (α–β) structural fold that is typically seen in triose phosphate isomerase (TIM). The TIM barrel structure of OPH has high similarities with similar structural fold found in urease, adenosine deaminase, and dihydroorotase, considered to be the prominent members of amidohydrolase superfamily (Holm and Sander, 1997). Though the protein-scaffold and key active-site residues of all amidohydrolases are conserved, the substrate specificity and reaction mechanisms of different family members vary due to differences in the sequence and length of the β- or α-loops, particularly in loops one, seven and eight which contribute for substrate-binding (Raushel, 2000). Interestingly, in these proteins, the zinc ions can be successfully replaced with other divalent cations such as $Mn^{2+}$, $Co^{2+}$, $Cd^{2+}$, and $Ni^{2+}$ without compromising with catalytic activity (Omburo et al., 1992). Removal of the metal ions with chelators such as EDTA, and 1,10-phenanthroline produces an inactive apoenzyme suggesting that metal ions are required for catalytic activity (Omburo et al., 1992). The catalytic site of OPH possess two Zn ions, which can be successfully replaced with cobalt without compromising with enzyme activity (Sharon et al., 1991; Omburo et al., 1992). In view of such structural and catalytic relationship, the members of triesterases such as OPHs have been proposed to have evolved from amidohydrolase superfamily of proteins, through divergent evolution (Holm and Sander, 1997).

The OPH is a homodimer having monomer with a binuclear metal center weighing approximately 36 kDa (Donarski et al., 1989). Mulbry and Karns have successfully established the membrane association of OPH (Mulbry and Karns, 1989). In their pioneering studies they have proposed that the maximum OPH activity was associated with the membrane. They have used the OPH purified from membrane for determining N-terminal sequence and shown that Ser Ile Gly Thr Gly Asp Arg Ile Asn Thr Val Arg Gly Pro Ile Thr Ile Ser Glu Ala Gly Phe Thr Leu as first 29 amino acids (Mulbry and Karns, 1989). When this N-terminal sequence was compared with amino acid sequence deduced from organophosphate degrading (*opd*) gene, it perfectly matched with the sequence of OPH starting from 30$^{th}$ Serine residue (Mulbry and Karns, 1989). Based on this observation, followed by secondary structure predictions, they have reported existence of a 29 amino acid long signal peptide in OPH (Mulbry and Karns, 1989). Though this work was reported way back in 1989, no successful attempts were made to investigate the molecular mechanism pertaining to the membrane targeting of OPH. Instead, a number of attempts have been made to achieve heterologous expression of *opd* gene and to use recombinant OPH for successful elimination of op-compound residues from the environment (Brown, 1980; Cheng et al., 1983; Abd-Alla, 1994; DeFrank and Cheng, 1991). While achieving the heterologous expression of OPH, most of the investigators have constructed expression plasmids by deleting the 5' region of *opd* gene that specify signal peptide of OPH (Serdar et al., 1989; Mulbry and Karns, 1989; Pandey et al., 2009). The constructs that were generated by deleting the 5' region of *opd* gene have always resulted in producing the mature form of OPH (mOPH) in reasonably high quantity (Pandey et al., 2009). Interestingly, the OPH expressed without signal peptide folded properly and showed identical catalytic properties with the OPH purified from the wild type strains of *B. diminuta* or *Flavobacterium* sp. ATCC27551 (Mulbry and Karns, 1989;

Pandey et al., 2009). Contrary to these observation when full length open reading frame of *opd* gene is cloned to express precursor form of OPH (preOPH) most of the recombinant protein remained unprocessed even in the organisms that have close taxonomic relationship with *B. diminuta* (our unpublished results). Such observation is rather unusual and prompted us to take up further investigations on membrane targeting of OPH using *B. diminuta* as a model organism.

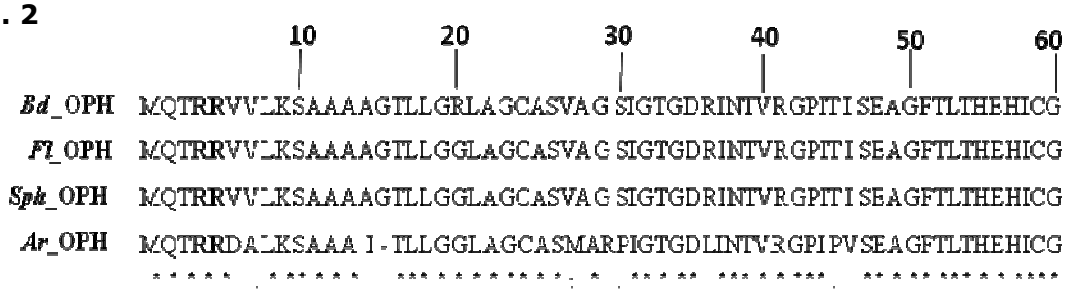
Before undertaking a detailed investigation into the membrane targeting of OPH, a thorough structural analysis was made by taking 29 amino acid long signal peptide of OPH. During the course of examination a structural motif (T-R-R-V-V-L-K) that show striking similarities to the recently discovered twin arginine transport motif (S/T-R-R-x-F-L-K) was identified (Berks, 1996; Weiner et al., 1998) (Fig. 1. 1). The twin arginine transport (Tat) motif is only seen in proteins that exclusively transport across/target membrane in a prefolded form (Weiner et al., 1998; Sargent et al., 2000).

**Fig. 1. 1**




Identification of Tat motif in the signal peptide of OPH. The Tat motif of OPH is shown in a box. The invariant twin arginines of Tat motif are shown with bold case (A). The Tat motif of OPH is compared with similar motif found in *E. coli* TorA protein, a typical Tat substrate (B).

**Fig. 1. 2**




Presence of Tat motif in OPH signal sequences. Alignment of the N-terminal amino acid sequences of organophosphate hydrolase (OPH) encoded by *opd* genes from *B. diminuta* (*Bd*), *Flavobacterium* sp. strain ATCC27551 (*Fl*), *Sphingomonas* sp. strain JK1 (*Sph*) and *Agrobacterium radiobacter* (*Ar*). The point of signal sequence cleavage, which has been determined experimentally in *B. diminuta* and *Flavobacterium*, is marked with a vertical line and the conserved twin arginine motif is shown in bold type. *, identity, :, similarity, ., weak similarity.

Encouraged by these results further study was undertaken to search if Tat motif is present in other OPH sequences isolated from taxonomically diversified group of soil microflora. When the signal peptide of *B. diminuta* was aligned, a well conserved Tat motif was also seen in the OPH sequences coded by *opd* genes of *Flavobacterium* sp. ATCC27551 (Sethunathan and Yoshida, 1973; Mulbry et al., 1986), *Agrobacterium radiobacter* (Horne et al., 2002) and *Sphingomonas* sp. JK1 (Genbank Access. No. ACD85809) (Fig. 1. 2). Such conservation signifies functional relevance of the Tat motif and prompted to hypothesize that OPH is a Tat (Twin arginine translocase) substrate. After formulating the hypothesis an attempt was made to elucidate the role of Tat motif in membrane targeting of OPH using *B. diminuta* as model organism.

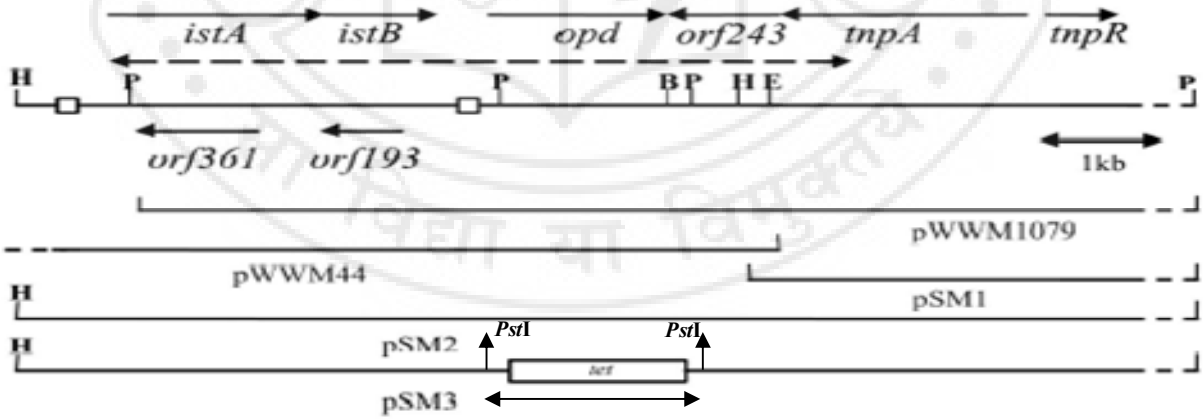
*B. diminuta* is isolated from the Texas agricultural soils, USA and no genetic tools are available to study membrane targeting of OPH in this organism. Therefore in the present study, as a initial requirement, we have developed broad host range, mobilizable expression plasmids to code OPH with C-terminal his tag. If such his-tagged OPH is produced in *B. diminuta,* it facilitates detection of precursor and mature forms of OPH through western blots by using anti

his antibodies. However, if such expression plasmids are mobilized into wild type *B. diminuta* cells they would be producing both normal (coded by the wild type *opd* gene) and his-tagged OPH. These two OPH variants are then expected to compete for membrane targeting causing difficulties in precise subcellular localization of the processed and mature forms of OPH. Therefore it is proposed to generate *opd* negative mutants of *B. diminuta,* before actually mobilizing the expression plasmid.

### 1.1. Generation of *opd* gene knockouts in *B. diminuta*

In our previous studies we have reported identification of transposon like *opd* gene cluster in *Flavobacterium* sp. ATCC27551 and generated plasmid pSM3 to study transposition of *opd* gene cluster (Siddavattam et al., 2003). As there are two *Pst*I sites flanking *opd* gene (Mulbry et al., 1987; Siddavattam et al., 2003) it was possible to excise the *opd::tet* from pSM3 as *Pst*I fragment (Fig. 1. 3).

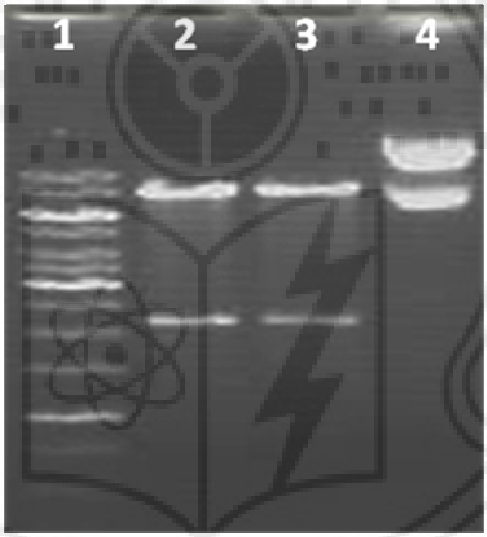
**Fig. 1. 3**




Organization of *opd* gene cluster: The *opd*::*tet* taken from pSM3 plasmid as *Pst*I fragment is shown with double headed arrow (Siddavattam et al., 2003).

Such fragment having *opd::tet* gene was then cloned in suicidal vector pSUP202 (Simon et al., 1983) digested with *Pst*I. The recombinant plasmid thus generated was designated as pSUP*opd::tet*. The restriction pattern of pSUP*opd::tet* is shown in Fig. 1. 4. The pSUP*opd::tet* was then transformed into *E. coli* S7-1 and used as donor in the conjugation studies intended to mobilize pSUP*opd::tet* into *B. diminuta.* The conjugation protocols described in materials and methods section were very successful and a number of exconjugants were present on LB plates containing tetracycline and polymixinB.

**Fig. 1. 4**




Agarose gel showing the presence of *opd::tet* in pSUP202. Lane 1 kb ladder, lanes 2 and 3 represent *Pst*I digests of pSUP*opd::tet*, lane 4 undigested pSUP*opd::tet*. Existence of 2.2 kb *opd::tet* is clearly seen along with vector backbone in lanes 2 & 3.

However no colonies were seen when parent strains were plated on selective plates. The exconjugants that have grown on polymixinB and tetracycline were then replica plated on polymixinB, chloramphenicol plates to examine the presence of vector backbone in *B. diminuta*. About ten percent of the colonies have shown sensitivity to chloramphenicol indicating the disappearance of vector in these exconjugants. Finally the exconjugants that showed resistant to polymixinB and tetracycline and sensitivity to chloramphenicol were selected for further genetic

analysis to confirm successful elimination of *opd* gene with *opd::tet* through homologous recombination. Two approaches were used for identification of successful *opd* knockout. One of them was PCR approach, in here we have used primers flanking *opd* gene to amplify *opd* sequences from the total genome isolated from the exconjugants and wild type strains of *B. diminuta*. In the PCR reaction, if wild type *opd* gene is replaced with *opd::tet,* a 2.2 kb amplicon is expected from the exconjugants as against the 1.1 kb amplicon generated from the wild type strains. The PCR screening indeed provided a quick reference to identify successful *opd* kncokouts. As shown in Fig. 1. 5 amplicon with a size of 2.2 kb and 1.1 kb were seen when PCR reaction mixture was analyzed on 0.8% agarose gel.

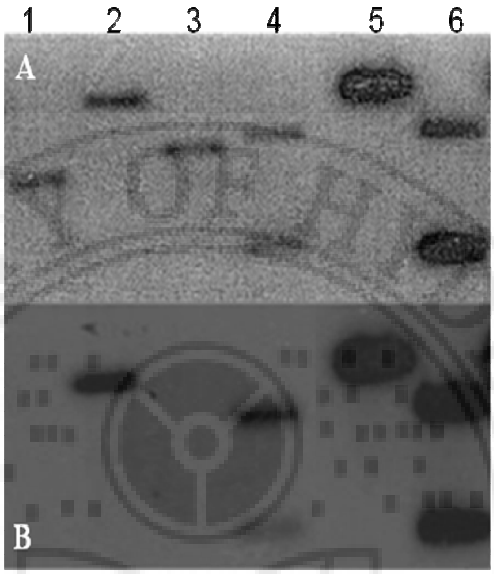
**Fig. 1. 5**


Agarose gel (0.8%) 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 show digestion of *opd::tet* with *Bam*HI, an unique site in tetracycline gene. Lane 4 represents amplification of 1.2 kb *opd* gene from wild type *B. diminuta* cells.

These positive colonies were then selected for performing a second confirmatory test involving southern hybridization. While performing hybridization experiment indigenous plasmid pCMS1 was isolated both from exconjugants and wild type strains of *B. diminuta.* The indigenous plasmid pCMS1 harboring *opd* gene, has a typical restriction pattern (Mulbry et al., 1987). If pCMS1 is digested with *Pst*I the entire *opd* gene will be seen in 1.1 kb *Pst*I fragment (Mulbry et al., 1987). If the wild type *opd* gene is replaced with *opd::tet* due to existence of tetracycline gene in the *opd* sequence its size will be elevated to 2.2 kb. Such difference in size can be seen when hybridization experiments are performed using internal *opd* sequence as a radio labeled probe. As expected, in southern hybridization experiment a 1.1 kb signal is clearly seen in the lanes loaded with *Pst*I digested pCMS1 isolated from wild type *B. diminuta* cells. No such signal was seen at the corresponding position in the lanes loaded with *Pst*I digests of pCMS1 isolated from exconjugants. Instead, a clear band was found at an elevated position corresponding to 2.2. kb region. Though this results confirm successful replacement of wild type *opd* gene with *opd::tet,* one more hybridization was performed using internal sequences of tetracycline gene as a radiolabeled probe. As expected no signal was seen in the lanes loaded with *Pst*I digests of wild type pCMS1. However, in the lane loaded with *Pst*I digests of pCMS1 isolated from the exconjugants the very same fragment that hybridized with *opd* sequence has given positive signal suggesting insertion of tetracycline gene in the coding sequence of *opd* gene (Fig. 1. 6). Based on these two approaches generation of *opd* knockout in *B. diminuta* is clearly established. The plasmid found in exconjugants where the *opd* gene is successfully replaced with *opd::tet* is designated as pCMS1::tet. The exconjugants having pCMS1::tet were then used to prepare cellular extracts for assaying OPH activity. In confirmation with genetic experiments, the exconjugants containing pCMS1::tet showed no OPH activity (Fig. 1. 7). After

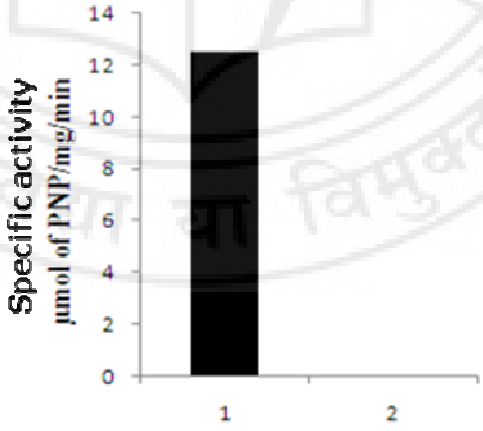
confirming generation of *opd* knockout using both genetic and biochemical tools, the *opd* negative mutants of *B. diminuta* were then used as recipients to mobilize expression plasmids coding OPH with C-terminal his tag.

**Fig. 1. 6**




Autoradiogram showing *opd* knockout in *B. diminuta.* Panel A & B showing autoradiograms developed using *opd* and tetracycline gene sequences as probes respectively. Lanes 1 and 2 represent *Pst*I digests of pCMS1 and pCMS1*::tet* plasmids, and lanes 3 and 4 represent *Bam*HI digests of similar plasmids; Plasmid pSM3 (Siddavattam et al., 2003) used as positive control was digested with *Pst*I and *Bam*HI and loaded in lanes 5 and 6 respectively.

**Fig. 1. 7**

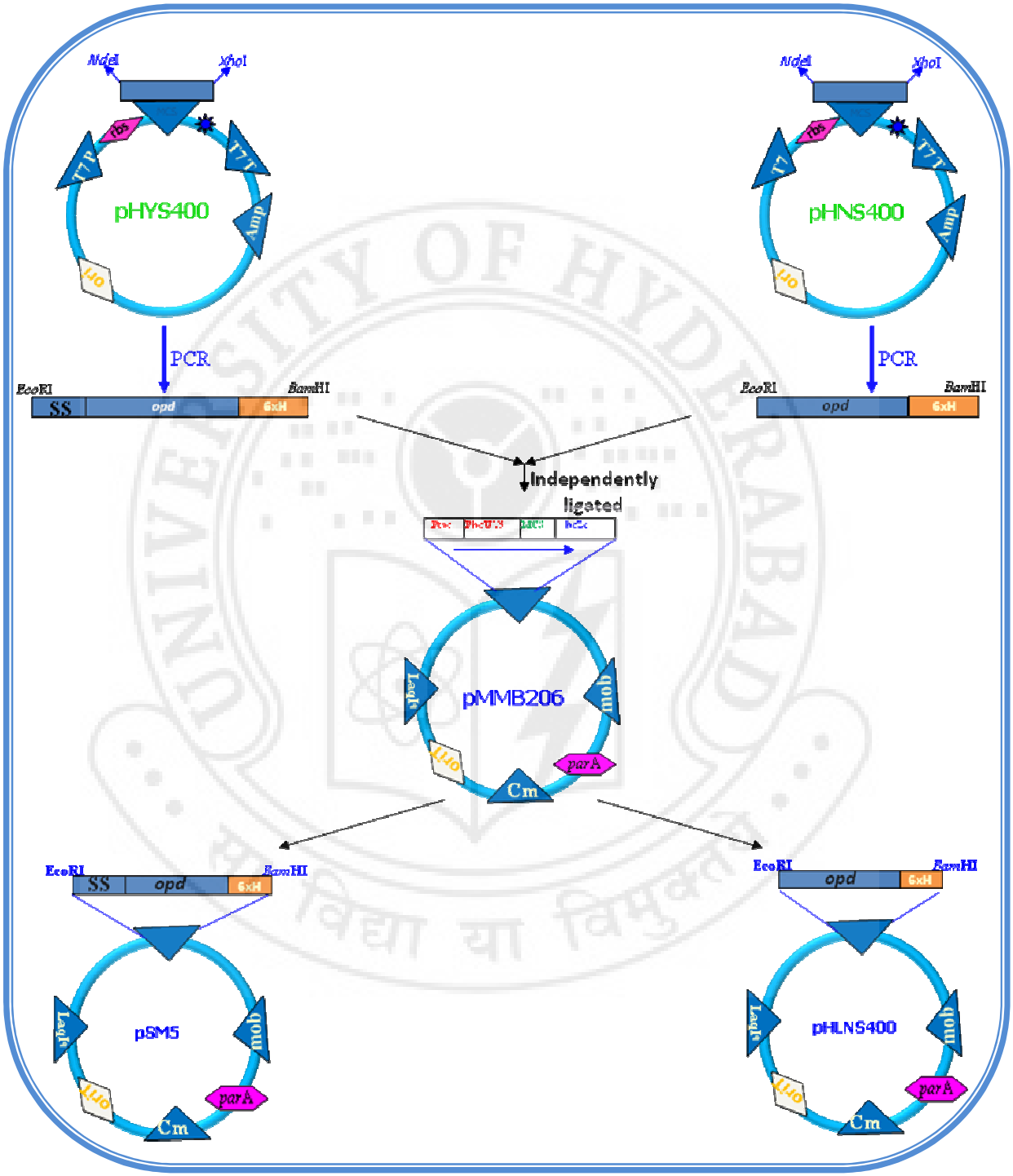



Specific activity of OPH 1. Wild type strains of *B. diminuta,* 2. *opd* negative mutants of *B. diminuta.*

### 1.2. Construction of expression plasmids

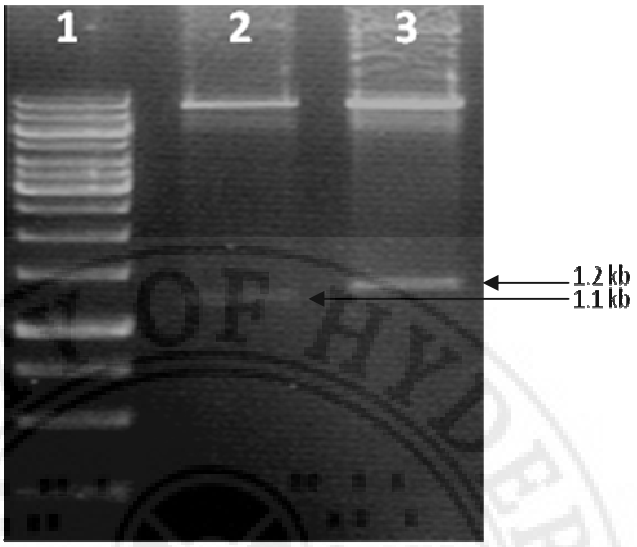
Generation of broad host range, mobilizable, low copy number expression plasmids coding OPH with C-terminal his tag were constructed by using vector pMMB206 (Morales et al., 1991). In our previous studies our laboratory has generated *E. coli* specific expression plasmids pHYS400 (Siddavattam et al., 2006) and pHNS400 (Pandey et al., 2009) to express precursor (preOPH) and mature forms (mOPH) of OPH with C-terminal his tag. In the present study, these two plasmids were used as templates to amplify corresponding *opd* genes coding for preOPH and mOPH. A forward and reverse primers were designed to amplify *opd* sequences from these two plasmids. A forward primer was designed taking vector specific sequences located upstream of ribosomal binding site with a recognition sequence for restriction endonuclease *Eco*RI (5'GGG AGA CCA CAA C**GAA TTC** CCT CTA GAA A3') and the reverse primer was similar to that of T7 terminator primer except that it has a recognition sequence for restriction endonuclease *Bam*HI (5'GCT CAG C**GG ATC C**AG CAG CCA ACT CAG C3'). The *opd* amplicons generated using pHYS400 and pHNS400 as templates were then digested with *Eco*RI and *Bam*HI and were independently cloned in plasmid pMMB206 to generate recombinant plasmids pSM5 and pHLNS400 respectively. Cloning of *opd* sequences in this manner places them directly under the control of p*tac lacUV5* promoters of expression vector pMMB206. As the *opd* amplicons used to generate pSM5 and pHLNS400 were obtained from pHYS400 and pHNS400 the precursor and mature form of OPH coded by these plasmids contain his tag epitope at C-terminus. Therefore by mobilizing plasmids pSM5 and pHLNS400 into *opd* negative mutants of *B. diminuta* expression of preOPH-6His and mOPH-6His can be induced by following standard protocols. A detailed strategy used for construction of expression plasmids and their restriction pattern was shown in Figs. 1. 8 and 1. 9.

**Fig. 1. 8**




Strategy used for construction of expression plasmids coding precursor (preOPH) and mature (mOPH) forms of OPH with C-terminal his-tags.

**Fig. 1. 9**

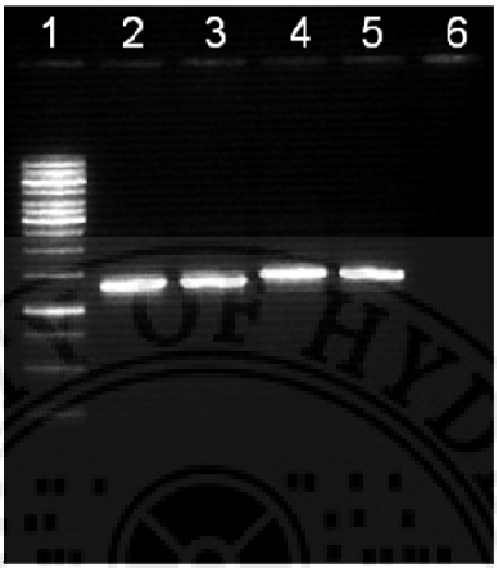



Agarose gel (0.8%) showing the release of (shown with arrow marks) 1.1 kb and 1.2 kb fragments containing *opd* sequences coding for preOPH (lane 3) and mOPH (lane 2) from vector pMMB206. Lane 1 kb ladder; lanes 2 and 3 represent expression plasmids pHLNS400 and pSM5 digested with *Eco*RI and *Bam*HI respectively.

### 1.3. Mobilization of expression plasmids into *B. diminuta*

The expression plasmids were mobilized into *B. diminuta* by following conjugation protocols described in materials and methods section. *E. coli* S17-1 cells containing either pSM5 or pHLNS400 were used as donors while the *opd* knockouts of *B. diminuta* served as recipient. The exconjugants were selected on polymixinB, chloramphenicol plates and the existence of the expression plasmids in exconjugants was determined by performing colony PCR using universal forward and reverse primers. Usage these primers facilitate amplification of insert cloned at multiple cloning site of expression vector pMMB206. The amplicons obtained from the exconjugants confirm existence of corresponding expression plasmids in exconjugants (Fig. 1. 10). No such amplification was seen in *opd* mutants of *B. diminuta.* This results confirm presence and stable maintenance of expression plasmids in *opd* knockouts of *B. diminuta.*

**Fig. 1. 10**



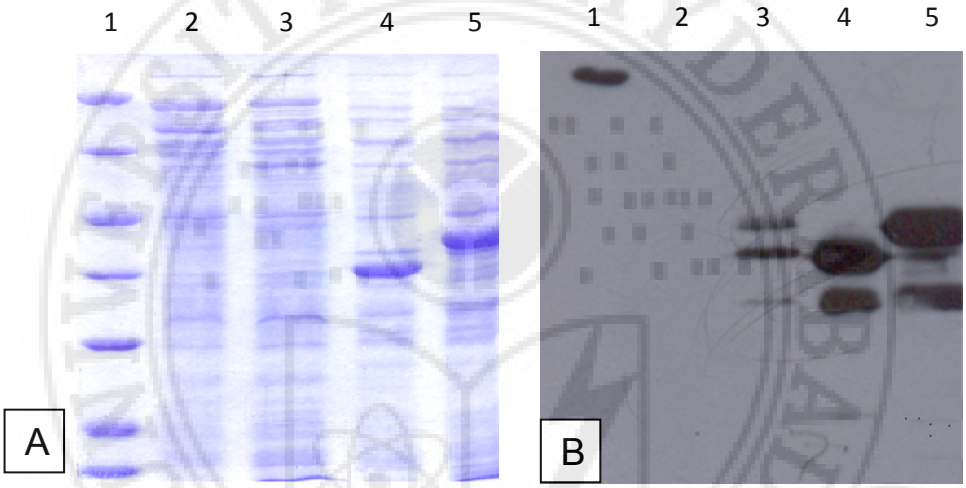

Colony PCR showing the existence and stable maintenance of expression plasmids pHLNS400 and pSM5 in *B. diminuta.* Lane 1 kb ladder, lanes 2 and 3 represent exconjugants of *B. diminuta* containing pHLNS400 similarly lanes 4 and 5 represent exconjugants containing expression plasmid pSM5. Lane 6 represents loading of PCR mixture obtained from *B. diminuta opd* negative mutant.

### 1.4. Expression and detection of preOPH-6His in *B. diminuta*

The exconjugants containing pSM5 was independently grown and expression of OPH was induced by following the protocols described in materials and methods section. Initially the induced cultures were taken to prepare total cellular protein extracts to analyze on 12.5% SDS-PAGE along with similar protein extracts prepared from uninduced cultures and cultures of *B. diminuta opd* knockouts. There was no visible difference among the protein profiles of these cultures when the coomassie stained gels were critically analyzed. However when western blots were performed using anti his antibody, two prominent signals were identified in the lanes loaded with protein extracts prepared from induced cultures. The sizes of these two signals were then compared with the sizes of his-tagged preOPH and mOPH proteins expressed and purified

from *E. coli* cells. Of these two signals the bigger one was identical in size with preOPH whereas the size of the lower signal perfectly matched with the size of the mOPH marker (Fig.1. 11). This results clearly demonstrate that the preOPH coded by stably maintained expression plasmid pSM5 is successfully processed facilitating the detection of both preOPH and mOPH forms in the protein extracts of induced cultures.

**Fig. 1. 11**




**Panel A** coomassie stained SDS-PAGE showing protein extracts prepared from *opd* negative mutants of *B. diminuta* (pSM5). Lane 1 Molecular weight marker, lanes 2 and 3 represent uninduced and induced cultures respectively. Purified recombinant mOPH and preOPH used as size marker are shown in lanes 4 and 5. **Panel B** corresponding western blot showing existence of preOPH and mOPH seen only in induced cultures of *B. diminuta* (pSM5) (Lane 3).

After gaining clear evidence of expression and processing of C-terminal his-tagged protein further experiments were conducted to establish the subcellular localization of precursor and mature forms of OPH in *B. diminuta*. In addition to these two signals, there was an additional signal at much lower size (Fig. 1. 11. Lane 3) which is also seen in lanes loaded with preOPH and mOPH proteins purified from *E. coli* to use as size markers. A brief discussion on this signal will be made in the discussion part.

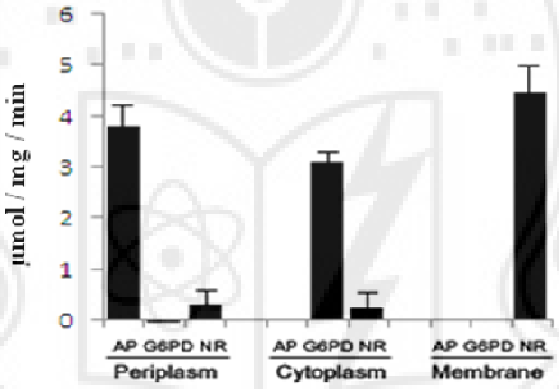
### 1.5. Preparation of subcellular fractions from *B. diminuta*

The subcellular fractions of *B. diminuta* were prepared following standard procedures described in materials and methods section. Before going for detection of preOPH and mOPH, the purity of these fractions was ascertained by assaying marker enzymes as described in materials and methods section. Acid phosphatase was used as marker enzyme to test the purity of periplasmic fraction. Acid phosphatase activity depends on formation of disulfide bonds (Atlung et al., 1989). The reducing environment prevailing in cytoplasm does not support disulfide bond formation to produce active acid phosphatase in cytoplasm (Atlung *et al*., 1989). However, after translocating to periplasm, in the presence of oxidizing environment help in formation of disulfide bonds required for producing active form of acid phosphatase. Therefore assay of acid phosphatase activity in subcellular fractions serves as an indicator to monitor periplasmic protein contamination in cytoplasmic and membrane fractions. Similarly, Glucose-6-phosphate dehydrogenase (G-6-PD, E.C.1.1.1.49), a key enzyme in pentose phosphate pathway (Warburg and Christian, 1936) is used as cytoplasmic marker. Its presence in cytoplasm were strictly demonstrated by number of workers and hence is used as cytoplasmic marker (Noltmann et al., 1961). If periplasmic/membrane fractions are contaminated with cytoplasmic proteins it can be monitored by assaying G-6-PD activity in these fractions.

Nitrate reductase has been studied in several different higher plants and bacteria including in some strains of *E. coli.* In bacteria the nitrate reductase is primarily a membrane bound protein and is linked functionally to cytochrome b1 (Moreno-vivia et al., 1999). Hence the nitrate reductase was used as membrane marker. After preparing the subcellular fractions each fraction was used as source for assaying these marker enzymes. The specific activities obtained were then compared by plotting histograms (Fig. 1. 12). As expected acid phosphatase

activity was more in periplasmic fraction. The activities of nitrate reductase and G-6-PD were found almost negligible, suggesting the periplasmic fraction obtained was pure enough to use in further studies. Similarly no detectable acid phosphatase and nitrate reductase activities were seen in cytoplasm despite of having substantial amount of G-6-PD activity. Such results form basis to claim the cytoplasmic fraction prepared in this study was pure. The membrane fractions prepared were thoroughly washed before redissolving them in phosphate buffer and using it as source for assaying marker enzymes. Only nitrate reductase activity was seen in membrane fractions and neither acid phosphatase nor G-6-PD activity was noticed in membrane fractions indicating the purity of isolated membrane (Fig. 1. 12).

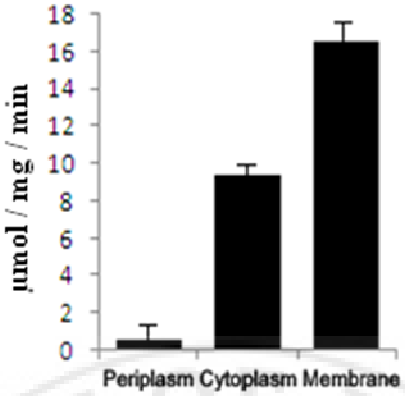
**Fig. 1. 12**




Marker enzyme assays: The bar diagram represents assay of acid phosphatase (AP), glucose-6-phosphate dehydrogenase (G6PD) and nitrate reductase (NR) in periplasmic, cytoplasmic and membrane fractions respectively.

After establishing the purity of subcellular fractions, the proteins prepared from these fractions were used for assaying OPH activity and for analyzing existence of precursor and mature forms of OPH through western blots. There was very negligible OPH activity in periplasm. However in cytoplasmic and membrane fractions substantial amounts of OPH activity was obtained (Fig. 1. 13).

**Fig. 1. 13**

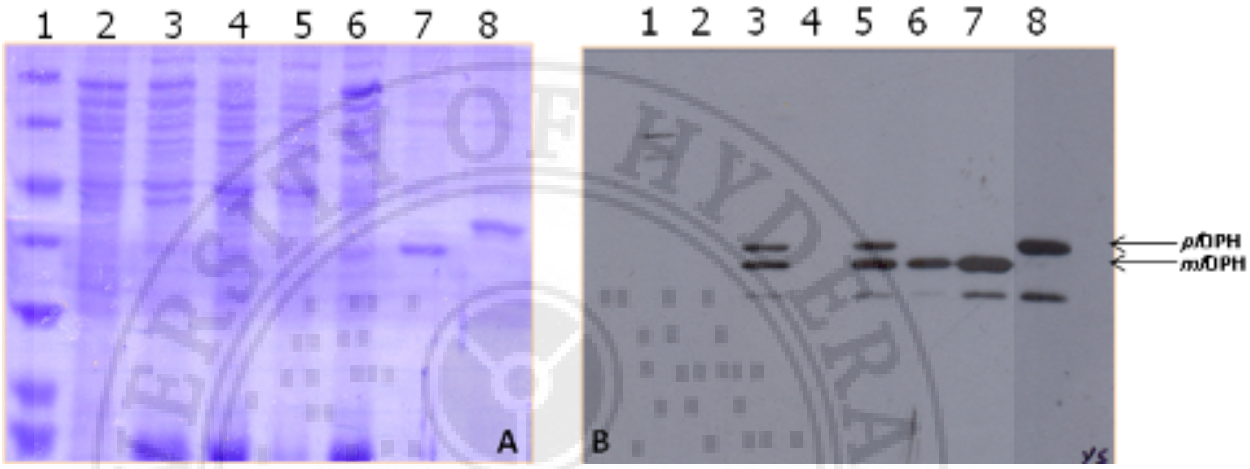



OPH activity in subcellular fractions. The bar diagram represents OPH activity obtained in periplasmic, cytoplasmic and membrane fractions prepared from *opd* negative mutant of *B. diminuta* (pSM5)

As revealed through OPH activity, there was no OPH specific signal in the lanes loaded with protein extracts prepared from periplasmic fraction (Fig. 1. 14 B Lane 4). A prominent signals that were similar in size with the precursor and mature forms of OPH were found segregated between membrane and cytoplasmic fractions. In membrane fraction the OPH signal that corresponds in size with mOPH was alone noticed (Fig. 1. 14B Lane 6). This is a clear indication to claim that plasmid pSM5 coded preOPH was successfully processed and targeted to the membrane in the *opd* negative mutants of *B. diminuta.* There exists no signal in membrane fraction that corresponds to the size of preOPH (Fig. 1. 14B lane 6). The entire preOPH was seen exclusively in cytoplasmic fraction suggesting that the signal peptide of preOPH is required for membrane targeting and is cleaved after OPH reaching to its destination i.e. membrane. A small, reasonably good amount of mOPH is also seen in cytoplasmic fraction. This may be due to existence of excess OPH as a consequence of its induction from strong inducible *lac* promoter. Such excess production of membrane protein is expected to saturate membrane leading to its accumulation in cytoplasmic fraction (Chanal et al., 2003). Nevertheless the experiment clearly

suggests the system developed in the present study is sensitive enough to use for establishing the role of Tat motif in membrane targeting of OPH.

**Fig. 1. 14**



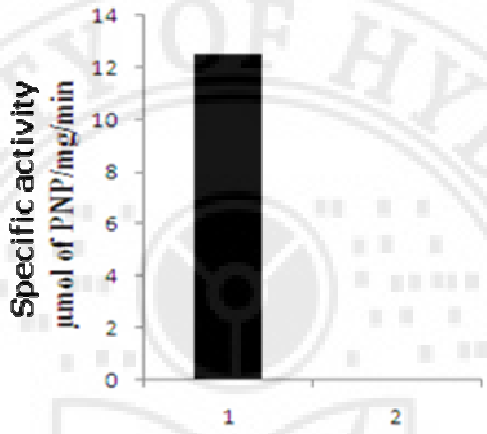

**Subcellular localization of OPH with signal peptide A)** SDS-PAGE analysis of subcellular fractions. **B)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pSM5. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pSM5). Lanes 7 and 8, purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

## 1.6. OPH folds prior to targeting membrane

As stated in introduction, Tat substrates are either periplasmic or membrane associated proteins that fold prior to their targeting / translocating across the membrane (Sargent et al., 2000). The prefolded protein transport is exclusively done by the Tat machinery (Philip et al., 2006). It is not clear if OPH folds prior to its targeting to the membrane. Such information can be obtained by expressing and monitoring the activity of cytoplasmically expressed OPH. In order to have such situation the signal peptide less OPH (mOPH) was expressed in *opd* negative mutants of *B. diminuta* and expression of mOPH is expected to display OPH activity only when it is correctly folded in the cytoplasm. In order to have such situation, the expression plasmid

pHLNS400 that codes mOPH was mobilized into *opd* negative mutants of *B. diminuta.* After expressing mOPH in *B. diminuta,* the OPH activity was measured using *opd* negative mutants of *B. diminuta* as controls. As shown in Fig. 1. 15 substantial amount of OPH activity was seen only in plasmid pHLNS400 containing *B. diminuta* cells. No OPH activity was seen in control cultures.

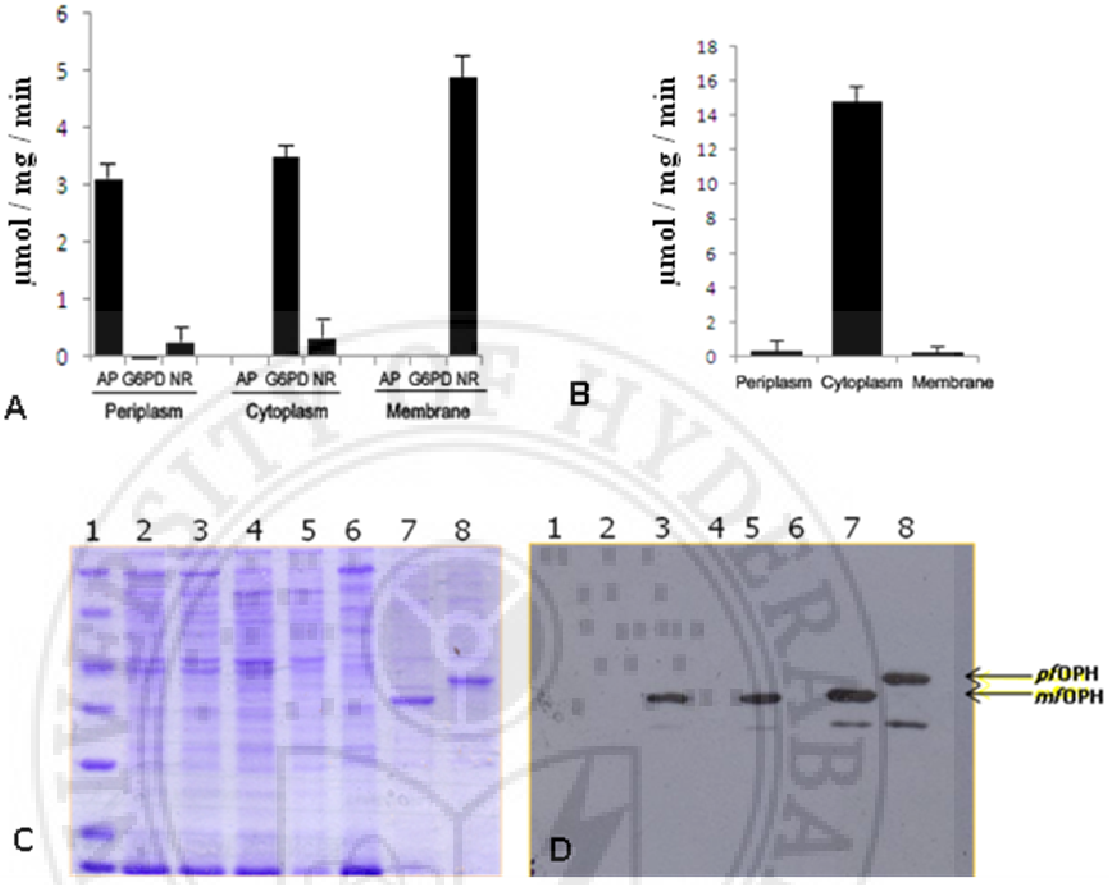
**Fig. 1. 15**




Specific activity of cytoplasmically located mOPH: The bar diagram shows OPH activity shown in *B. diminuta* cells containing pHLNS400 (1) and no OPH activity was seen in *opd* negative mutants of *B. diminuta* (2)

To gain supporting evidence western blots were performed using anti his antibodies. A clear signal corresponding to the size of mOPH was only found in induced cultures (Fig. 1. 16 Lane 2) and as expected no signal was seen in uninduced and control cultures (Fig. 1. 16 Lane 3). These induced cultures were then fractionated into periplasm, cytoplasm and were probed by western blots to detect the existence of mOPH. Almost all mOPH was seen only in cytoplasmic fraction. Not even traces of mOPH was found in the membrane fractions highlighting the importance of the signal peptide in membrane targeting of OPH (Fig.1. 16 Lane 6).

**Fig. 1. 16**




Subcellular localization of mOPH in *B. diminuta* **A)** Assay of marker enzymes acid phosphatase (AP), glucose-6-phosphae dehydrogenase (G6PD), and nitrate reductase (NR) in periplasmic, cytoplasmic and membrane fractions respectively. **B)** OPH activity (µmoles of *p*-nitrophenol formed per minute per mg of protein) in periplasmic, cytoplasmic and membrane fractions. **C)** SDS-PAGE analysis of subcellular fractions. **D)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pPHLNS400. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pPHLNS400). Lanes 7 and 8, purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

The above mentioned studies, have given clear indication to show that the OPH is folded prior to targeting of membrane. Therefore, further studies were planned to elucidate the role of Tat motif found in the signal peptide of OPH. Most of the Tat substrates are extracellular proteins that acquire folded conformation before transporting/targeting the membranes. A number of studies have shown importance of Tat motif in membrane targeting of prefolded proteins (Sargent et al., 2000; Philip et al., 2006). The Tat machinery of *E. coli* fails to recognize

the well known Tat substrate TorA, if the invariant arginines of Tat motifs are changed to lysines (Stanley et al., 2000; Chanal et al., 2003) suggesting that arginines of the twin arginine motif play a key role in membrane targeting of Tat substrates. Therefore, in the present study an attempt was made to generate OPH variants by substituting twin arginines, with amino acids that have either identifical charge or shape. Such variants were then expressed in *opd* negative background to assess their ability to target the membrane.

### 1.7. Expression of OPH variants having substitutions in the invariant arginines of Tat motif

Mutant *opd* genes coding OPH variants in which the invariant arginines of Tat motifs were replaced with amino acids having either similar charge or shape were generated by PCR mutagenesis. In earlier sections a description is made about the construction of expression plasmid pSM5 coding precursor form of OPH. While generating OPH variants plasmid pSM5 (Siddavattam et al., 2006) was used as a template. The forward primer containing the desired mutation and $T_7$ terminator primer was used as the reverse primer. An *Eco*RI and *Bam*HI recognition sequences were introduced in forward and reverse primers to facilitate cloning of *opd* variants in mobilizable vector pMMB206. The forward primers used in the present study and amino acid substitutions made to the invariant arginines of Tat motif are shown in Table.1.1. As shown in Table 1. 1, primer 1 (R5A) and primer 2 (R5K) were used to generate OPH variants having alanine and lysine substitutions at the second invariant arginine residue of the Tat motif. Similarly primer 3 and primer 4 were used to introduce glutamine (R4Q, R5Q) and lysine substitutions (R4K, R5K) at the invariant arginines of Tat motif.

**Table 1. 1.** Oligonucleotides designed to generate amino acid substitutions at the invariant arginines of Tat motif. Mutations introduced are shown with bold case. Start codon of *opd* gene in underlined.

| S. No. | Oligonucleotides | Amino acid substitution |
|---|---|---|
| **1.** | 5' TTG TT*G AAT TC*T AAG AAG GAG ATA TAC ATA TGC AAA CGA GA**G CG**G TTG TGC TC 3**'** **R A** | **R5A** |
| **2.** | 5' TTG TT*G AAT TC*T AAG AAG GAG ATA TAC ATA TGC AAA CGA GA**A AG**G TTG TGC TC 3' **R K** | **R5K** |
| **3.** | 5' TTG TT*G AAT TC*T AAG AAG GAG ATA TA CAT ATG CAA A CG **CAA CAG** GTT GTG CTC AAG 3' **Q Q** | **R4Q, R5Q** |
| **4.** | 5' TTG TT*G AAT TC*T AAG AAG GAG ATA TAC ATA TGC AAA CG**A AAA AG**G TTG TGC TC 3' **K K** | **R4K, R5K** |

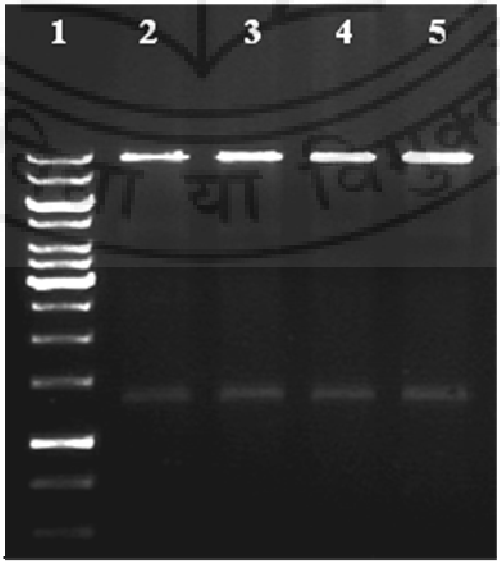
The rational behind selecting these mutations is to identify if similar shape or charge serve as substitutes for twin arginines of Tat motif. After performing PCR using these primers the PCR amplicons (Fig. 1. 17) containing *opd* variants were independently cloned in pMMB206 and the recombinant plasmids containing *opd* variants were confirmed by excising the insert from the multiple cloning site of the vector by digesting with *Eco*RI and *Bam*HI (Fig. 1. 18). These recombinant plasmids coding OPH variants were designated as pHRA400, pHRK400, pHKK400 and pHQQ400. These plasmids coding OPH variants were then independently transformed into *E. coli* S17-1 and used as donors to mobilize them into *opd* negative mutants of *B. diminuta* following the protocols described in materials and methods sections. The exconjugants of *B. diminuta* having expression plasmids coding OPH variants were selected on chloramphenicol, tetracycline plates and were randomly picked to screen for the existence of the expression plasmids by colony PCR using universal forward and reverse primers. All most all

exconjugants screened gave amplification of 1.1 kb amplicon showing the existence of *opd* gene containing expression plasmids (Fig. 1. 19). After ascertaining the existence of expression plasmids in the *opd* negative mutants of *B. diminuta* they were selected for expression of OPH variants by following induction protocols described in materials and methods section.

**Fig. 1. 17**


Agarose gel (0.8%) showing PCR amplification of OPH variants containing amino acid substitutions at invariant arginines of Tat motif. Lane 1 kb ladder, lanes 2 and 3, represent amplicons having alanine and lysine substitution at the second invariant arginine, and lanes 4 and 5 represent amplicons having lysine and glutamine substitutions at both the invariant arginines of Tat motif respectively.

**Fig. 1. 18**



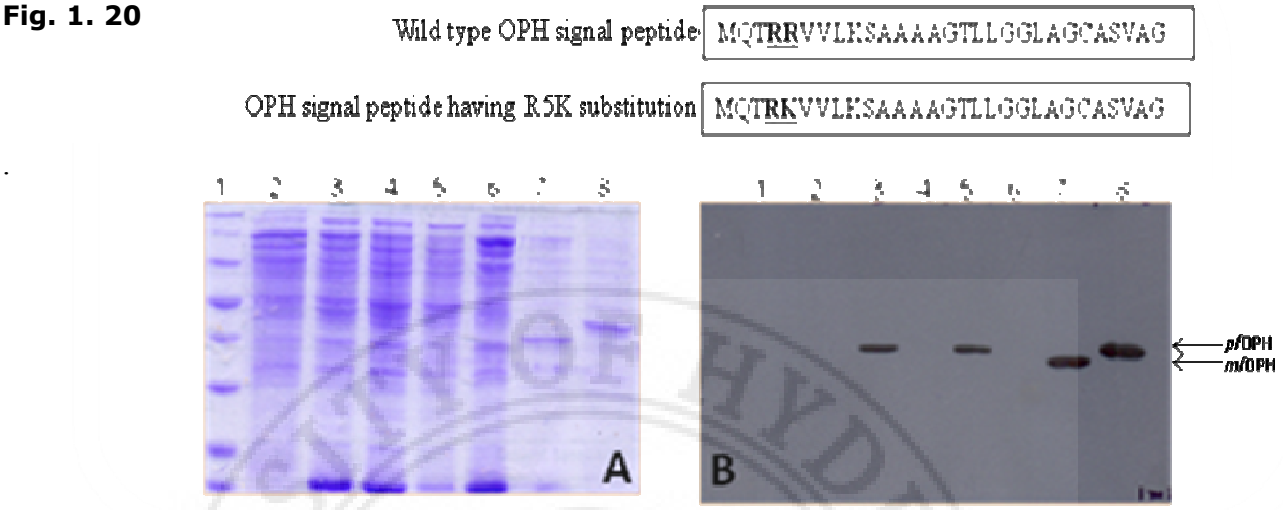

Agarose gel electrophoresis showing the release of 1.2 kb inserts from pHRA400 (lane 1), pHRK400 (lane 2), pHKK400 (lane 3) and pHQQ400 (lane 4) upon *Eco*RI and *Bam*HI digestion.

**Fig. 1. 19**


Agarose gel (0.8%) electrophoresis showing amplification of *opd* variants from *B. diminuta* exconjugants. Lane 1 kb ladder; lanes 2, 3, 4 and 5 represent *B. diminuta* (pHRA400), *B. diminuta* (pHRK400), *B. diminuta* (pHKK400) and *B. diminuta* (pHQQ400) respectively.

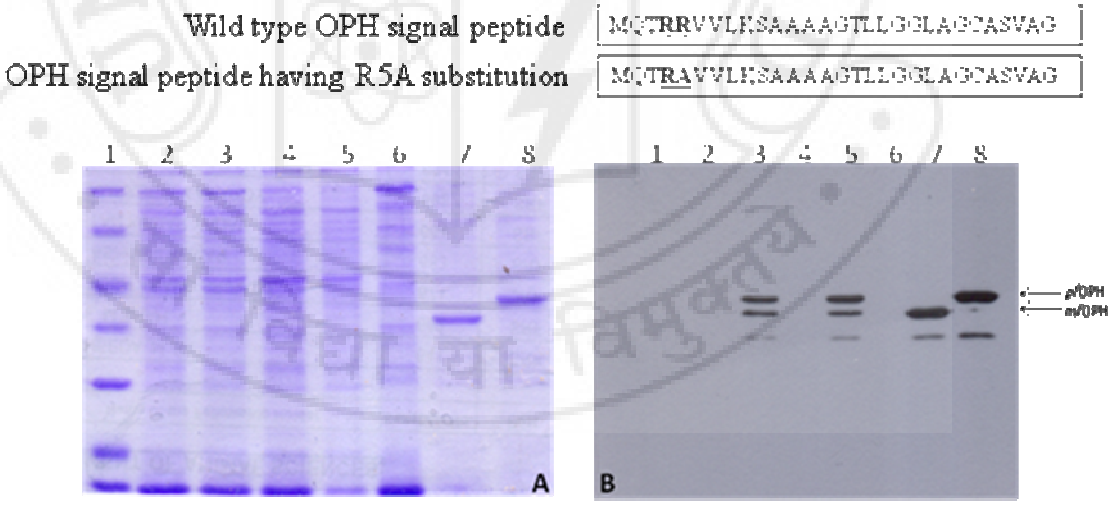
After successful induction, the cells were fractionated and subcellular localization of OPH variants was monitored by performing western blots using anti his anitobody. Among four OPH variants the R5K variant was completely unprocessed and remained exclusively in the cytoplasm (Fig. 1. 20B Lane 5) while all other three variants showed a small extent of processing (Lane 5 of Figs.1. 21B, 1. 22B and 1. 23B). Even in those variants that showed some processing none of the mOPH was targeted to the membrane (Lane 6 of Figs. 1. 21, 1. 22 and 1. 23). The properties of these four variants clearly confirm the dependency of OPH on the arg-arg dipeptide of Tat motif for membrane targeting. Such observation, if seen together with its ability to fold in cytoplasm clearly strengthens proposed hypothesis to claiming OPH as a Tat substrate.

**Fig. 1. 20**



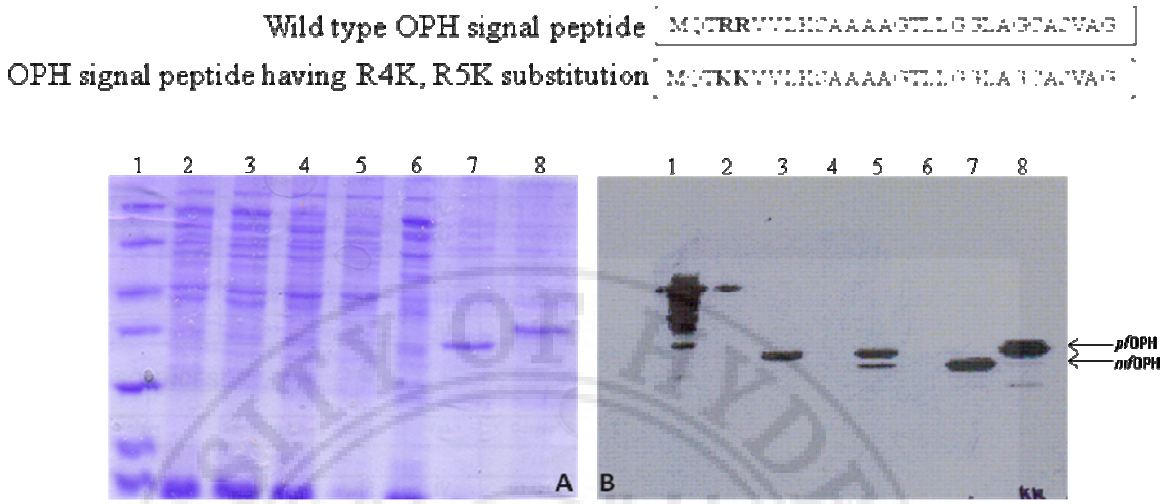

Expression and subcellular localization of OPH (R5K) with C-terminal His tag in *B. diminuta.* **A)** SDS-PAGE analysis of subcellular fractions. **B)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pHRK400. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pHRK400). Lanes 7 and 8, represent purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

**Fig. 1. 21**



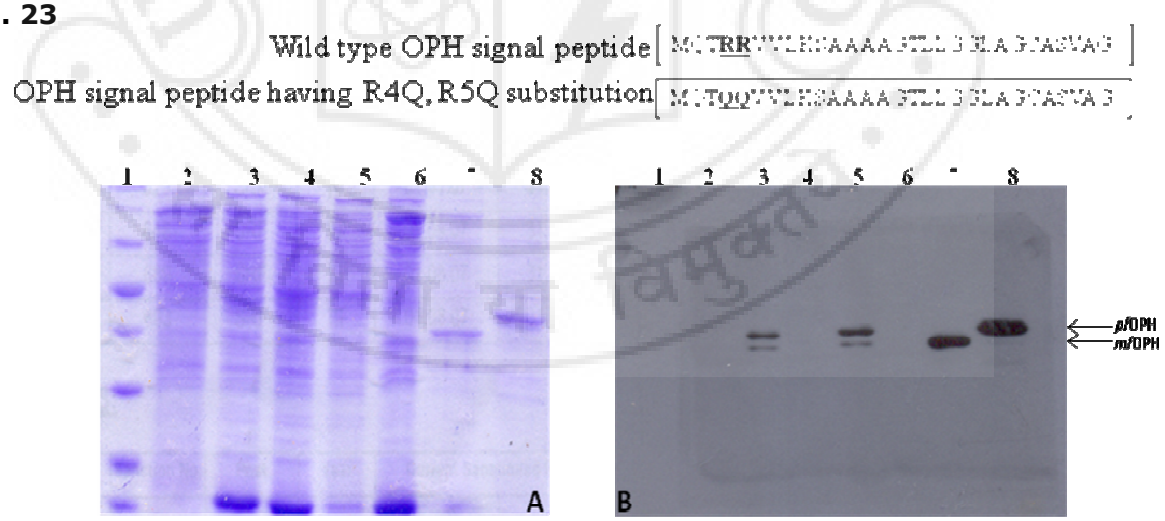

Expression and subcellular localization of OPH (R5A) with C-terminal His tag in *B. diminuta.* **A)** SDS-PAGE analysis of subcellular fractions. **B)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pHRA400. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pHRA400). Lanes 7 and 8, represent purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

**Fig. 1. 22**




Expression and subcellular localization of OPH (R4K, R5K) with C-terminal His tag in *B. diminuta.* **A)** SDS-PAGE analysis of subcellular fractions. **B)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pHKK400. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pHKK400). Lanes 7 and 8, represent purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

**Fig. 1. 23**




Expression and subcellular localization of OPH (R4Q, R5Q) with C-terminal His tag in *B. diminuta.* **A)** SDS-PAGE analysis of subcellular fractions. **B)** Western blot (using anti-His antibody) to detect OPH-6His in SDS-PAGE of total cellular protein of *B. diminuta* (*opd::tet*) without (lane 2) and with (lane 3) expression plasmid pHQQ400. Lanes 4, 5 and 6, periplasmic, cytoplasmic and membrane fractions from induced cultures of *B. diminuta* (pHQQ400). Lanes 7 and 8, represent purified recombinant mOPH and preOPH proteins expressed from *E. coli* BL21 cells.

In an ideal situation, to reinforce these results obtained from western blot analysis, further experiments have to be performed in the *tat* (twin arginine translocases) negative background. In gram negative *E. coli* transport of prefolded proteins depend on novel translocases, often referred as Tat machinery (Sargent et al., 2000). In *E. coli* Tat machinery contains three components coded by *tat* genes, *tatA, tatB* and *tatC* (Berks, 1996; Weiner et al., 1998; Sargent et al., 1998; Bogsch et al., 1998; Sargent et al., 1999). A number of reports are available highlighting the role of Tat components in targeting/translocating prefolded proteins across the inner membrane in a manner driven by pH gradient (Bogsch et al., 1998; Brink et al., 1998; Robinson and Bolhuis, 2004; Philip et al., 2006). The experiments described in the preceeding sections of this chapter provide substantial evidence to claim that OPH is a Tat substrate. However to gain supporting evidence the membrane targeting of OPH needs to be demonstrated both in presence and absence of Tat machinery. Unfortunately the *tat* negative mutants of *B. diminuta* are not available to conduct such studies. In our experience, the preOPH produced in *E. coli* remained completely unprocessed (unpublished data). Therefore the *tat* negative mutants of *E. coli* cannot be used to gain experimental evidence supporting Tat dependency of OPH for membrane targeting. Under these conditions the best option would be to generate a chimeric protein by fusing the signal peptide of OPH to a known Tat substrates that functions efficiently in *E. coli.* Expression of such a chimeric protein in *tat* negative background of *E. coli* helps in assessing the ability of OPH signal peptide in membrane targeting of chimeric protein.

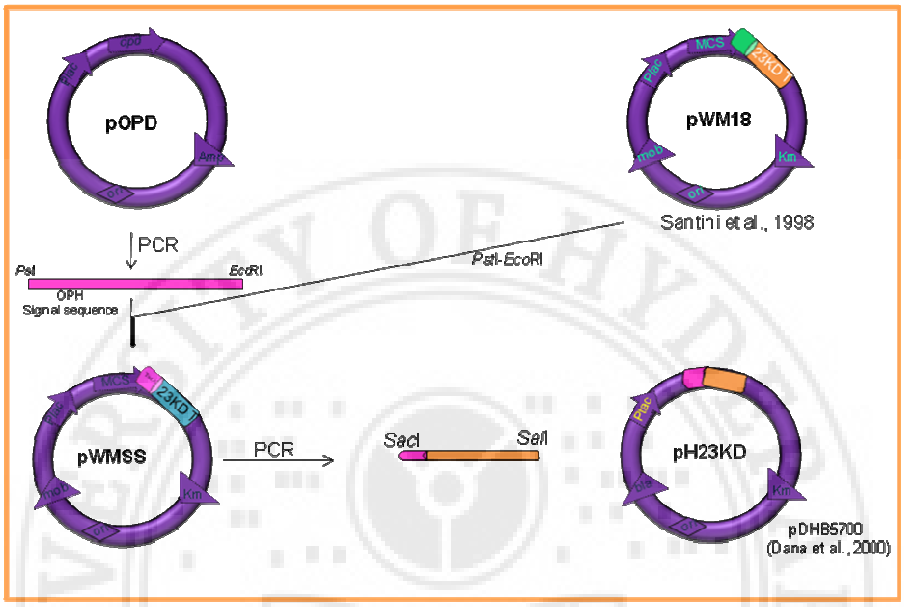
### 1.8. Expression of preOPH-23K in wild type and *tat* mutants of *E. coli* MC4100

The 23K thylakoid protein, that serve as part of oxygen evolving protein complex is a known Tat substrate and it has been successfully used as reporter system for validating the Tat dependency of signal peptides (Chaddock et al., 1995). Therefore in the present study a chimeric

protein, designated as preOPH-23K, was produced by fusing signal peptide coding sequence of *opd* gene to the cDNA region coding for 23K thylakoid membrane protein. A detailed strategy used for generating preOPH-23K fusion is shown in Fig. 1. 24. While generating the inframe fusions, the coding sequence of OPH signal peptide was amplified by PCR using a forward primer (GCA AGC A**CT GCA G**TA AGC AAT CGC AAG GGG GCA GCA TGC AAA CG) and a reverse primer (5' CGC CTG C**GA ATT C**TC CAG CCA CG 3') having *Pst*I and *Eco*RI sites respectively. The amplicon was then cloned in pWM18 (Santini et al., 1998) using *Pst*I and *Eco*RI sites, such that the OPH signal peptide was inframe with the spinach 23kDa oxygen evolving complex protein (23k). The resulted plasmid was designated as pOPS400. The coding sequence for preOPH-23K protein was then cloned as a *Sac*I and SalI fragment into expression plasmid pBDH5700 (De Gier et al., 1998) to give pTLS400. *E. coli* MC4100 wild type and *tat* mutants were transformed with pTLS400 and expression of preOPH-23K was induced by performing standard protocols. After fractionating the induced cultures, the subcellular components were probed with anti-23K antibody to detect preOPH-23K. In support of our claim the 23K protein specific signal was only detected in the membrane fractions of wild type *E. coli* MC4100 and it was totally undetectable in the membrane fractions isolated from Δ*tatAE,* Δ*tatB* and Δ*tatC* mutants (Fig. 1. 25). In *tat* mutants, most of the protein was found in cytoplasm supporting the proposal that OPH is a Tat substrate. However in *E. coli tat* mutants, but not in wild type, a small amount of OPH-23K was detectable in the periplasmic fractions. It may be due to the membrane translocation of OPH-23K as a consequence of aberrant Sec-dependent translocation that occurs when protein fails to be correctly processed by the Tat system. Unlike studies with native OPH these experiments with the preOPH-23K fusion protein did not resolve into the processed and unprocessed forms (Fig. 1. 25). Nevertheless, the data provide clear

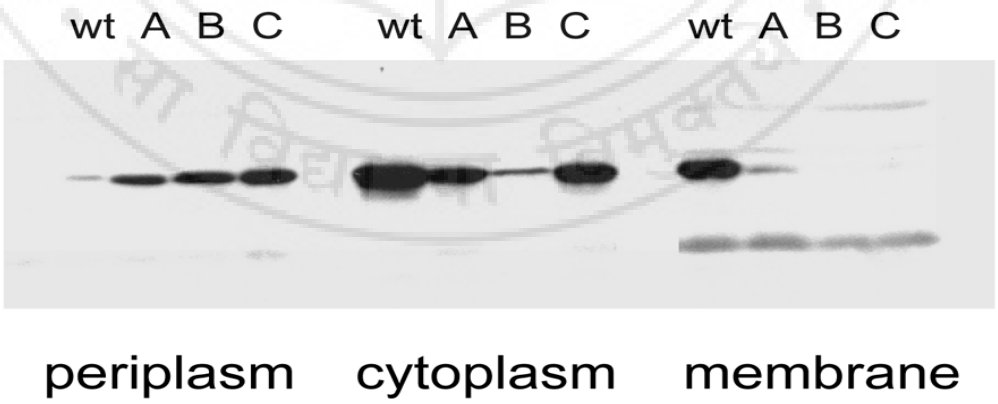
supporting evidence that the OPH signal peptide does facilitate membrane targeting in a Tat dependent manner.

**Fig. 1. 24**




**Strategy used for construction of preOPH-23K fusion**

**Fig. 1. 25**




Western blot analysis and subcellular localization of OPH-23K fusion protein in wild type *E. coli* MC4100 and *tat* mutants. (wt: wild type; A: *ΔtatAE* mutant; B: *ΔtatB* mutant; and C: *ΔtatC* mutants of *E. coli* MC4100)

## 1.9. Discussion

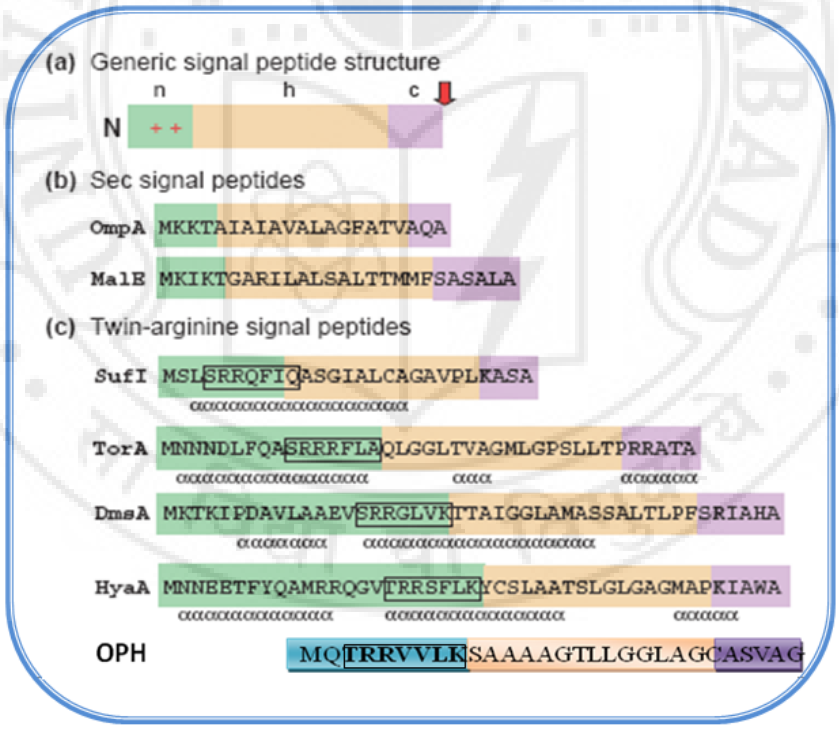
Targeting/transport of proteins to and across the biological membrane is a crucial feature of cellular life (Berks, 1996). In prokaryotes two major routes are used to achieve protein targeting/translocation across the cytoplasmic membrane (Wickner et al., 1991; Driessen et al., 2001; Berks et al., 2005). One of them is Sec-pathway. It is responsible for targeting nearly 90% of the extracellular proteins (Effrosyni et al., 2007). In Sec-pathway the unstructured proteins are transported using energy provided by ATP hydrolysis (Effrosyni et al., 2007). The second important route of protein transport is the recently discovered twin arginine transport pathway (Berks, 1996; Danese et al., 1998; Hinsley et al., 2001; Sanders et al., 2001; Yen et al., 2002; Ize et al., 2003; Robinson and Bolhuis, 2004; Berks et al., 2005; Muller and Klosgen, 2005; Sturm et al., 2006). In Tat pathway proteins are targeted/transported across the membrane using solely the transmembrane proton electrochemical gradient (Palmer and Berks, 2003) to achieve the export of folded proteins across the cytoplasmic membrane. In enteric bacteria mainly four kinds of proteins use Tat machinery for targeting/translocating across the membrane. These are the proteins that depend on large cofactors for activity and the proteins that acquire folded conformation quickly after completion of translation process (Sargent et al., 2000). The third category of proteins are those which do not have favorable extracellular environment for acquiring folded conformation. This category of proteins fold before transport to escape from the proteolytic degradation in the periplasmic space (Philip et al., 2006). In addition to these three classes fourth category of proteins are multisubunit proteins destined to target/translocate across the membrane (Rodrigue et al., 1999; Jack et al., 2004). In this export oriented multiprotein complex only one of the subunits have Tat specific signal peptide (Rodrigue et al., 1999). This subunit recruits interacting partners while in cytoplasm and take them across the membrane as

complex (Philip et al., 2006). These four categories of proteins, if clearly examined, provide valuable clues on preferring the Tat machinery for transport/targeting the membrane. The large cofactor containing proteins, if selected Sec route for translocation have to wait for the export of cofactor to acquire active conformation. Such situation demands concomitant evolution of energy demanding exclusive transport machinery for targeting cofactor to the periplasm. This situation may not be favored during the process of molecular evolution. Berks and his associates have extensively reviewed various aspects of bacterial twin arginine transport pathway (Berks et al., 2005). In one of the review articles they have provided the list of Tat substrates in *E. coli* K-12 strain (Berks et al., 2005). In that exhaustive list, most of the proteins are dependent on large cofactors for activity. However, this list also contained metallo-enzymes such as AmiA, AmiC and CueO (Berks et al, 2005). It is therefore interesting to revisit for the reasons of their Tat dependency. Long Fie and his associates have described the challenges faced by the extracellular metallo enzymes during membrane transport (Long Fie et al., 2000). They have suggested three distant mechanisms of formation of metallo centres in the metallo-enzymes. The first one is represented by the respiratory cytochromes (Thony-Meyer, 1997). Apocytochrome precursors are exported through Sec machinery and covalent attachment of heme appears to be enzyme catalyzed and takes place in periplasm (Long-Fei et al., 2000). The second possible pathway implies a concomitant formation of the metallo centre with the translocation of a precursor crossing the membrane. However, existence of such mechanism is yet to be demonstrated in enteric bacteria. In the third mechanism, the metal ions are incorporated in to proteins while they are in cytoplasm and covert them in to active folded protein before translocating across the membrane (Long-Fei et al., 2000). Discovery of Tat pathway involved in transport of folded proteins indeed supports the third mechanism.

Protein transport depends on the presence of a signal peptide at their extreme N-terminus (Philip et al., 2006). Signal sequence in general have a tripartite structure where a short basic n-region proceeds a longer hydrophobic stretch of amino acids (h-region) followed by the c-region which normally contains a recognition sequence for the enzyme signal peptidase (De Buck et al., 2004; Dilks et al., 2005). The signal peptides are proteolytically removed before release of the mature protein from the transporter (Palmer et al., 2005). As shown in Fig. 1. 26, designed based on the recently appeared review article, the signal sequences that target proteins to the Tat machinery also conform to the overall tripartite structure seen in the signal peptides of Sec-dependent proteins (Palmer et al., 2005). However, they have additional features that distinguish them from those present on Sec targeted substrates. The most striking feature is existence of a consensus Tat motif at the n-region. This motif is designated as SRRxFLK where consecutive arginine residues are almost invariant (Berks, 1996; Stanley et al., 2000; Berks et al., 2005; Palmer et al., 2005). The other amino acids are found with a frequency exceeding 50% and the 'x' represent any polar amino acid (Berkes, 1996; Stanley et al., 2000). Due to the presence of conserved arginine pair the name Tat (twin arginine translocation) is given to the cognate transport system. Existence of Tat protein export system was first demonstrated in chloroplast and then in *E. coli* (Berks, 1994; Settles et al., 1997; Bogsch et al., 1998, Chanal et al., 1998, Santini et al., 1998, Sargent et al., 1998, Weiner et al., 1998). Initially the transport system was designated as membrane targeting and translocation (mtt) and then due to dependency on twin arginine motif it was redesignated as twin arginine translocation (Tat) pathway (Sargent et al., 1998; Weiner et al., 1998). The twin arginines of the signal peptide are crucial for efficient targeting of passenger proteins (Berks, 1996; Brink et al., 1998; Cristobal et al., 1999; Stanley et al. 2000; Sargent et al., 2000). Even a conservative substitutions of either residues for lysines

usually blocks transport (Cristobal et al., 1999; Buchanan et al., 2001). There are other several key differences between twin arginine and Sec targeting signal peptides. In Tat substrates the h-region is generally less hydrophobic than that of their Sec counter parts (Cristoball et al., 1999). Additionally the Tat signal peptides frequently contain a basic residue in the c-region that is almost never found in Sec signals and is proposed to serve as an effective Sec-avoidance mechanism in *E. coli* (Bruser et al., 1998; Wexler et al., 1998). In addition to these primary differences the Tat signal peptides are much longer than the Sec signal peptides. This is due to existence of extended h-region which can be anything up to 23 amino acids long (Cristobal et al., 1999; Philip et al., 2006). A comparison of OPH signal peptide is made with both Sec and Tat signal peptides (Fig. 1. 26).

**Fig. 1. 26**




**Comparison of OPH signal peptide with both Sec and Tat signal peptides. OPH signal peptide showing Tat motif was shown with a box.**

If the sequence of signal peptide of OPH is carefully examined it clearly shows existence of more Tat specific features. It has almost identical Tat motif (T**RR**VVLK) at the n-region

which is followed by a long stretch of h-region. The typical alanine stretch that is found in the well known Tat substrate NapA, the catalytic subunit a nitrate reductase, is also seen in the n-region of OPH signal peptide (MQT**RR**VVLKS**AAAA**GTLLGG LAGCASVAG). It starts from +2 to +9$^{th}$ position relative to RR motif (Maillard et al., 2007). Interestingly there are certain basic differences in the c-region. As described above, in Tat substrates a net positive charge is seen at c-region of signal peptide (Stanely et al., 2000). Existence of net positive charge at the junction of signal peptide cleavage site is considered as one of the critical feature of Tat signal peptides. However, there exists no positive charge at c-region of OPH signal peptide. In certain cases, the positive charge available at the n-terminus of the mature protein also contribute for the overall positive charge required for Sec-avoidance (Bruser et al., 1998; Wexler et al., 1998). If such reports are taken into consideration, the 7$^{th}$ residue of mOPH is arginine and probably it might be contributing for the required Sec-avoidance directing the OPH towards Tat pathway. However, this speculation can only be strengthened through generation of experimental evidence. As Sec-dependent signal peptides have net negative charge at similar position existence of positive charge serves as a critical features to escape Tat proteins from Sec machinery (Wexler et al., 1998; Bruser et al., 1998). Absence of such structural signature in the c-region of OPH signal peptide is indeed worth mentioning.

Tat machinery is well characterized in *E. coli.* (Stanley et al., 2000). Membrane bound TatA, TatB and TatC constitute Tat translocation machinery in *E. coli* (Weiner et al., 1998; Sargent et al., 1998; Bogsch et al., 1998; Sargent et al., 1999; Robinson and Bolhuis, 2004). In this model organism *tatABC* genes are organized as one operon (Weiner et al., 1998; Sargent et al., 1998; Bogsch et al., 1998; Sargent et al., 1999; Wexeller et al., 2000). The *tatE* is a cryptic gene and is shown as duplication of *tatA* (Sargent et al., 1998). Interestingly the products of *tatA*

and *tatB* share considerable sequence similarity (Alami et al., 2003). Both of them possess a hydrophobic transmembrane α-helix localized at the cytoplasmic side of the membrane (Chan et al., 2007). The remaining C-terminal regions of both the proteins are predicted to be unstructured and have been shown to be non-essential for Tat dependent transport (Mori and Cline, 2002; Alami et al., 2003; Gohlke et al., 2005). In some bacteria such as *Bacillus subtillis* there exists no *tatB* gene, it has minimal Tat translocation system comprising just *tatA* and *tatC* (Jongbloed et al., 2006; Barnett et al., 2008). The TatC protein is an absolute necessity for transport of prefolded proteins (Jongbloed et al., 2000; Bolhuis et al., 2001; Allen et al., 2002; Buchanan et al., 2002). It contains six transmembrane domains and it is the largest and most highly conserved Tat components required for transport of prefolded proteins (Behrendt et al., 2004; Ki et al., 2004). These Tat components form complex with a relative mass of 600kDa (Bolhuis et al., 2001). In this complex Tat B and Tat C are found in equimolar ratio together with a variable amounts of TatA (Bolhuis et al., 2001; De Leeuw et al., 2002; Oates et al., 2003). Current evidence suggests the TatBC complex interacts with Tat signal peptides and serve as a specific recognition factors for the twin arginine motif of the signal peptide (Alami et al., 2003). TatA forms a transport channel which takes folded proteins across the membrane when transmembrane proton electrochemical gradient is present (Leake et al., 2008).

The evidence gathered in the present study implicates OPH as a Tat substrate. Initially, the OPH is capable of acquiring folded conformation while it is still found in cytoplasm. This claim is supported by the experiments done involving expression of mOPH in *opd* negative mutants of *B. diminuta.* As shown in Fig. 1. 15. Lane 5, most of the expressed mOPH remained in cytoplasm. The cytoplasmically accumulated mOPH was active, suggesting that OPH has acquired active conformation like the OPH purified from membrane fractions of wild type *B.*

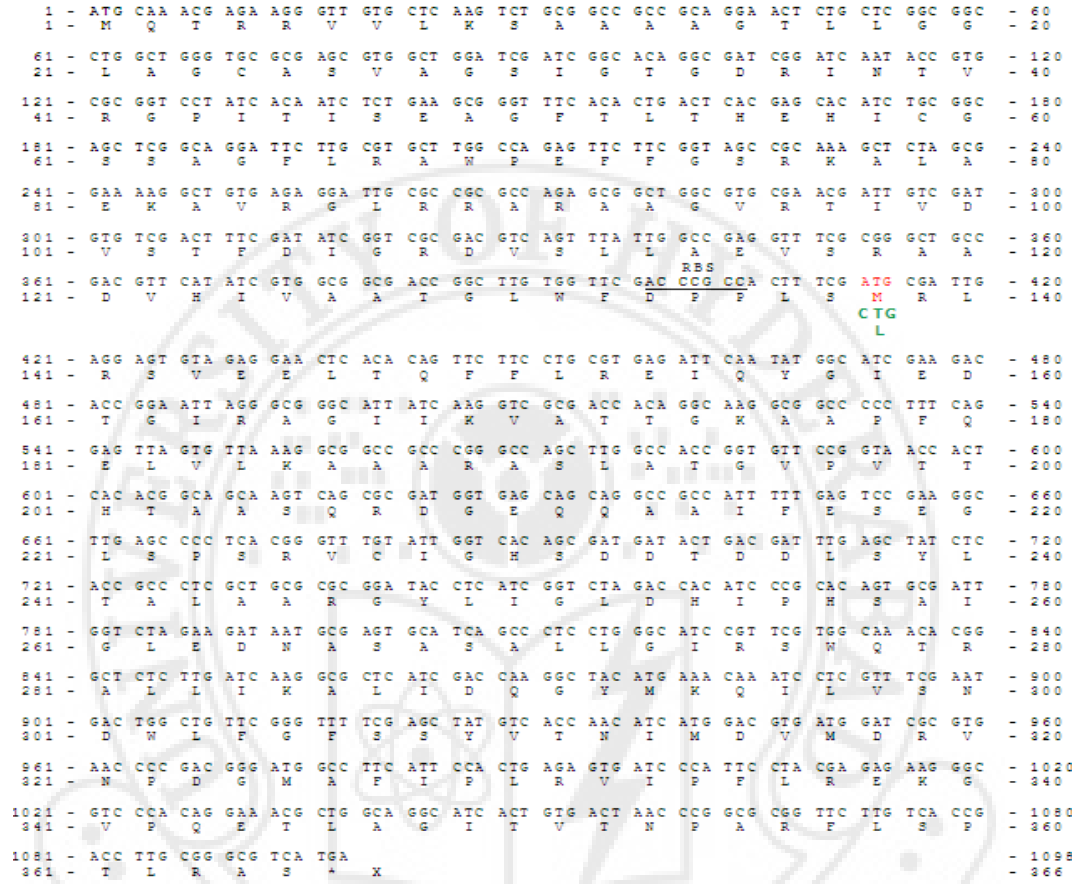
*diminuta* cells. This is in fact one of the primary requirements needed for recognizing a OPH as Tat substrate. The Second line of evidence that was obtained in support of claiming OPH as a Tat substrate, is existence of Tat motif in the n-region of OPH signal peptide. This conserved Tat motif shares 100% structural identity with the consensus Tat motif of well known Tat substrates. The arg-arg dipeptide of Tat motif is absolutely required for membrane targeting of OPH. The OPH variants having R5K and R4K,R5K substitutions at Tat motif were neither processed nor targeted to the membrane (Lane 6 of Figs. 1. 20 and 1. 22). Substitutions of glutamines at arg-arg dipeptide of Tat motif did not facilitate membrane targeting of OPH. Even, the OPH variants having R4Q,R5Q substitutions remained unprocessed. A small portion of processed form generated from these variants remained in cytoplasm. Failure in membrane targeting of OPH variants having substitutions in invariant twin arginines of Tat motif provide an unequivocal support to claim OPH as Tat substrate. In addition to these experiments we have also expressed preOPH-23K fusion protein in *tat* negative background of *E. coli* MC4100. Here again the signal peptide of OPH has successfully driven the 23KDa thylakoid membrane protein to the membrane only in wild type cells of *E. coli* MC4100. The membrane preparation made from the *tat* negative mutants showed no signs of 23KDa protein implicating that Tat machinery is required for successful targeting of 23KDa protein in a OPH signal dependent manner. In the absence of *tat* negative mutants of *B. diminuta* this is the most ideal available situation to gain supporting evidence to show OPH as Tat substrate. Prior to expression of preOPH-23K in *tat* negative mutants, preOPH-6His was expressed in *E. coli* wild type and *tat* negative mutants, imagining that membrane targeting of OPH can be monitored through western blots using anti-his antibodies. Interestingly most of the expressed preOPH, despite of acquiring active conformation, as shown through OPH assay, remained unprocessed. There was no formation of

mOPH either in *tat* mutants or in wild type strains of *E. coli* MC4100 (data not shown). Even in *Pseudomonas aeruginosa PAO11610,* which has close taxonomic link with *B. diminuta,* the preOPH-6His remained unprocessed. Interestingly, the heterologously expressed preOPH-6His acquired active conformation and showed substantial OPH activity when methylparathion was used as assay substrate. Though it is a surprising observation, precedents of having species specific Tat machinery are not uncommon in literature. Successful transport/targeting of heterologously expressed Tat substrates is seen only when cognate Tat machinery is co-expressed (Emmy et al., 2008). In the light of these reports possible existence of unique Tat machinery in *B. diminuta* serving, exclusively for transport of prefolded proteins cannot be ruled out. Alternatively the OPH may be part of multisubunit protein complex. Assembly of such complex in cytoplasm may be prerequisite for successful targeting to the membrane. Absence of such partner proteins might be preventing formation of transport ready OPH complex in *E. coli* and *PAO11610.* Further experiments involving immunoprecipitation of OPH complex from the membranes of *B. diminuta* is required to identify if OPH is really a multisubunit complex. Similarly, cloning of *tat* genes that contribute for Tat machinery of *B. diminuta* is essential to examine variations associated with Tat components.

In western blots used for detecting OPH-6His always gave a third OPH specific signal in addition to pre and mOPH specific signals. The size of the signal as estimated from SDS-PAGE is 23kDa. The presence of this signal was consistent and it was even found when OPH is expressed in *E. coli* and *PAO11610.* Such consistent appearance of this signal prompted to revisit the *opd* sequence for presence of second translational site. Exactly 7 bases upstream of the codon that specify 138 metheonine residue of OPH (Mulbry and Karns, 1989), a sequence that serve possible ribosomal binding site was noticed. Existence of such strong ribosomal

binding site, that to 7 bases upstream of ATG codon suggests for possible second translation site within OPH coding sequence (Fig. 1. 27).

**Fig. 1. 27**




Sequence of *opd* gene and deduced amino acid sequence. The possible second translation site is shown with red color and leucine (M138L) is indicated with green.

If the ATG specifying 138$^{th}$ metheonine of OPH is taken as initiation codon, the size of the predicted translation product would be 23KDa, which has size identity with the third OPH specific signal. In view of size similarity with the observed signal and existence of a possible ribosomal binding site, a second translational start site was predicted within the *opd* coding sequence. To gain an experimental evidence for such prediction, the metheonine was successfully replaced with leucine by performing site directed mutagenesis (data not shown). If the 23K OPH specific signal is formed due to existence of second translation site, expression of

OPH variant, (M138L) should abolish its generation. Interestingly the cultures expressing (M138L) have also given OPH specific signal at 23kDa region, suggesting that the 23kDa product might be a consequence of mild proteolytic cleavage.

### 1.9.1. Signal Recognition Particle

The signal recognition particle (SRP), a conserved ribonucleoprotein particle, recognizes the signal sequence as soon as it emerges from the ribosomal polypeptide exit tunnel (Richardson, 2000). In fact *E. coli* contains a combination of 23 redox enzymes which include dimethyl sulfoxide (DMSO), trimethylamine-*N*-oxide (TMAO), formate, nitrate, and others (Richardson, 2000). Some redox enzymes are anchored at the cytoplasmic side of the membrane, but the majority of them are located at the periplasmic side (Chan et al., 2009). They also contain cofactors such as molybdopterin, Fe-S, and Ni-Fe clusters that are incorporated into the protein prior to their targeting and translocation across the cytoplasmic membrane (Chan et al., 2008).

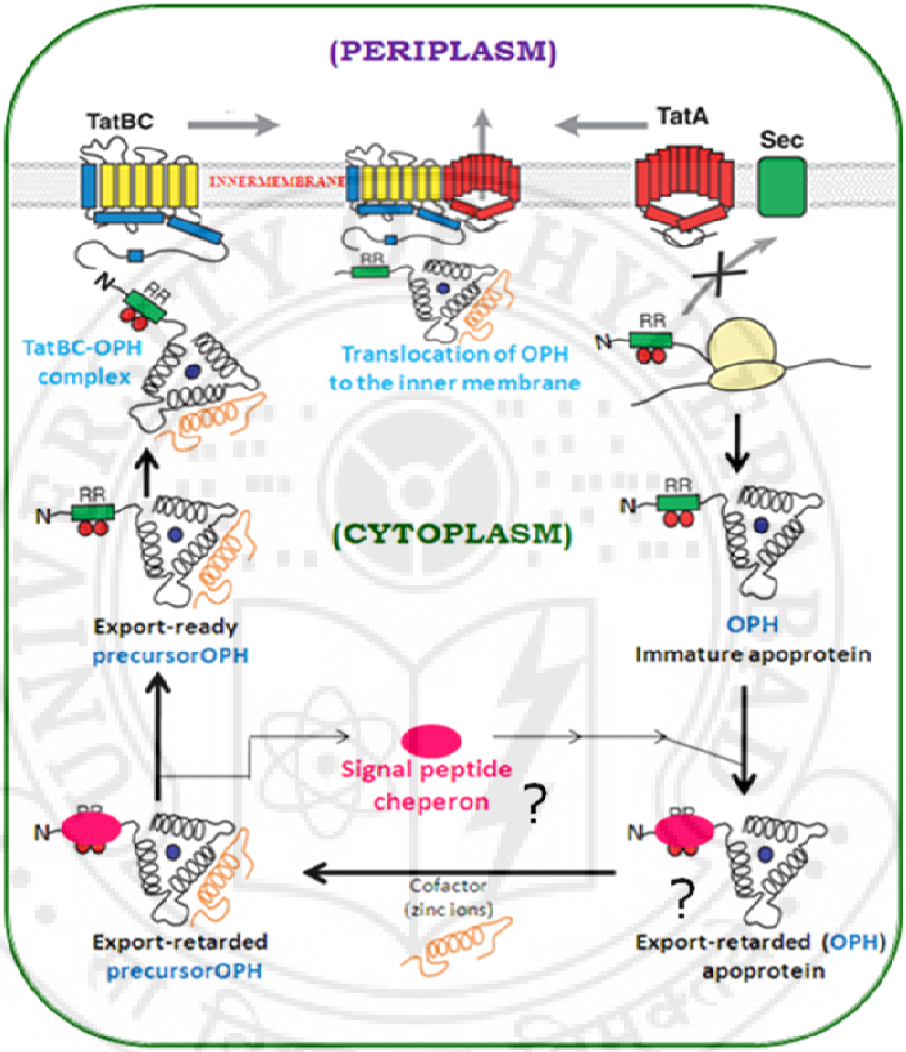
In multisubunit containing proteins such as the *E. coli* trimethylamine N-oxide reductase TorA is subject to a chaperone-mediated quality control process known as ‘‘Tat proofreading,’’ by TorD which prevents premature targeting of TorA until all biosynthetic processes are complete (Jack et al., 2004). The TorD chaperone recognizes the TorA signal peptide and binds directly to it, thus shielding it from the Tat transporter (Jack et al., 2004; Hatzixanthis et al., 2005). *E. coli* TorD is a member of a family of peptide-binding proteins that share a common fold comprised almost entirely of α-helices and completely devoid of β-strand (Pommier et al., 1998; Jack et al., 2004; Sargent, 2007; Genest et al., 2008). Tat proofreading involves a direct, specific interaction between TorD and the TorA signal peptide (Jack et al., 2004; Hatzixanthis et al., 2005), however the mechanism of signal peptide recognition and release is not fully

understood (Grant et al., 2008). Similarly Chan and his associates reviewed in one of their pioneering articles that the operons encoding many redox enzymes appeared to contain an extra gene product that did not appear to be part of the final holoenzyme (Chan et al., 2009). These genes were found to encode chaperone proteins specific to the redox enzymes in that operon with apparent roles in the activation or assembly of the holoenzyme complexes. Based on the roles of such proteins they are designated collectively as redox enzyme maturation protein (REMP), where a protein is defined as such when it is involved in the assembly of a complex redox enzyme but does not constitute part of the final holoenzyme (Natale et al., 2008).

TorD and DmsD are required for Tat proofreading and targeting of molybdoproteins TMAO reductase (TorA) and DMSO reductase (DmsA) to the periplasm via twin arginine translocation pathway (Sargent, 2007). However, when the signal peptides of TorA and DmsA are fused to the heterologous proteins, they were successfully targeted to the periplasm via Tat pathway but did not depend on TorD and DmsD chaperons for Tat proofreading activity. Hence the authors concluded that the Tat proofreading proteins bind to the signal peptides of their cognate substrates and appear to be important only for the assembly and export of these specific proteins (Sargent, 2007). If this observation is taken into consideration, it would be logical to expect existence of OPH specific signal recognition particle (SRP) in *B. diminuta* and *dmsD,* which code for signal recognition particles are closely located and found as part of the operon with the genes that code for their cognate substrates (Sargent, 2007; Cristina et al., 2008). In the light of these reports we have examined the *opd* gene cluster, to identify if there is an ORF that show similarity with either TorD or DmsD. In the transposon-like organization of *opd* gene cluster such homologue was not identified. However, its existence is certain in *B. diminuta* chromosome or in indigenous plasmid pCMS1. Based on the experimental evidences gained in

the present study a model is generated to highlight salient features of membrane targeting of OPH in *B. diminuta* and the same is shown in Fig. 1. 28.

**Fig. 1. 28.**




Proposed model showing the mechanism of membrane targeting of OPH via twin arginine translocation pathway in *B. diminuta.*

# MATERIALS & METHODS

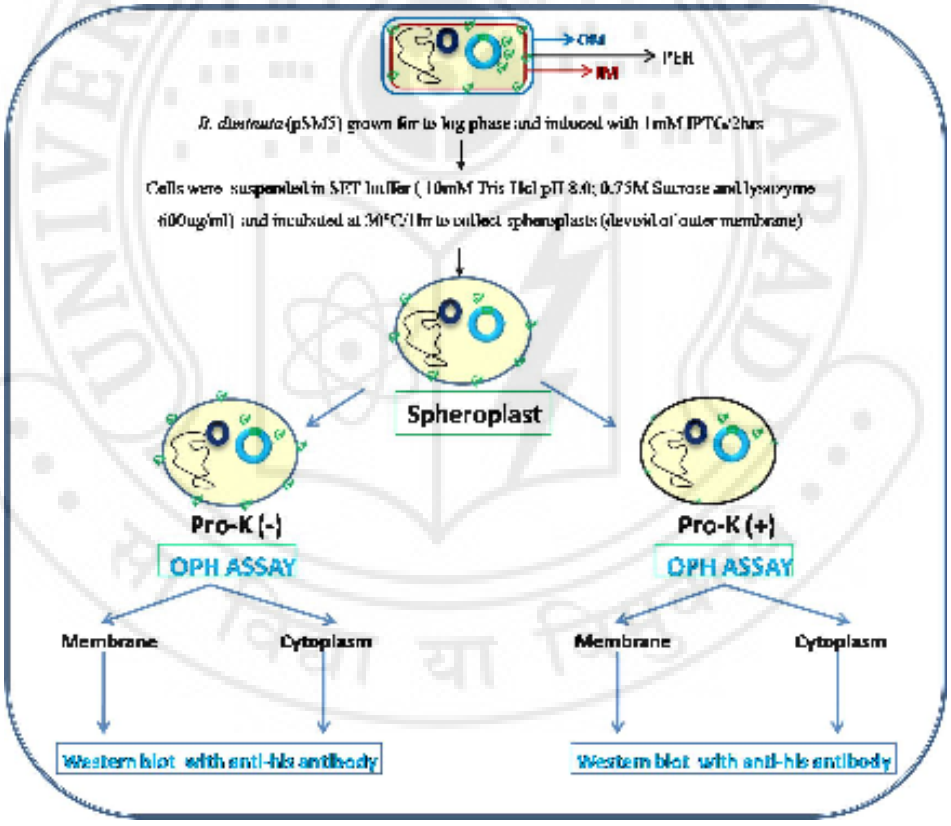
Experiments designed in the proceeding chapter have provided substantial evidence to show that OPH is a Tat substrate and is targeted to the inner membrane in prefolded form. However these studies have not given any clue about its membrane topology. The biochemistry of OPH is very well known and its crystal structure provided no clue on existence of transmembrane domain (Benning et al., 2001). In the absence of such structural fold its association to the inner membrane must be through interaction with membrane anchored proteins. Before gaining evidence on mechanism of membrane anchoring / association, it is important to determine topology of OPH in the inner membrane of *B. diminuta.* Generally PhoA fusions are successfully used for establishing the topology of membrane proteins (Matthew et al., 2008). The conditional folding of PhoA in periplasmic environment serves as marker to identify if the protein / domain fused to PhoA is facing towards cytoplasmic or periplasmic face (Matthew et al., 2008). Unfortunately such experimental design may not be successful in case of proteins targeting in prefolded form. As PhoA fails to fold in cytoplasm, if it is fused to a protein destined to the membrane in prefolded form, it is expected to create a semi folded chimeric protein that serves as a bad substrate for Tat machinery. Precisely for this reason, reports of using PhoA for determining the membrane topology of prefolded proteins are scarce. Therefore, in the present study alternative techniques were used to elucidate membrane topology of OPH. One of them is based on proteinase K treatment. If spheroplasts prepared from *opd* negative mutants of *B. diminuta* (pSM5) are treated with proteinase K, it should eliminate OPH if it is facing periplasmic space. Alternatively, if OPH is on the cytoplasmic face it will not gain access to proteinase K and hence remains unaffected. Monitoring the OPH activity followed by detecting OPH

through western blots in membrane samples prepared from proteinase K treated spheroplasts is expected to provide primary evidence on the membrane topology of OPH.

## 2. 1. Spheroplasts – Proteinase K treatment

The *opd* negative mutants of *B. diminuta* having expression plasmid pSM5 were induced before using them for preparation of spheroplasts. The schematic representation showing preparation of spheroplasts and procedure followed for proteinase K treatment is depicted in Fig. 2.1.

**Fig. 2.1.**

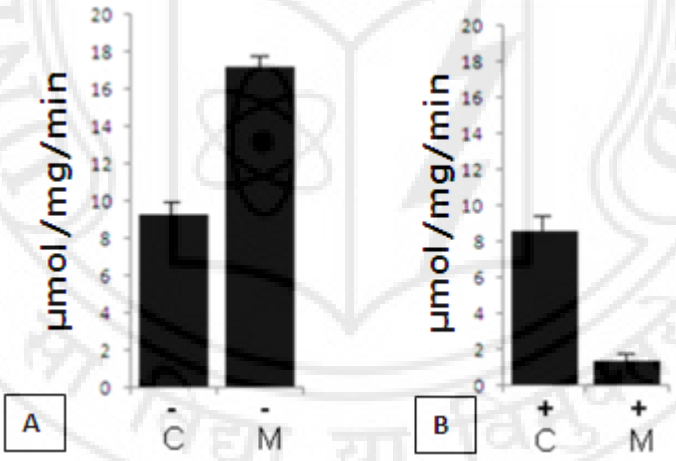



Preparation of spheroplasts and procedure followed for proteinase K treatment

The spheroplasts prepared were divided into two portions. One of them was used for proteinase K treatment, whereas the second portion served as control. Membrane and

cytoplasmic fractions prepared from the proteinase K treated and untreated portions were then used for performing OPH activity. The specific activity of OPH obtained from all these fractions was calculated and a bar graph showing the specific activity is shown in Fig. 2. 2. The data obtained in the study clearly shows that the OPH activity levels remained unaffected in the proteinase K treated and untreated cytoplasmic fractions. However, there was huge difference in the specific activity levels obtained from membrane fractions. In membrane fractions prepared from proteinase K treated spheroplasts the OPH activity was very low when compared to the activity levels obtained from the membrane preparations obtained from the control samples (Fig. 2. 2 B). This data clearly suggests that the OPH is associating the membrane facing periplasmic space.

**Fig. 2. 2**

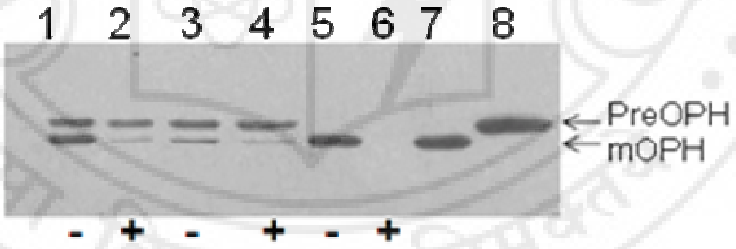



Specific activity of OPH in cytoplasmic and membrane fractions prepared from spheroplasts untreated **(A)** and treated **(B)** with proteinase K. (C: cytoplasmic fraction, M: membrane fraction)

Western blots performed to detect OPH in proteinase K treated and control samples supported this claim (Fig. 2.3). In Fig. 2.3 lanes 1 and 2 represent typical protein extracts

prepared from proteinase K untreated (Lane1) and treated spheroplasts (Lane 2). In these two lanes, the intensity of signal corresponding to mOPH has lot of significance. In proteinase K treated spheroplasts its intensity is very weak, when compared to the intensity of similar signal obtained from control spheroplast proteins (Fig. 2.3 Lanes 1 and 2). Substantial amount of mOPH was eliminated due to proteinase K treatment. Further, when membrane and cytoplasmic fractions prepared from these control and proteinase K treated spheroplasts were probed to detect OPH signals, a clear picture has resulted on membrane topology of OPH. As shown in Fig. 2.3 no difference is seen in the signal intensity of preOPH, exclusively found in the cytoplasmic fractions prepared from both control and proteinase K treated spheroplasts (Fig. 2.3 Lanes 3 and 4). However, in the membrane fractions, obtained from the proteinase K treated sample, proteinase K has eliminated most of the OPH and as a consequence there was no OPH positive signal was obtained (Fig. 2.3 Lanes 6).

**Fig. 2.3.**




Localization of OPH by Proteinase K treatment of spheroplasts. Western blots (using anti-$His_6$ antibody) of *B. diminuta* (pSM5). Lanes 1 and 2, spheroplasts; lanes 3 and 4, cytoplasmic fractions from spheroplasts; lanes 5 and 6, membrane fractions from spheroplasts. Lanes 1, 3 and 5, without Proteinase K (-); lanes 2, 4 and 6, after Proteinase K treatment (+). Lane 7, *E. coli* BL21 (pHNS400), expressing mOPH; lane 8 *E. coli* BL21 (pHYS400), expressing preOPH.

The proteinase K gain no access to the preOPH due to existence of membrane and hence the signal corresponding to preOPH in both proteinase K treated and control cytoplasmic samples were unaffected (Fig. 2.3 Lanes 3 and 4). The data gained from proteinase K treated

spheroplasts of *B. diminuta* provided substantial clue on existence of OPH on the periplasmic face of inner membrane (Fig. 2.3). In order to provide supporting evidence, the preOPH-GFP proteins were generated and expressed in *B. diminuta* cells.

GFP is one of the most efficient reporter systems available today. It has been successfully used to study subcellular localization of proteins both in prokaryotes and eukaryotic systems (Danielle et al., 2007). Usage of GFP for studying Tat mediated protein targeting/translocating across the membrane received much attention through the pioneering workers of Danielle and his associates (Danielle et al., 2007). Using elegant experimental design they have shown that GFP retained fluorescence only when exported to periplasm through Tat machinery. If the GFP is fused to Sec dependent signal peptide it is successfully exported into the periplasm via Sec pathway and failed to gain the folded confirmation to emit the characteristic fluorescence. Finally such unfolded GFP got degraded due to existence of proteolytic machinery in periplasmic space (Matthew et al., 2002). This excellent findings have kept GFP based reporter systems ahead of other techniques used for studying the membrane topology of Tat substrates. In order to exploit such eligent system for understanding the membrane topology of OPH, the signal peptide of OPH was fused to GFP and the chimeric preOPH-GFP was expressed in *B. diminuta.* If fluorescence is noticed in the cells expressing OPH-GFP, it clearly indicate transport of OPH through Tat pathway. Presence of fluorescence in preOPH-GFP expressed *B. diminuta* was monitored through fluorescence microscopy, FACS and by using anti-GFP antibodies. Further, an SsrA signal was also added at the C-terminus of OPH-GFP to know if OPH-GFPSsrA can escape cytoplasmically located ClpXP proteolytic machinery by crossing the membrane.

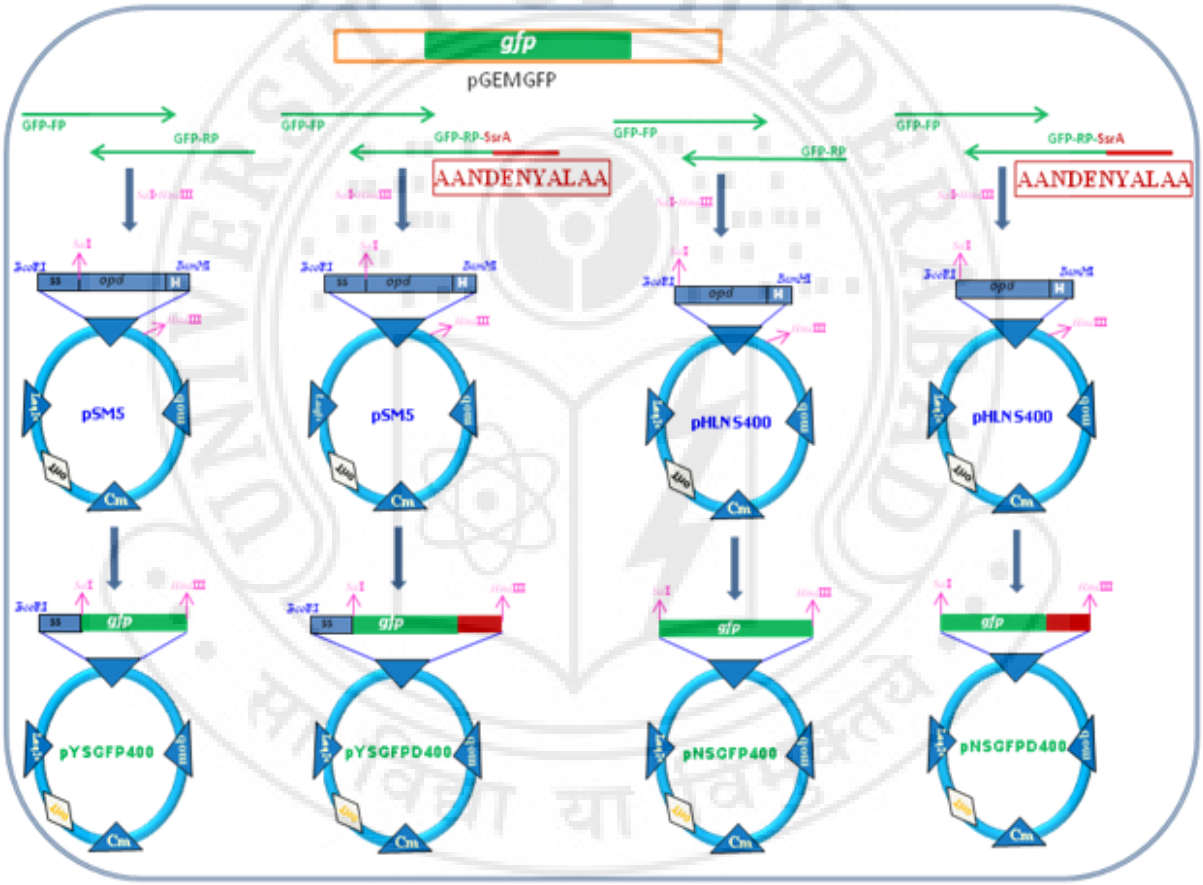
**2. 2. Construction of *opd-gfp* fusions**

Plasmid pSM5 (Siddavattam et al., 2003) was used to generate inframe fusions of *gfp* with the *opd* gene. The *opd* gene contains a unique *Sal*I site at 300bp downstream of the start codon ATG and pSM5 has a unique *Hind*III site in the multiple cloning site of the vector. Hence digestion with *Sal*I and *Hind*III removes the 3' end of the *opd* gene. To generate an expression plasmid encoding preOPH-GFP, the *gfp* gene was amplified from pBQGFP using forward (5'GGA GAT AT***G TCG AC***G GCT AGC AAA GGA GAA GAA CTC 3') and reverse (5'GCT TTG TTA GCA ***AAG CTT*** TCC TCA GTT GTA CAG 3') primers containing recognition sequences for restriction endonucleases *Sal*I and *Hind*III (shown in bold case), and cloned into the *Sal*I and *Hind*III sites of pSM5 to generate pYSGFP400 in which *gfp* is fused in frame with the 5' end of *opd* that includes the coding region for the signal peptide of OPH.

A similar cloning strategy was followed to introduce the SsrA signal (Tu et al., 1995) downstream of the GFP coding sequences, except that an oligo (5' GCG ATG AAG CTT GCA TGC TTA AGC TGC **TAA AGC GTA GTT TTC GTC GTT TGC TGC GTC GAC** GTT GTA CAG TTC ATC CAT GCC 3') with codons (shown in bold case) that specify amino acids (AANDENYALAA) at the C-terminal end of GFP was used as the reverse primer. The plasmid encoding preOPH-GFP with an SsrA signal was designated as pYSGFPD400. Comparable plasmids encoding the mature form of OPH (mOPH) were constructed using an identical strategy, except that pHLNS400 was used as the recipient plasmid. The plasmids coding for mOPH-GFP and mOPH-GFP with an SsrA signal were designated as pNSGFP400 and pNSGFPD400 respectively. A diagrammatic representation showing generation of *opd-gfp*

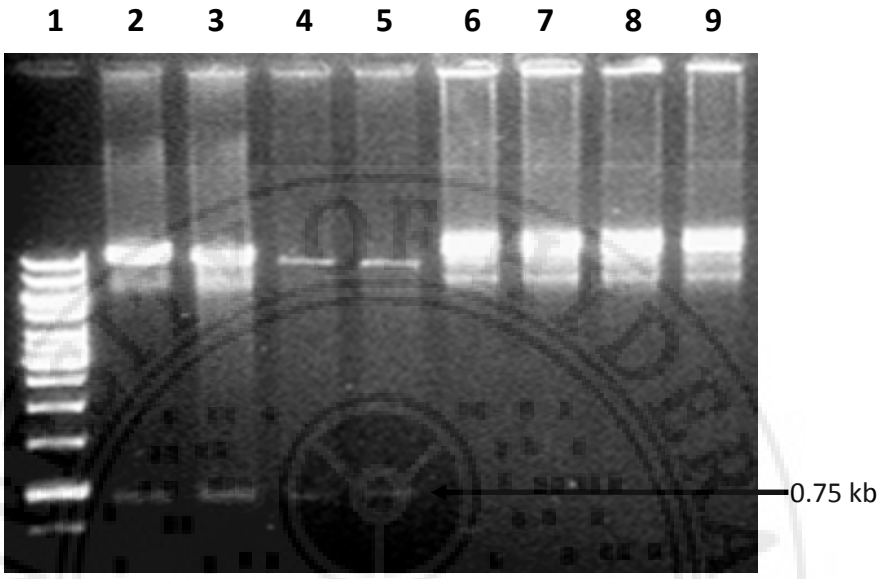
constructions is shown in Fig. 2. 4. After successful construction, these recombinant plasmids were isolated and existence of insert was confirmed by digesting them with *Sal*I and *Hind*III. An agarose gel showing the presence of *gfp* insert in pYSGFP400, pYSGFPD400, pNSGFP400 and pNSGFPD400 is shown in Fig. 2. 5.

**Fig. 2.4**




A diagrammatic representation showing the construction of expression plasmids coding for preOPH-GFP (pYSGFP400) and mOPH-GFP (pNSGFP400) with and without SsrA signal at the C-terminus.

**Fig. 2.5.**




Construction of *opd-gfp* and *opd-gfpssrA* fusions. Lanes 1 kb ladder, lanes 2, 3, 4 and 5 represents pYSGFP400, pYSGFPD400, pNSGFP400 and pNSGFPD400 digested with restriction endonucleases *Sal*I and *Hind*III respectively; lanes 6, 7, 8 and 9 are corresponding uncut plasmids. Release of 0.75kb insert containing *gfp* gene is shown with arrow mark.

### 2. 3. Mobilization and expression of OPH-GFP fusion protein

After constructing expression plasmids coding preOPH-GFP (pYSGFP400), mOPH-GFP (pNSGFP400), preOPH-GFPSsrA (pYSGFPD400) and mOPH-GFPSsrA (pNSGFPD400) were mobilized into *B. diminuta* by performing biparental conjugation described in materials and methods section and the cells containing these expression plasmids were used for conducting FACS analysis and to obtain fluorescence image.

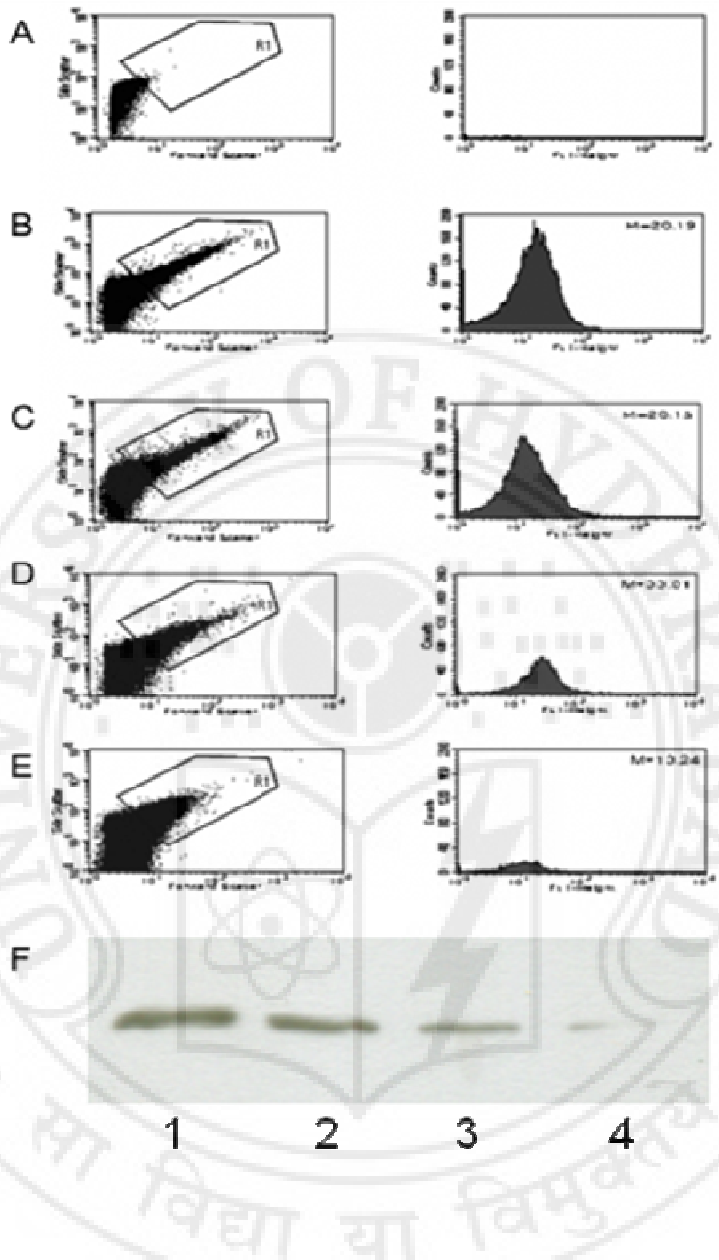
### 2. 4. FACS analysis

*B. diminuta* cells having various OPH-GFP fusions were induced before diluting the culture (1:10) in phosphate-buffered saline (pH 7.6). These cells were then analyzed in a BD

Pharmingen FACSort machine, following the protocols described in materials and methods section. Fluorescence emission was read with a 515/40 band pass filter. As shown in Fig. 2.6., the fluorescence channel boundaries (gates) were set to sort the population with highest fluorescence intensity. Out of a total events, approximate number of events that fell within the imposed gates were collected and amplified for the FACS selection. As shown in Fig. 2. 6 panel A shows the FACS scan of fluorescence from *B. diminuta* cells having no expression plasmid which served as reference control. Panels B to E shows the sorted population for FACS selection of high fluorescence from the selected gates (R1) emitted from *B. diminuta* cells expressing preOPH-GFP (B), preOPH-GFP-SsrA (C), mOPH-GFP (D), and mOPH-GFP-SsrA (E). Scan of an individual enhanced-GFP expression (box/gate R1) was compared with wild type *B. diminuta* cells having no GFP expression (Panel A gate R1).

The mean fluorescence value was calculated for cells expressing preOPH-GFP (20.15) was not affected by the presence of a C-terminal SsrA (20.19) suggesting that the preOPH-GFP-SsrA (with OPH signal peptide and SsrA) was not accessible to the intracellular proteolytic ClpXP machinery and had been translocated across the inner membrane in *B. diminuta* (Fig. 2.6 B, C). In accordance with the above results, when the expression plasmids coding mOPH-GFP (without OPH signal peptide and SsrA) and mOPH-GFPSsrA (without OPH signal peptide and with SsrA) were induced in *B. diminuta,* the mean fluorescence (33.01) was retained with cells expressing mOPH-GFP, but a significant level of reduction in mean fluorescence was noticed with the cells expressing mOPH-GFPSsrA (Fig. 2.6 D, E).

**Fig. 2.6.**




FACS analysis of *B. diminuta* cells expressing OPH-GFP fusion proteins. Left column, dot plots of raw data. The polygon in each graph identifies the gate R1 used for generation of histograms. Right column, histograms of counts from gate R1 for each strain. **(A)** Wild-type *B. diminuta* cells. (B-D) *B. diminuta* cells expressing preOPH-GFP **(B)**, preOPH-GFP-SsrA **(C)**, mOPH-GFP **(D)**, and mOPH-GFP-SsrA **(E)**. **(F)** Presence of OPH-GFP fusion proteins in *B. diminuta* cells having expression plasmids coding for preOPH-GFP (lane 1), preOPH-GFP-SsrA (lane 2), mOPH-GFP (lane 3), and mOPH-GFP-SsrA (lane 4). Total cell proteins from the respective cultures were analyzed by SDS-PAGE, and Western blots were performed using anti-GFP antibodies.

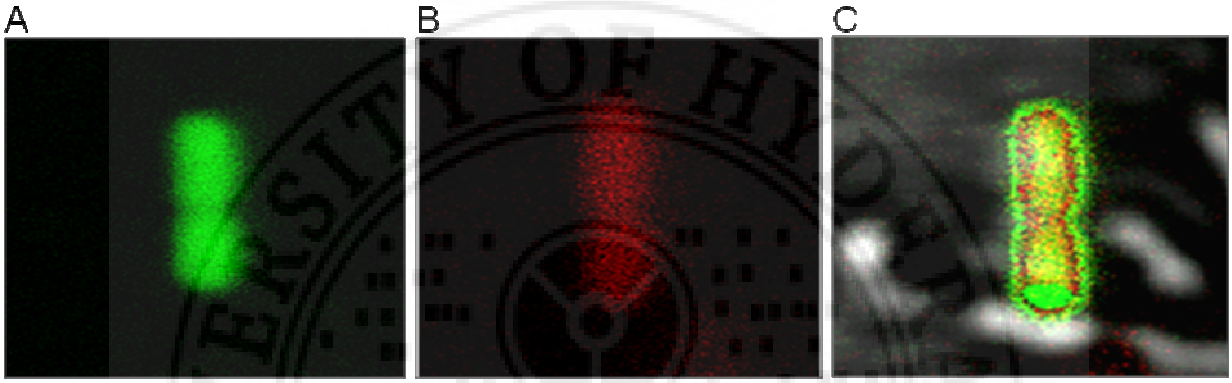
Such reduction in mean fluorescence in cells containing mOPH-GFPSsrA is suggestive of ClpXP mediated degradation of the fusion protein, obviously due to its accumulation in cytoplasm. In cells expressing preOPH-GFPSsrA, due to presence of signal peptide, the fusion protein crossed the membrane before getting degraded by the cytoplasmically located ClpXP proteolytic machinery, which is why there was no significant reduction in mean fluorescence of the cells expressing preOPH-GFPSsrA. In fact, this result rather complements the data obtained from proteinase K treated spheroplasts of *B. diminuta* (pSM5). The proteinase K eliminated much of the membrane associated OPH as it was found on the periplasmic face of the inner membrane. Instead, due to OPH signal peptide mediated insertion at the periplasmic face of the membrane, the GFP retained its fluorescence despite of having SsrA signal at the C-terminus. Even the western blot results generated using anti-GFP antibody supported this claim. The GFP specific signal was unaffected in cells expressing preOPH-GFP and preOPH-GFPSsrA (Fig. 2.6. F, Lanes 1 and 2). However, a weak GFP specific signal (Fig. 2.6 F, Lane 4) indicates elimination of mOPH-GFPSsrA from the cells and support reduced fluorescence in cells expressing mOPH-GFPSsrA.

**2. 5. Confocal fluorescence microscopy**

To reinforce the data obtained through proteinase K treatment and OPH-GFPSsrA reporter system, fluorescence images were obtained for cells expressing preOPH-GFP in *B. diminuta. B. diminuta* cells carrying preOPH-GFP were incubated with FM4-64 (N-(3-triethylammoniumpropyl)-4-(p-diethylaminophenylhexatrienyl)-pyridinium 2Br), a membrane specific dye (Feilmeier et al., 2000) before taking confocal microscopic images (Fig. 2.7). A brightened fluorescent image obtained (Fig. 2.7) was then overlaid on identical size image to

detect the localization of green fluorescence emitted due to presence of OPH-GFP. The image clearly shows the presence of green fluorescence at the periphery of the cell in induced cultures expressing OPH-GFP (Fig. 2.7 C). This is rather a direct evidence to show location OPH at the periplasmic face of the inner membrane in *B. diminuta.*

**Fig. 2.7.**




Fluorescence microscopy localization of preOPH-GFP. **(A)** Fluorescence microscopic image of *B. diminuta* cells expressing preOPH-GFP. **(B)** Image taken after staining the cells with the membrane-specific dye FM4-64. **(C)** Overlay of images shown in panels A and B.

## 2. 6. Discussion

The penetration of a protein into a membrane is one of the most basic cellular localization processes, which is recognized by a translocation machinery in a cell. Determination of the topology of a particular site on a membrane protein is an important step in elucidating its structure. Even if the structure of a membrane protein is known, topology measurements is useful for verifying its correct fold and orientation in the membrane (Ladokhin et al., 2002). Measurements of topology have been especially important in the determination of membrane insertion pathways for a variety of spontaneously inserting nonconstitutive proteins, such as bacterial toxins (Ladokhin et al., 2002), colicins (Parker et al., 1990; Tory and Merrill, 1999),

and annexins (Isas et al., 2000). Based on functional importance of membrane proteins, several groups have investigated the common features of membrane proteins as well as transmembrane regions (Eisenberg et al., 1984) and developed software systems for the prediction of transmembrane domains (Rost et al., 1996; Cserzo et al., 1997; Sonnhammer et al., 1998; Tusnady et al., 2001).

Various biochemical and molecular techniques were used for elucidation of the membrane topology of integral membrane proteins. These include introduction of antibody epitopes and generation of reporter gene fusions to N-terminal or C-terminal regions (Marleen et al., 2000). PhoA, LacZ and β–lactamase fusions have been highly useful for elucidating membrane topology (Marleen et al., 2000). However, the unique mechanism required for the process of folding and assembly of PhoA is considered as one of the serious limitations in using it as reporter system for determination of topology of Tat substrates. In contrast to PhoA, LacZ (β-galactosidase) exhibits enzymatic activity in the cytoplasm (Marleen et al., 2000). When LacZ was attached downstream of an export signal, the enzyme is trapped in the membrane, preventing proper folding of the enzyme to gain active conformation (Silhavy et al., 1977). Thus, LacZ fusions destined to a periplasmic site of a membrane protein are inactive, while fusions to a cytoplasmic domain are active. Therefore, the LacZ reporter system cannot be used to determine a complete membrane topology of an integral membrane protein (Marleen et al., 2000). The third commonly used reporter molecule, which has been used as an alternative to PhoA and LacZ, is the mature form of the monomeric periplasmic enzyme β-lactamase, encoded by the *bla* gene (Broome-Smith et al., 1990). Periplasmic β-lactamase protects cells against lysis by β-lactam antibiotics such as ampicillin, which otherwise inactivate enzymes essential for cell wall biosynthesis that are anchored to the outer surface of the cytoplasmic membrane. Cells

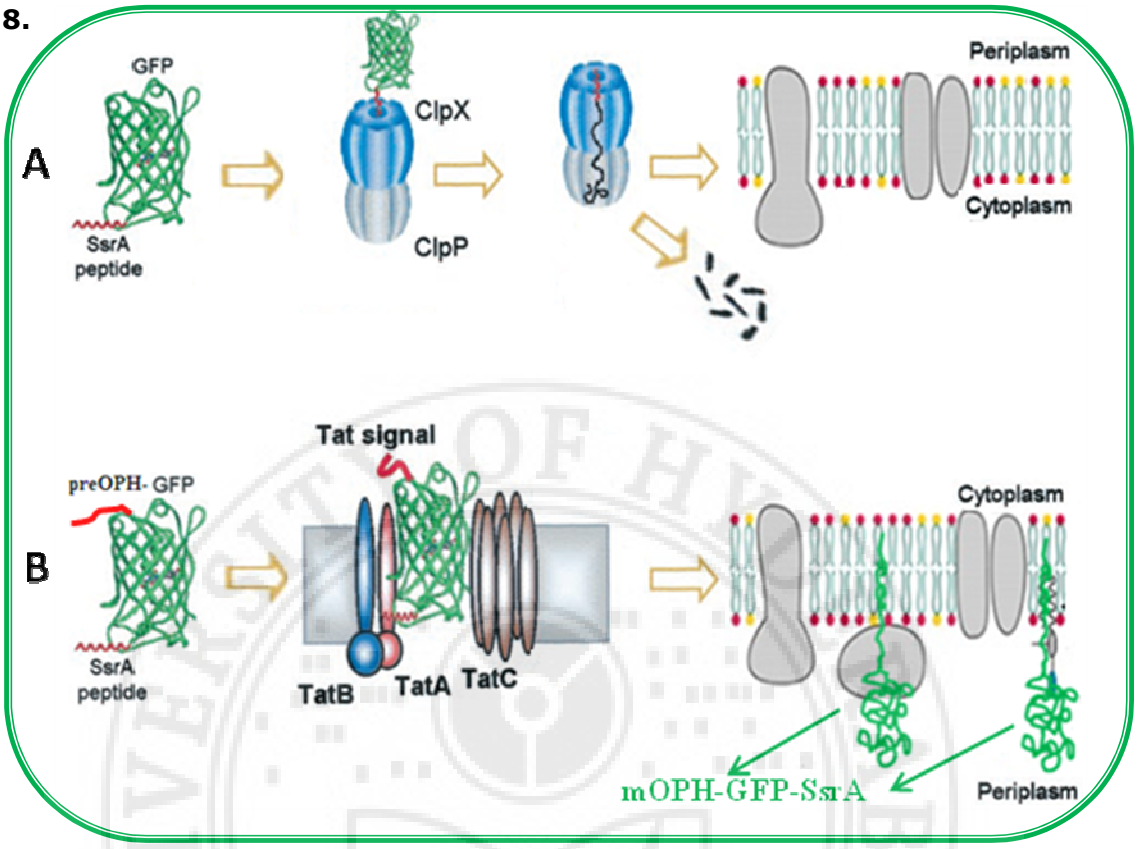
expressing cytoplasmic β-lactamase are sensitive to the antibiotics because cytoplasmic β-lactamase cannot inactivate the antibiotic found in periplasmic space. Thus, cells expressing fusion proteins in which the mature form of β-lactamase is fused to a cytoplasmic domain of a membrane protein fail to grow on plates with a high ampicillin concentration while cells expressing fusions to periplasmic domains are able to grow on these plates (Broome-Smith and Spratt, 1986).

In contrast to the above reporter fusions, a recent discovery in using green fluorescence protein as a reporter system has been widely used and it has been shown to overcome the limitiations in determining the membrane topology of integral membrane proteins (Stuart et al., 2002). Especially, for determining the topology of proteins that dependent on Tat machinery for membrane targeting, the GFP is found to be one of the favorite reporter systems (Santini et al., 2000; Barret et al., 2003). The 29 kDa green fluorescent protein (GFP) which provides unique environment for three residues in its primary sequence to act as a fluorophore (Chalfie and Kain, 1998). The GFP chromophore consists of a cyclic tripeptide derived from Ser-Tyr-Gly (amino acids 65–67) in the primary protein sequence (Cody et al., 1993) and is only fluorescent when embedded within the complete GFP protein. Of the complete 238-amino-acid polypeptide, amino acids 7–229 are required for fluorescence (Li et al., 1997). Nascent GFP is not fluorescent, since chromophore formation occurs post-translationally. The chromophore is formed by a cyclization reaction and an oxidation step that requires molecular oxygen (Reid and Flynn, 1997). Chromophore formation may be the rate-limiting step for maturation of the fluorescent protein (Brain and Gregory, 1997). Detection of GFP in living cells can be performed and is amenable to real-time analysis of molecular events (Westphal et al., 1997). Use of GFP also eliminates the need for fixation or cell permeabilization (Chalfie et al., 1994). In several gram negative

bacterial systems, GFP was shown to be efficiently translocated by the bacterial TAT machinery across the cytoplasmic membrane into the periplasmic space. When fused to the TAT-specific targeting signal peptide domain from trimethylamine N-oxide (TMAO) reductase (TorA) the GFP got successfully translocated across the membrane and retained its characteristic fluorescence (Santini et al., 2001; Barret et al., 2003). However, upon fusion of GFP to the maltose binding protein, a substrate of the Sec pathway, only insignificant amounts of GFP could be transported into the periplasm of *E. coli*. Further, the protein secreted by the Sec pathway did not emit fluorescence and was particularly prone to proteolytic degradation, suggesting a misfolded conformation (Feilmeier et al., 2000). It is therefore evident that GFP retains its fluorescence only when transported to the periplasm via twin arginine translocation pathway (Claire et al., 2003).

In addition to versatile applications of GFP, there exists another system which cellular proteases use for the removal of non-functional or damaged proteins which serve no functional benefit to the cells. One of the best studied system is SsrA proteolysis of incompletely translated proteins (Wickner et al., 1994). They naturally stop translation of proteins upon encoding a stop codon. In addition to natural phenomenon, the ribosomes stall translation process at broken ends of mRNA or when a rare codon for which no cognate tRNA is available in the cells. Such abrupt termination of translation generate truncated proteins. Addition of an SsrA signal comprising amino acid sequence AANDENYALAA serves as an identification mark for incomplete translated proteins and quickly degraded by an intracellular ClpXP and ClpAP proteolytic machinery (Gottesman et al., 1998; Bohn et al., 2002). Delisa and coworkers have used SsrA system containing GFP for demonstration of periplasmic location of Tor (Delisa et al., 2002).

**Fig. 2.8.**




Diagrammatic representation of SsrA mediated degradation of incompletely translated proteins **(A)** and its usage for demonstrating OPH-GFP-SsrA association to the inner membrane in *B. diminuta* **(B)**.

The present study used three independent strategies for determining membrane topology of OPH in *B. diminuta*. The first strategy depended mainly on degradation of surface expressed proteins by proteinase K. In this strategy, the spheroplasts prepared from the *B. diminuta* (pSM5) cells were treated with proteinase K and subsequently the existence of OPH was monitored by detecting OPH through western blots and by measuring the activity of OPH. In both the cases the study could prove that OPH is associating with the membrane facing the periplasmic space. The second approach was by showing the Tat-transport specific fluorescence of *B. diminuta* cells. The *B. diminuta* cells expressing preOPH-GFP have successfully retained fluorescence suggesting that OPH is targeted to the membrane via Tat pathway. Such

translocated protein was further found on membrane as seen by analyzing the images generated for normal fluorescent cells and cells that were treated with membrane specific dye.

The third experiment is in fact a combination of Tat-specific fluorescence and SsrA mediated proteolysis. This was measured by FACS analysis. The preOPH-GFPSsrA successfully escaped SsrA dependent proteolysis suggesting that the preOPH-GFPSsrA has successfully translocated across the membrane before getting elevated by the cytoplasmically located proteolytic machinery. In case of mOPH-GFPSsrA, there was no such scope and hence most of it got degraded loosing associated fluorescence. All these independent studies have unanimously pointed towards existence of OPH on the periplasmic face of the membrane. Based on the results presented in the chapter II and the results discussed in previous chapter, it can be convincingly stated that OPH is a Tat substrate and localized on the periplasmic face of the inner membrane of *B. diminuta.*

Having shown the membrane association of OPH, the next question that remains to be answered is how is it associating with the membrane. Existing literature suggests very well established ways of proteins anchoring the membrane (Marleen and Juke, 2000). The first one is the role of cysteine residue and its applications in determining membrane anchoring via sulfhydril reagents, and the second important one is the presence of transmembrane domains. The cysteine residues present in the extracellular or intracellular loops of a membrane protein are highly useful in determining orientation of an integral membrane protein (Marleen and Juke, 2000). This method could be analyzed by membrane-permeable and -impermeable sulfhydryl reagents which contain either a biotin group (Loo and Clarke, 1995), a fluorescent group (Zhou et al., 1995), or a radiolabel (Kimura et al., 1997), allowing detection of labelled proteins by

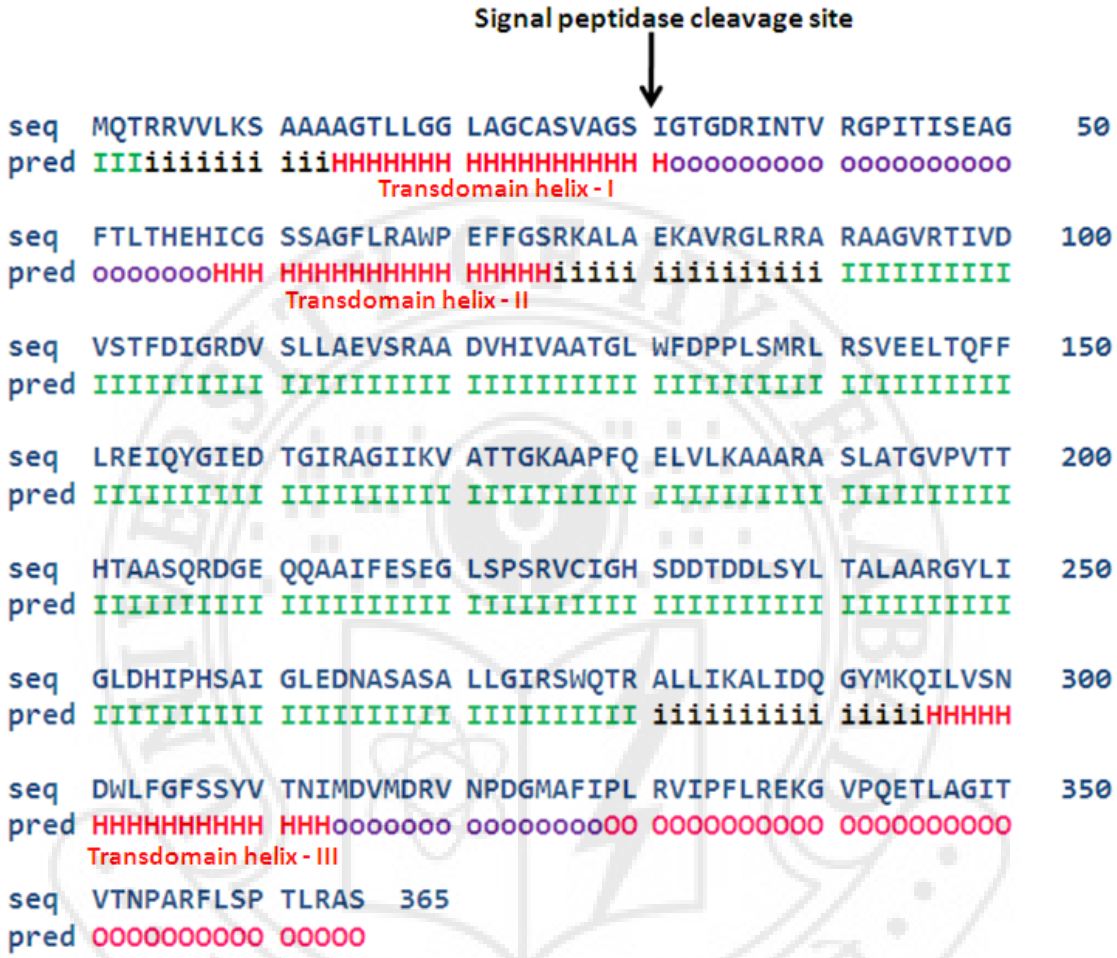
streptavidin binding, fluorescence, or autoradiography, respectively. The labelling of a cysteine residue with membrane-permeable or impermeable reagents is an indicative of a cytoplasmically located cysteine residue, while the absence of labelling reveals a periplasmic location (Marleen and Juke, 2000). In the case of OPH there exists no cysteine residues which thought to play in membrane anchoring and studies continued to look for the existence of transmembrane domains.

The integral membrane proteins are divided into two distinct classes based on the structural difference in transmembrane domains. The class I proteins grouped based on the structure of the transmembrane spanning segments which form an α-helical structure with lengths of 17 to 25 amino acid residues (von Heijne, 1994). The class II proteins are only known in the bacterial outer porins that have a 16-stranded β-barrel structure (Marleen et al., 2000). All these membrane anchoring signals are well characterized and well defined systematic tools are also available for predicting such signals in any protein molecules. Membrane protein topology predictions are generally based on the overall hydrophobicity of transmembrane α-helices or the charge distribution of the hydrophilic loops that connect the transmembrane segments. The hydrophilic loops follow the "positive inside" rule, which states that nontranslocated loops are enriched in positively charged residues compared to translocated loops (von Heijne, 1992). The first observation is used to identify the transmembrane segments in the amino acid sequence by analyzing the hydropathic properties of the amino acid sequence (von Heijne, 1992; Claros et al., 1994), and the second observation is used to predict the overall orientation of the protein in the membrane (Marleen et al., 2000).

In the present study one of the best ranking membrane helices prediction method HMMTOP2.0 (Tusnady et al., 2001) was used to predict transmembrane helices in OPH.

Interestingly analysis of OPH sequence using the HMMTOP2.0 software, indicated existence of three putative transmembrane helices in OPH (Fig. 2.9).

**Fig. 2.9**



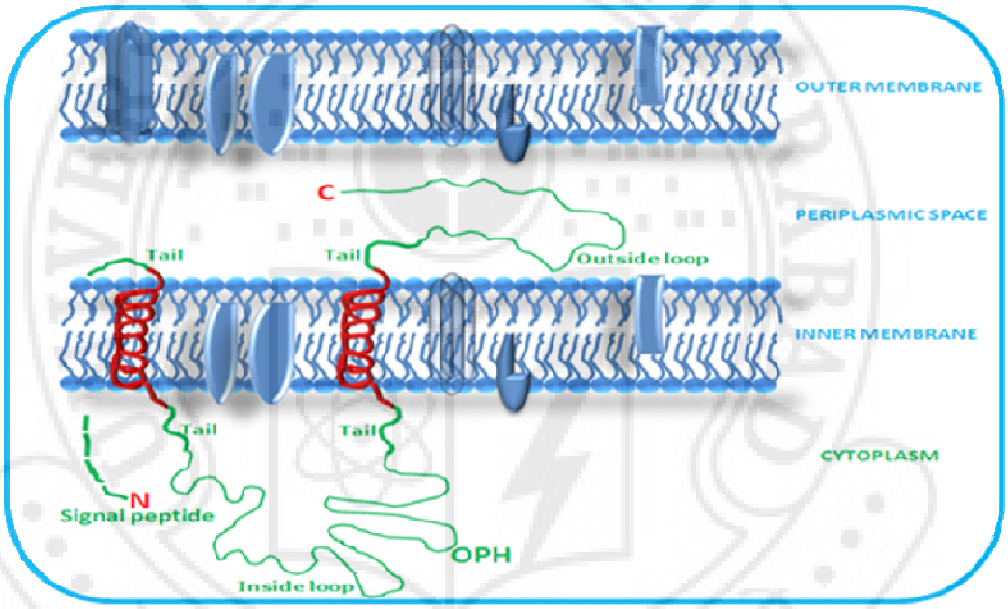

Structural states defined for OPH having a putative typical helical transmembrane feature. The five states are: inside loop (I), inside tail (i), membrane helix (h), outside tail (o) and outside loop (O).

The first transmembrane domain (14-31) has no significance with respect to the membrane anchoring as most of it (up to 29$^{th}$ amino acid) is available in the signal peptide which is cleaved after membrane targeting. As shown in predictions, the second transmembrane domain is present between amino acid residues 58-75. If this region is seen together with other structural states predicted by the programme, in OPH there exists a bit cytoplasmic loop (I) before having a third transmembrane domain (H) between amino acid residues 296-313. After

the third transmembrane domain C-terminal periplasmic tail (o) and loop (O) is predicted. This study has clearly shown existence of OPH on the periplasmic face of inner membrane.

The western blots were done to detect C-terminal his tag and the SsrA signal was kept at the C-terminus of the protein. These two independent studies have shown that the protein is facing periplasmic space of inner membrane, which is in agreement with the predictions made in the study. Based on the results gained from this study a model proposed on membrane anchoring of OPH is shown in Fig. 2. 10.

**Fig. 2. 10.**




Proposed model indicating the topology of OPH in the inner membrane in *Brevundimonas diminuta* showing mOPH with C-terminal outloop facing towards periplasmic space

# RESULTS & DISCUSSION

The work described in this chapter is initiated as a consequence of a novel observation made during heterologous expression of *opd* gene. When expression plasmids were constructed to express *opd* gene to code organophosphate hydrolase (OPH) with and without signal peptide, the amount of protein coded by the plasmid expressing mOPH (without signal peptide) was several folds high when compared to the expression levels obtained from plasmid coding for preOPH (OPH with signal peptide). Such unusual increase in expression level, despite of having identical vector backbone indicates existence of regulatory elements in the *opd* region specifying signal sequence (Fig. 3. 1). Therefore in the present study a detailed investigation is made to understand the regulation of *opd* gene expression. Before describing experimental strategies for studying the regulation of *opd* gene expression, a thorough bioinformatic analysis was made to identify regulatory elements that contribute for expression of *opd* gene in *B. diminuta.* Using Bioinformatic tools developed by Munch et al., 2005 (www.prodoric.de/vfp) two putative promoter elements were identified 33bp upstream of start codon ATG. One of them has shown resemblance to the classical sigma 70 ($\sigma^{70}$) dependent promoters (Fig. 3. 1). The second one has shown identity to $\sigma^{54}$ dependent promoter element. Though existence of these two promoter elements was predicted by earlier investigators their functional status was not examined till to date (Harper et al., 1988; Mulbry and Karns, 1989). Further, on careful examination we have also identified an inverted repeat sequence at 32 bp downstream of start codon ATG. This IR sequence is found within the signal peptide coding sequence of *opd* gene. Removal of signal peptide coding sequence has resulted in elevated expression levels of OPH. Therefore, the IR sequence is expected to play a role in downregulation of *opd* gene expression. This chapter describes the experiments designed to assess the functional status of the predicted regulatory

elements. Based on the inference drawn from the experiments a logical conclusion is made explaining the regulation of *opd* gene expression in *B. diminuta.*

**Fig. 3. 1**


Identification of putative promoter elements upstream of start codon ATG of *opd* gene. The putative $\sigma^{70}$ and $\sigma^{54}$ dependent promoters are highlighted with red and pink colors. The putative ribosomal biding site is shown with dotted lines. Important restriction sites used to generate *opd-lacZ* fusions are mapped. The inverted repeat sequence found 32 bp downstream of start codon ATG is shown with bold case.

### 3. 1. Prediction of putative promoter elements of *opd* gene

The sequence of *opd* gene was analyzed using motif search programme (www.prodoric.de/vfp). The search revealed existence of two putative promoter elements upstream of start codon ATG. The conserved hexamers of one of the promoters have shown strong similarity to the consensus $\sigma^{70}$ dependent promoter. To gain better understanding, the predicted promoter motif was aligned to the consensus $\sigma^{70}$, $\sigma^{32}$, $\sigma^{E}$ and $\sigma^{S}$ dependent promoters. As shown in Table. 3. 1 the putative promoter motif has shown more resemblance to the consensus $\sigma^{70}$ motif. The -35 hexameric sequence has shown more conservation than the hexameric sequence found at -10 region. Downstream of this promoter motif, existence of

dinucleotides GG and GC with a gap of 11 bp, points towards existence of putative $\sigma^{54}$ dependent promoter element upstream of the start codon of *opd* gene. Further, to validate the functional status of these putative promoters, they were independently cloned upstream of the promoter less *lacZ* gene of the promoter probe vector pMP220 (Spaink et al., 1987). The *opd-lacZ* fusions thus constructed were mobilized in *B. diminuta* and promoter activity was assessed through β-galactosidase activity.

**Table. 3. 1**

| σ Factor | Promoter consensus sequence | | Space between -35 and -10 / -24 and -12 regions | Reference |
|---|---|---|---|---|
| | -35 | -10 | | |
| *B. diminuta* *opd* putative $\sigma^{70}$ | TTGACA | TAAAAG | 16 | This work |
| $\sigma^{70}$ | TTGACA | TATAAT | 16 ± 17 | Gruber and Gross, 2003. |
| $\sigma^{S}$ | TTGACA | TATAAT | 17 | Typas and Hengge, 2006. |
| $\sigma^{32}$ | NCTTGAA | NCCCCATNT | 17 | Wang and deHaseth, 2003. |
| $\sigma^{28}$ | CTAA | CCGATAT | 11 ± 12 | Yu et al., 2006. |
| | -24 | -10 | | |
| *B. diminuta* *opd* -putative $\sigma^{54}$ | CGGC | TGCA | 11 | This work |
| $\sigma^{54}$ | TGGC | TGCN | 10 | Barrios et al., 1999. |

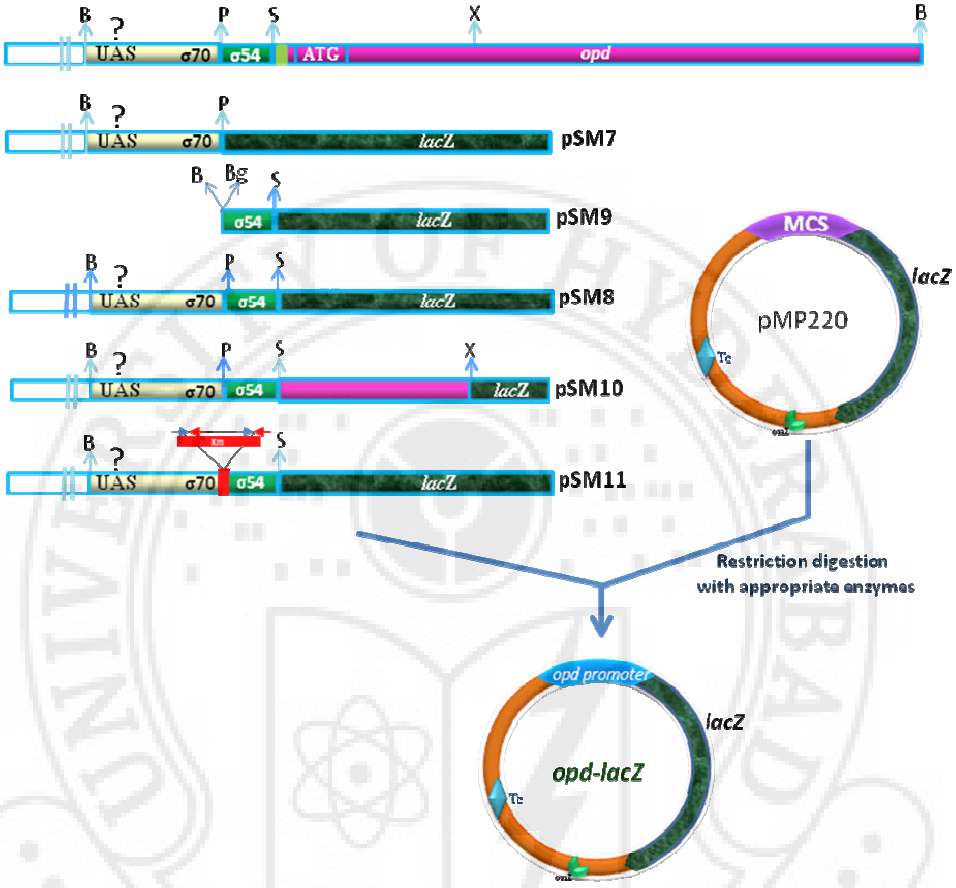
Comparison of putative $\sigma^{70}$ and $\sigma^{54}$ dependent promoter sequences of *opd* gene of *B. diminuta* with consensus $\sigma^{70}$, $\sigma^{S}$, $\sigma^{32}$, $\sigma^{28}$ $\sigma^{54}$ promoters of *E. coli.*

### 3. 2. Construction of *opd-LacZ* fusions

In order to determine functional status of putative *opd* promoter element predicted upstream of the start codon of *opd* gene, a series of promoter *lacZ* fusions were generated by cloning upstream region of *opd* gene in a broad host range mobilizable vector pMP220 (Spaink et al., 1987). The *opd* gene cloned from *B. diminuta* was taken as a source plasmid and the

5'region of *opd* gene was taken as various restriction fragments and cloned in pMP220. A detailed cloning strategy used to generate *opd-lacZ* fusions is shown in Fig. 3. 2.

**Fig. 3. 2**




A diagrammatic representation showing generation of *opd-lacZ* fusions. Plasmids pSM9 and pSM8 contain either $\sigma^{70}$ or $\sigma^{54}$ dependent promoters. Both $\sigma^{70}$ and $\sigma^{54}$ dependent promoters are fused to *lacZ* gene in plasmid pSM8. The IR sequence along with $\sigma^{70}$ and $\sigma^{54}$ promoters are fused to *lacZ* gene in pSM10. In pSM11 an omega (Ω) fragment is inserted between putative $\sigma^{70}$ and $\sigma^{54}$ dependent promoters. Abbreviations B, P, S, and X represent recognition sequences for restriction endonucleases *Bam*HI, *Bgl*II, *Pst*I, *Sph*I, and *Xba*I, respectively.

### 3. 3. Mobilization of *opd-LacZ* fusions and assay of promoter activity in *B. diminuta*

The *opd-LacZ* fusions generated were transformed into *E. coli* S17-1 and were used as helper strains to mobilize them into *B. diminuta* by following conjugation protocols described in materials and methods section. *E. coli* S17-1 cells containing *opd-LacZ* fusions were used as

donors while *B. diminuta* served as recipient. The exconjugants were selected on LB plates containing polymixinB and tetracycline.

The promoter activity for various *opd-lacZ* fusions was determined by measuring *β*-galactosidase activity (Miller, 1972) and the values obtained are shown in Table. 3. 2. A significant *β*-galactosidase activity (2040 miller units) was found in *B. diminuta* cells containing $\sigma^{70}$ dependant promoter-*lacZ* fusion (pSM7) indicating the predicted $\sigma^{70}$ dependant promoter is functional and responsible for transcription of *opd* gene in *B. diminuta.* In an attempt to assess the functional status of predicted $\sigma^{54}$ dependent promoter, an *opd-lacZ* fusion was constructed (pSM9) by including only $\sigma^{54}$ dependent promoter (Fig. 3. 2). When *B. diminuta* cells containing pSM9 were used for monitoring LacZ activity, only a negligible amount of LacZ activity is seen (Table. 3. 2). Such reduction in β-galactosidase activity points towards existence of a sequence motif that has no functional significance. However in literature, the dependency of $\sigma^{54}$ on upstream activating sequences (UAS) is well documented (Su et al., 1990). The UAS is normally present several base pairs upstream of the promoter element. In plasmid pSM9, the *opd* region cloned upstream of *lacZ* gene does not represent much of the upstream sequence of *opd* gene. In this construct there is no scope for existence of UAS. In order to avoid such situation another *opd-lacZ* fusion (pSM8) was generated by including the entire upstream region of *opd* gene. Hence in plasmid pSM8, an unique *Pst*I site is there in between $\sigma^{70}$ and $\sigma^{54}$ dependent promoters. This unique *Pst*I site was exploited to introduce an omega (Ω) fragment separating these putative promoters. The omega fragment is designed by introducing transcription terminator sequences flanking to the kanamycin gene. Existence of such transcription terminator sequence blocks transcription for $\sigma^{70}$ dependent promoter and hence the LacZ activation will be seen only when transcription initiation is made from predicted $\sigma^{54}$ dependent promoter. Further this situation

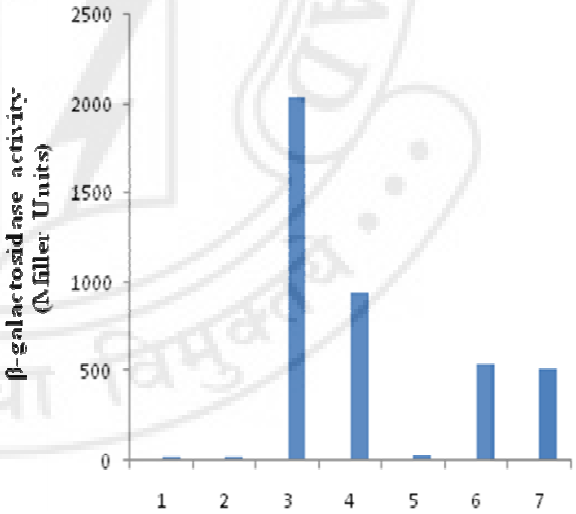
does not deprive promoter-UAS interactions, if there exists one upstream of putative $\sigma^{54}$ dependent promoter. Surprisingly, in *B. diminuta* cells, as shown in Fig. 3. 3, the *lacZ* fusions having both $\sigma^{70}$ and $\sigma^{54}$ dependent promoters (pSM8) and the construct generated by introducing an omega fragment between putative $\sigma^{70}$ and $\sigma^{54}$ promoters (pSM11) have shown comparable promoter activities. Though the activity levels were very low when compared to the activity levels obtained for a construct (pSM7) having only $\sigma^{70}$ dependent promoter (2040 miller units), the functional status of $\sigma^{54}$ cannot be ignored. The $\sigma^{54}$ dependent genes are induced in response to a number of physiological conditions. Unless suitable physiological conditions are created maximal promoter activity is not achieved from these promoters (Reitzer and Schneider., 2001). Though such physiological condition is yet to be identified, usage of $\sigma^{54}$ dependent promoter for expression of *opd* gene in *B. diminuta* cannot be ruled out.

**Table. 3. 2**

| S.NO. | *B. diminuta* containing *opd-lacZ* fusion | *β*-galactosidase activity (Miller Units) |
|---|---|---|
| 1. | *B. diminuta* | 19.4 |
| 2. | *B. diminuta* (pMP220) | 16.6 |
| 3. | *B. diminuta* (pSM7) | 2040.0 |
| 4. | *B. diminuta* (pSM8) | 943.61 |
| 5. | *B. diminuta* (pSM9) | 30 |
| 6. | *B. diminuta* (pSM10) | 549.44 |
| 7. | *B. diminuta* (pSM11) | 514.44 |

β-galactosidase assay for *B. diminuta* cells containing *opd-lacZ*fusions.

**Fig. 3. 3**

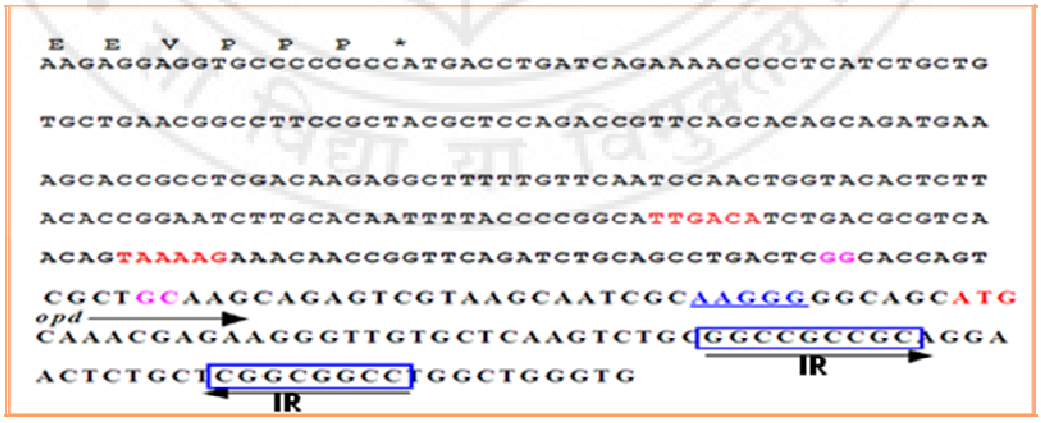



β-galactosiadase activity of *opd-lacZ* fusions. Bar 1 and 2 represent *B. diminuta* without and with pMP220 vector and serves as negative controls. Bar 3, 4, 5, 6 and 7 represent levels of β-galactosiadase activity obtained for *B. diminuta* cells having *opd-lacZ* fusions, pSM7, pSM8, pSM9, pSM10 and pSM11 respectively.

### 3. 4. Existence of IR sequences in the signal peptide coding region of *opd* gene

The *β*-galactosidase activities obtained for various *opd-lacZ* fusions pSM7, pSM10, besides indicating the functional status of putative promoter elements, are suggestive of existing regulatory sequences in the coding region of *opd* gene. In the *opd-lacZ* fusion (pSM10) constructed by including the downstream region (downstream of start codon ATG), especially the signal peptide coding region, there was a significant reduction in *β*-galactosidase activity (Table 3. 2). Such reduction in *β*-galactosidase activity is indeed supportive of elevated levels of OPH expression from expression plasmids coding for mOPH (mature form of OPH), where the signal peptide coding region of *opd* gene is omitted while construction of expression plasmid. Such observation prompted to examine for presence of regulatory sequence in this region. When the sequence found downstream of translational start site of *opd* gene was examined, a 22bp long inverted repeat (IR) was found in the *opd* region specifying signal peptide (Fig. 3. 4). Existence of such inverted repeat is expected to form a strong secondary structure at its 5' end of *opd* specific mRNA.

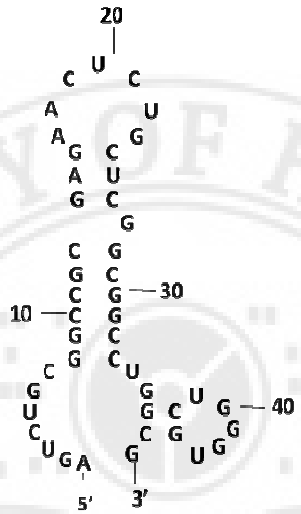
**Fig. 3. 4**




Nucleotide sequence of upstream region of *opd* gene showing the existence of IR sequences (indicated in blue colored closed boxes) located exactly 2 9 base pairs downstream of the initiation codon ATG.

Therefore, the *opd* mRNA sequence of *B. diminuta* was used (using mFOLD analysis program) to predict existence of a secondary structure (www.bioinfo.rpi.edu/applications/ mfold). As expected, existence of a stable stem loop structure requiring a free energy (ΔG) of -23.3 kCal / mol for destabilization was predicted (Fig. 3. 5).

**Fig. 3. 5**

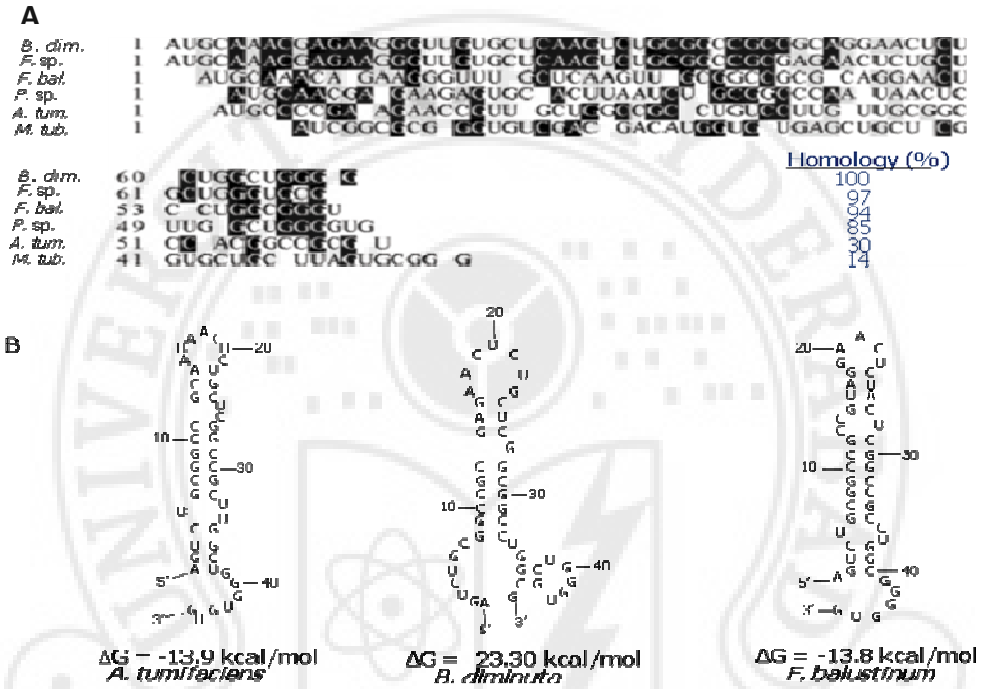



Secondary structure prediction and calculation of free energy requirement of *opd* mRNA using mFOLD programme (www.rna.ccbb.utexas.edu). (ΔG= -23.3 kCal/mol)

On perusal of literature, such stable stem loops, especially at the 5' end of mRNA, are shown to modulate the expression levels of cognate proteins (Punginelli et al., 2004). If the identified IR sequence has any regulatory role, it should have shown conservation among *opd* sequences found in other prokaryotic organisms. In order to gain information on existence of such IR sequences and associated secondary structures, the *opd* mRNA sequences of *B. diminuta* (Serdar et al., 1982), *Flavobacterium* sp. ATCC27551 (Sethunathan and Yoshida, 1973), *Flavobacterium balustinum* (Somara and Siddavattam, 1995), *Pseudomonas* sp. (Chaudhary and Wheeler, 1988) *Agrobacterium tumefaciens* (Horne et al., 2002) and *Mycobacterium tuberculosis* (Fleischmann et al., 2004) were downloaded from the data base and analyzed using mFOLD programme. In some of them the IR sequence was seen at an identical positions. The predicted secondary structures and ΔG values have shown close resemblance with the secondary

structure and ΔG value of *B. diminuta opd* mRNA (Fgi. 3. 6), suggesting that the structural conservation might be having functional significance. To gain further insights in to the IR sequence mediated modulation of OPH expression, a series of expression plasmids were constructed either by destabilizing or eliminating the IR sequence. Expression pattern obtained from these expression plasmids are presented below.

**Fig. 3. 6**



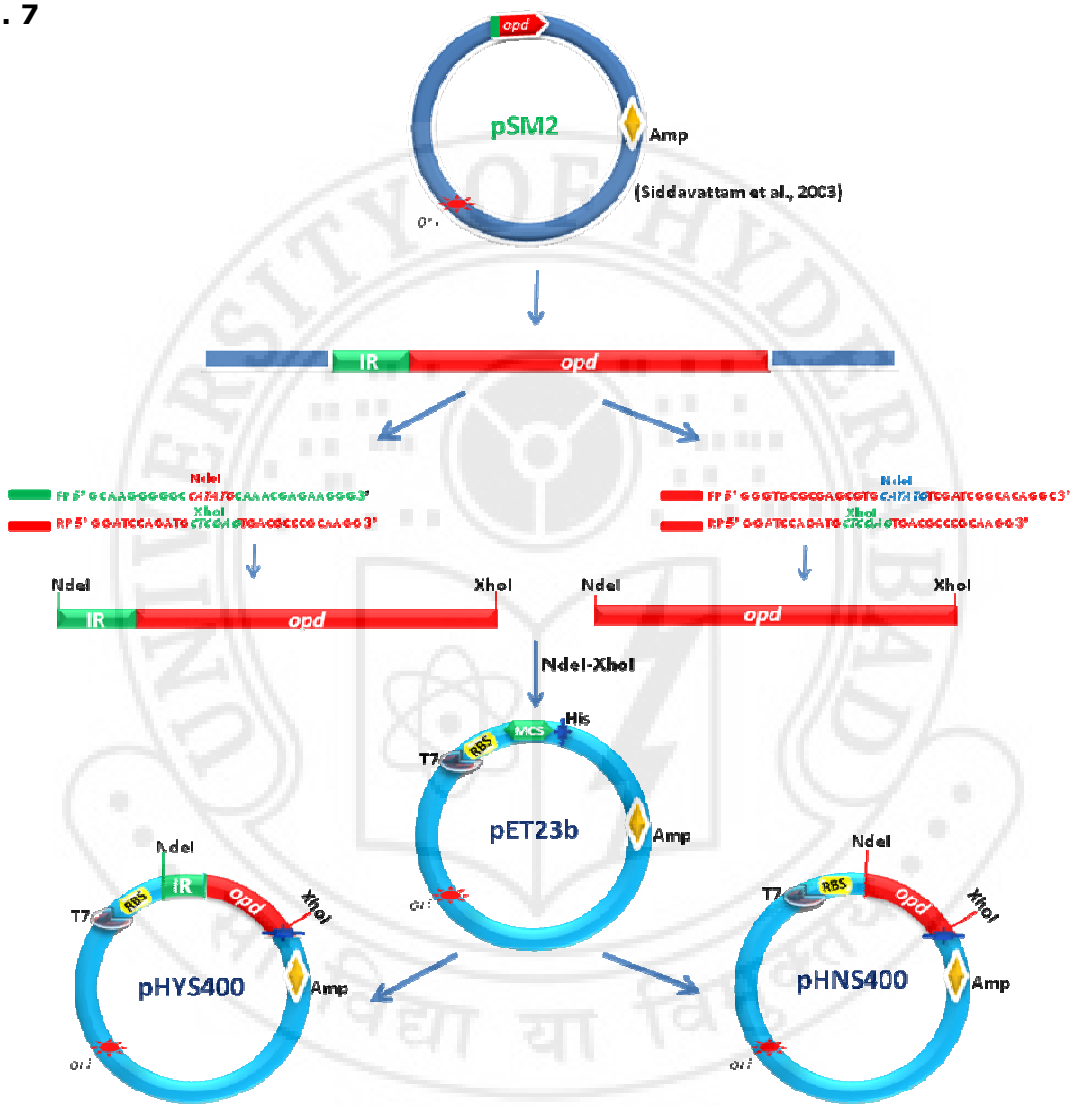

**Panel A** alignment of *opd* mRNA sequences from closely related bacteria. The percent homology is shown against each molecule. **Panel B** indicates the prediction of secondary structures found in *opd* mRNA sequences cloned from taxonomically diverse group of bacteria. *B. dim: B. diminuta; F. sp.: Flavobacterium sp. ATCC27551; F.bal: Flavobacterium balustinum; P.sp: Pliseomonas sp.; A.tum: Agrobacterium tumifaciens; M.tub: Mycobacterium tuberculosis.*

### 3. 5. Expression of OPH with and without IR sequence

To gain better insights on IR mediated modulation of OPH expression, the *opd* gene was truncated by performing PCR and its variants generated with and without IR sequence were cloned under the control of vector driven transcriptional and translational signals. One of the expression plasmid pHYS400, was generated by cloning the entire ORF of *opd* gene in pET23b as *Nde*I and *Xho*I fragment (Fig. 3.7). These sites were engineered through PCR mutagenesis

(Fig. 3. 7). The *Xho*I site was generated by modifying the stop codon of *opd* gene. Hence cloning of *opd* gene as *Nde*I-*Xho*I fragment generates inframe fusions of *opd* sequence with the vector specific sequence coding for six consecutive histidine residues.

**Fig. 3. 7**




A diagrammatic representation showing strategy used to construct expression plasmids coding *opd* gene with and without IR sequence in pET23b.

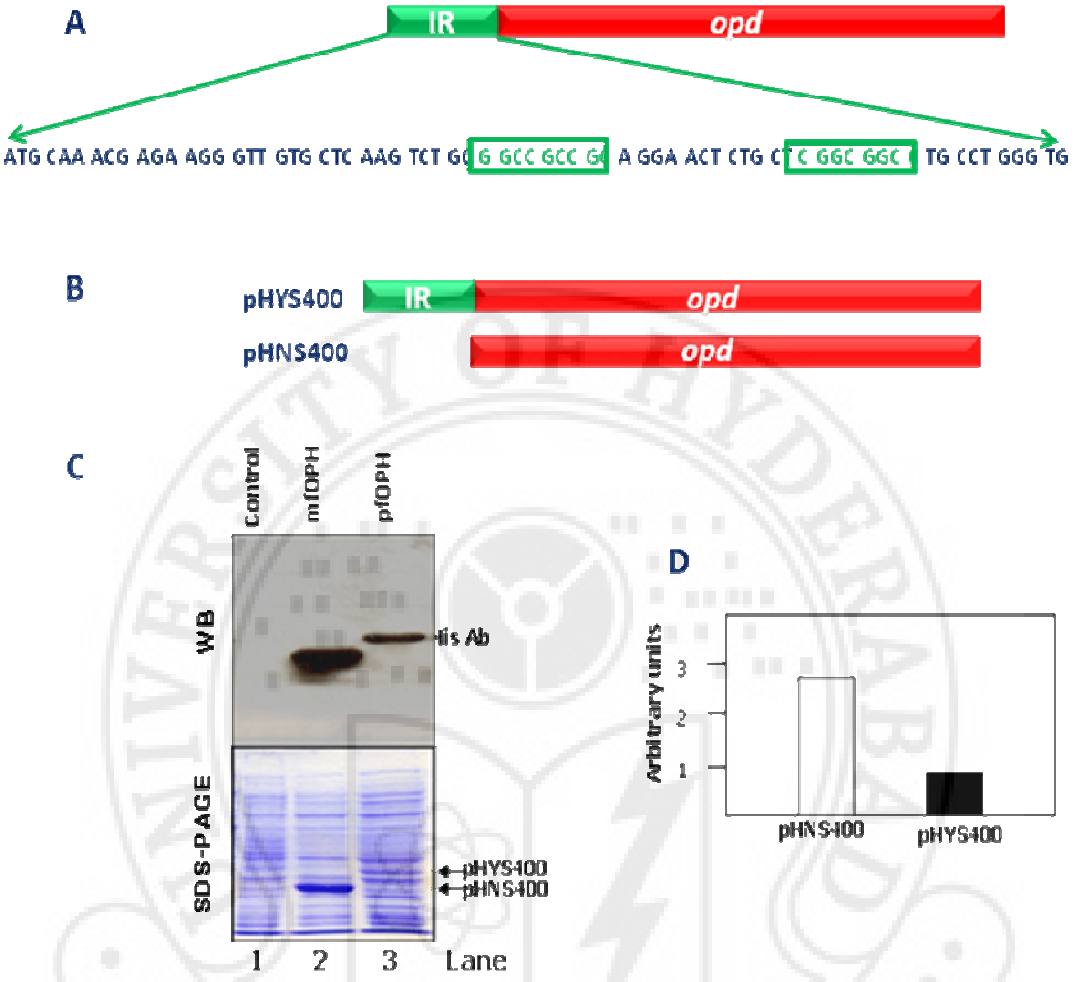
The second expression plasmid was generated by cloning the *opd* gene without IR sequence. While achieving this target an artificial start codon was generated downstream of the IR sequence through PCR mutagenesis and cloned in pET23b as *Nde*I and *Xho*I fragment. The

detailed cloning strategy is shown in Fig. 3. 7 and the plasmid having *opd* gene without IR sequence was designated as pHNS400. As OPH, coded by pHYS400 and pHNS400 contain his tag, they can be detected through western blots using anti-his antibodies.

### 3. 6. Assessment of expression levels coded by pHYS400 and pHNS400

The *E. coli* cells having expression plasmids were induced and the induced proteins were analyzed on SDS-PAGE. The protein gel, if carefully analyzed, indicates existence of a thick band in lanes where protein extracts prepared from *E. coli* cells containing pHNS400 are loaded (Fig. 3. 8 Panel C, Lane 2). The size of this protein corresponds to the mature form of OPH (mOPH). In Fig. 3. 8 Lane 3 of panel C, where protein extracts prepared from *E. coli* cells having pHYS400 are loaded, there is no such protein band. However a faint band corresponding to the precursor form of OPH (preOPH) is seen in this lane. Western blots were performed to prove if these additional protein bands really represent OPH. As shown in panel C, WB, these two bands gave positive signals when probed with anti-his antibodies. The densitometry performed for these two signals indicated about 2.5 fold increase in expression levels of OPH in cultures expressing OPH from the expression plasmid pHNS400. This is rather a clear indication to show that the *opd* region coding signal peptide contains a *cis* element that contributes for modulation of OPH expression. Though the predicted IR sequence is seen in this region, the above described experiments do not justify its role in expression of OPH. Therefore, further experiments are designed to know if IR sequence is really responsible for down regulation of *opd* gene expression.

**Fig. 3. 8**



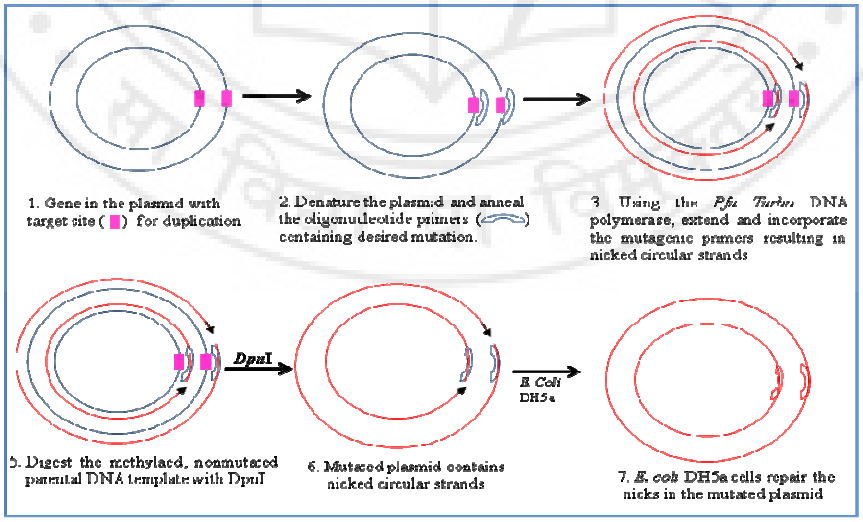

Expression of OPH with and without IR sequence in *E. coli* BL21 cells. **Panel A** shows existence of IR sequences in the signal peptide coding sequence of *opd* gene. **Panel B** shows extent of *opd* gene cloned in expression plasmids pHYS400 and pHNS400. **Panel C** indicates SDS-PAGE showing expression of OPH coded by pHNS400 (lane 2) and pHYS400 (lane 3). The corresponding western blot (WB) indicates expression levels of OPH coded by pHYS400 and pHNS400. **Panel D** shows the densitometry of OPH expression.

### 3. 7. Site directed mutagenesis and destabilization of secondary structure in *opd* mRNA

In order to generate *opd*' variant having no potential to form the secondary structure in the mRNA sequence, a PCR mutagenesis was performed. While performing the PCR mutagenesis, sense 5'GTG CTC AAG TCT GC**T** GC**A** GC**T** GCA GGA ACT CT**A** CT**A** GG**T** GG**A** CTG GCT GGG TG3' and anti-sense 5'CAC CCA GCC AGT CC**A** CC**T** AG**T** AGA

GTT CCT GC**A** GC**T** GC**A** GC**A** GAC TTG AGC AC3' primers were designed by incorporating changes only at the inverted repeat region. These changes were found exclusively at the third base of the codon so that the protein coded by *opd* variant *opd'* is identical to the OPH coded by the wild type *opd* gene. Further, a non-complementary base is used to change the third base of the codon, the transcript generated from *opd'* is expected to loose its ability to form a secondary structure. Therefore, the *opd* variant *opd'* generated in this manner is expected to code for an identical OPH without forming a secondary structure at the 5' end. Plasmid pHYS400 was used as template to perform PCR directed mutagenesis. The strategy used for performing site directed mutagenesis is shown in Fig. 3. 9. The generated PCR product was purified and treated with *Dpn*I to eliminate methylated parent plasmid used as template. The unmethylated mutated amplicon will be rescued from the *Dpn*I digestion. The amplicon was then transformed in to *E. coli* DH5α cells which repairs the nicks of plasmid and maintains its integrity inside the host system. The plasmid containing mutant *opd* gene was then isolated and used for confirmation of introduced mutations by sequencing analysis.

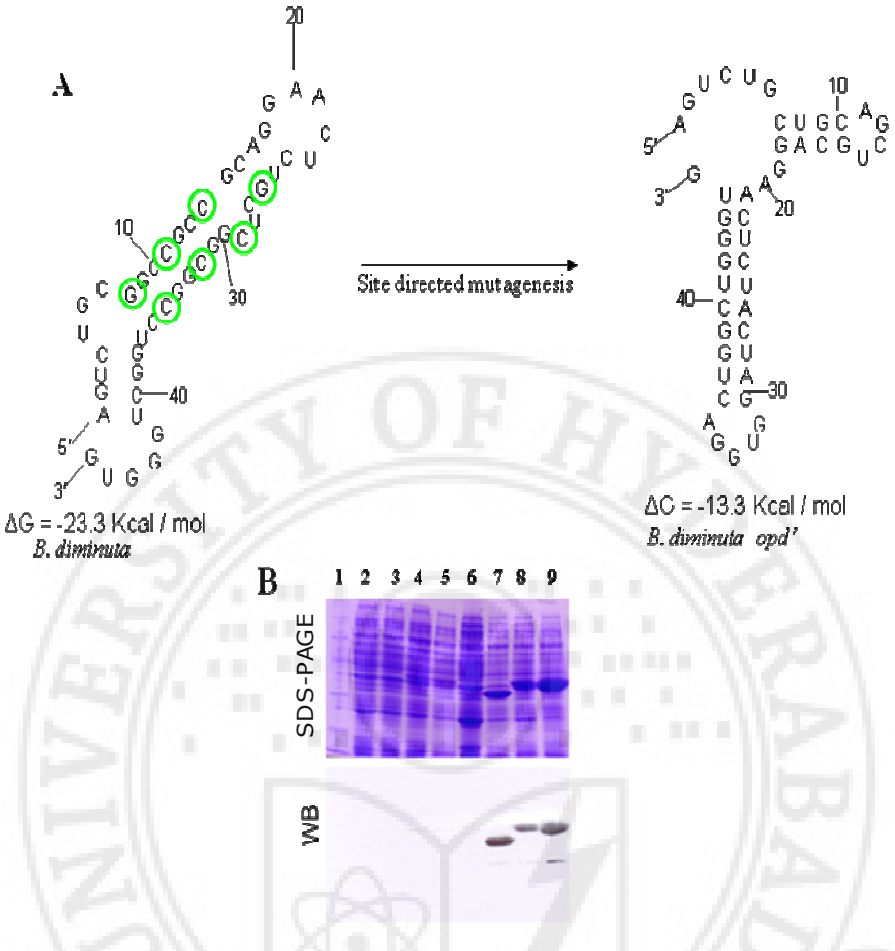
**Fig. 3. 9**




Schematic representation showing PCR mediated site directed mutagenesis performed to destabilize IR sequence induced secondary structure in OPH mRNA.

After confirming the mutations, the mRNA coded by *opd* variant *opd'* was subjected to mFOLD analysis programme to obtain secondary structure and ΔG values. Interestingly, as shown in Fig. 3. 10, the mRNA transcript of *opd* variant *opd'* showed no strong secondary structures. Even, the free energy (ΔG) calculated for destabilization of the weak secondary structure was just -13.3 kcal/mol. The expression plasmid containing the *opd* variant *opd'* was designated as pHIR400 and transformed in to *E. coli* BL21. The expression levels of OPH coded by pHYS400, pHNS400 and pHIR400 were compared both by running SDS-PAGE for induced cultures and by performing western blots. The results obtained are shown in Fig. 3. 10. Here, Lanes 7 and 9 are of having lot of significance. These two lanes contain protein extracts prepared from *E. coli* cells containing pHNS400 (lane 7) and pHIR400 (lane 9). As stated in earlier sections, the pHNS400 contains *opd* gene without inverted repeat sequence, whereas the plasmid pHIR400 has *opd* variant *opd'* where IR sequence is modified in such a way that the transcript made from it is no longer having potential to form stable secondary structure (Fig. 3. 10 panel A). If expression levels coded by these plasmids are seen, the *opd'* coded OPH (Fig. 3. 10, panel B, Lane 9) is equal to the OPH coded by pHNS400, the plasmid generated by removing signal peptide. Such observation is certainly an indication to show that IR sequence has a negative influence on the expression of OPH. While making such a bold statement, the results need to be supported through quantification of the *opd* transcript coded by all these plasmids. Such attempt was made by performing Real-Time PCR.

**Fig. 3. 10**



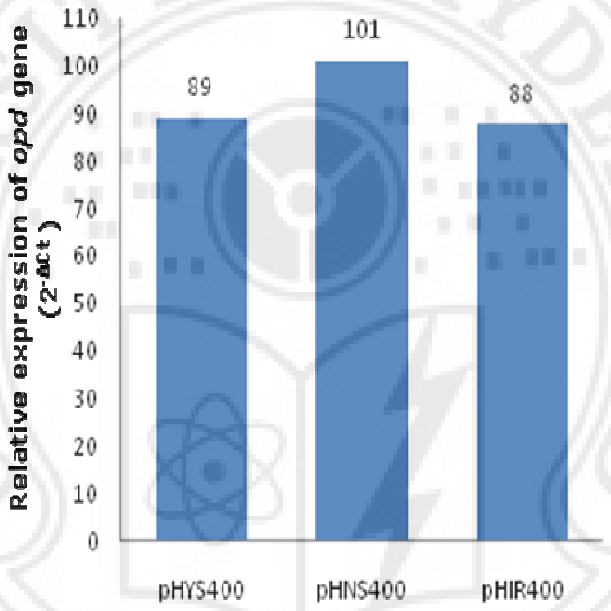

Expression of *opd'* variants in *E. coli*. **Panel A** shows the prediction of secondary structure and ΔG values of mRNA transcript made from *opd* variant *opd'*. The nucleotide changes are shown in green colored circles. **Panel B** SDS-PAGE and western blot analysis indicating expression profile of OPH coding from *opd* gene and its variants. Lane 1 represents Molecular weight marker; lane 2 represents protein extracts prepared from *E. coli* BL21 cells; lanes 3, 4, 5 and 6 represent protein extracts prepared from uninduced cultures of *E. coli* BL21 having pET23b, pHNS400, pHYS400 and pHIR400 and lanes 7 to 9 similar extracts prepared from induced cultures.

## 3. 8. Quantification of mRNA transcripts of *opd* variants

The OPH expression plasmids have identical vector backbone. These are all expressed using vector specific transcriptional and translational signals. Such a situation rules out involvement of transcriptional level regulation mediated by IR sequence. However to further justify this claim, a Real Time PCR was performed to quantify *opd* specific RNA from *E. coli* cells expressing OPH using expression plasmids pHYS400 (*opd* gene with IR sequence),

pHNS400 (*opd* gene without IR sequence) and pHIR400 (*opd* variant where IR sequence is destabilized). The quantity of *opd* specific mRNA as determined through RT-PCR is shown in Fig. 3. 11. As shown in Fig. 3. 11, the level of mRNA transcript of *opd* gene in all three cultures are almost identical and there exist no variation. Therefore, the observed elevation in levels of OPH expression in cultures having plasmids pHIR400 and pHNS400 is due to IR mediated translation modulation of *opd* mRNA.

**Fig. 3. 11**




Real Time PCR analysis. Histogram showing cyclic threshold values obtained for three independent experiments using total RNA isolated from *E. coli* cells containing expression plasmids pHYS400, pHNS400 and pHIR400. Bar lanes 1, 2 and 3 represent quantification of mRNA transcripts formed from expression plasmids pHYS400, pHNS400 and pHIR400 respectively.

### 3. 9. Discussion

The transcription of all eubacterial genes or operons is initiated by recognizing the promoter regions and its upstream regulatory elements by RNA polymerases (Morett and Segovia, 1993). The eubacterial RNA polymerase is a heterotetrameric enzyme comprising one

β, one β' and two α subunits (Helmann and Chamberl in, 1988). The 'core' enzyme interacts with the initiation factor σ to form the transcriptionally active enzyme (Eσ). The bacterial σ factors are divided into two different families. All the sigma factors that are evolutionarily related to the *E. coli* housekeeping factor $\sigma^{70}$ are kept in the first group. The sigma factors that show considerable similarity to the alternative sigma factor $\sigma^{54}$ are grouped in to second family (Gross et al., 1992). The $E\sigma^{70}$ does not form stable closed-promoter complex, therefore transcription can be initiated spontaneously in the absence of activator proteins. In contrast, the $E\sigma^{54}$ forms physically detectable closed-promoter complexes and fails to initiate transcription spontaneously (Gralla, 1990). Hence the $\sigma^{54}$ dependent RNA polymerase is completely dependent on additional transcriptional factors such as enhancer binding proteins (EBPs) to initiate the formation of mRNA transcript (Morett and Segovia, 1993). The $\sigma^{54}$ controls several auxillery processes in bacterial system, which include degradation of xylene and toluene, transport of dicarboxylic acids, pilin synthesis, nitrogen fixation, hydrogen uptake (Fischer, 1994; Kustu et al., 1989), flagellar assembly (Arora et al., 1997), arginine catabolism (Gardan et al., 1995), alginate production (Zielinski et al., 1992), rhamnolipid production (Pearson et al., 1997), acetoin catabolism (Priefert et al., 1992), mannose uptake (Martin-Verstraete, 1995) and proline iminopeptidase activity (Albertson and Koomey, 1993).

The basic promoters recognized by the $\sigma^{70}$ family are normally organized around two hexamers centered between -10 and -35 nucleotides upstream from the transcriptional start site though little diverse in their recognition sequence (Hawley and McClure, 1983). The initiation complex is further stabilized by the C-terminal domain of the core enzyme (aCTD), which can either interact directly with upstream DNA or with regulatory proteins (Benoff et al., 2002) and

its been reported that recognition of the -10 alone can also be sufficient for initiation to occur (Keilty and Rosenberg, 1987; Barne et al., 1997; Kumar et al., 1993).

Upstream sequence of *opd* gene has been used to identify promoter and other regulatory motifs using Bioinformatic tools. The results obtained in this study coincided with the predictions made by earlier investigators (Harper et al., 1988; Mulbry and Karns, 1989). The study conducted gathered fairly good evidence to show that a dual promoter system is involved in expression of *opd* gene. One of the promoters show highest similarity to the consensus $\sigma^{70}$ dependent promoter. Almost identical -35 hexameric sequence and optimal spacing between these conserved sequences (Hawley McClure, 1983) are some of the important features to believe that the predicted $\sigma^{70}$ dependent promoter is responsible for transcriptional activation of *opd* gene in *B. diminuta.* Upstream of *opd* gene existence of conserved dinucleotides, GG and GC with a gap of 11 nucleotides is the only reason to believe on existence of second $\sigma^{54}$ dependent promoter. In order to establish the functional status of these promoters, they were independently fused to the *lacZ* gene of promoter probe vector pMP220 (Spaink et al., 1987). This promoter test vector containing *opd-lacZ* fusions were subsequently mobilized into *B. diminuta* using *E. coli* S17-1 as helper strain. Significant β-galactosidase activity levels (2040 miller units) were noticed when consensus $\sigma^{70}$ dependent promoter is fused to the *lacZ* gene. This clearly indicates that the consensus $\sigma^{70}$ dependent promoter is responsible for transcriptional activation of *opd* gene in *B. diminuta*.

The *lacZ* fusions generated by including both $\sigma^{70}$ and $\sigma^{54}$ dependent promoters gave considerable β-galactosidase activity (943.61). However, the *lacZ* fusion (pSM9) obtained by fusing only putative $\sigma^{54}$ promoter motif failed to activate *lacZ* gene of promoter test vector.

Before concluding that $\sigma^{54}$ promoter motif non-functional, available literature on $\sigma^{54}$ dependent promoters was reviewed. In the $\sigma^{54}$ class of promoters the two conserved dinucleotides GG and GC at –12 and –24 regions are separated with a gap of 10 nucleotides (Bordes et al., 2004). The $\sigma^{54}$ has three functional domains, the carboxyl terminus is required for the binding of promoter DNA, the amino-terminal region is required for activation and the domain between these two regions is for binding core RNA polymerase (Sasse-Dwight and Gralla 1990). It has unique ability to bind promoter without the help of core polymerase. Upon binding to the promoter, recruits core polymerase and mediates interaction with transcription activator bound to UAS. This molecular events bring conformational changes in the promoter-DNA complex and facilitate opening of promoter to generate open promoter complex (Gralla, 1991; Wang et al., 1992; North et al., 1993). Yin and his co-workers have extensively studied regulation of $\sigma^{54}$ dependent *glnAP2* promoter (Yin et al., 1995). The authors have demonstrated molecular mechanism involved in opening and closing of transcription complex formation by RNAP-$\sigma^{54}$ and showed that when sufficient nitrogen is present, $\sigma^{54}$-holoenzyme forms a closed complex that occupies the *glnAp2* promoter in an inactive state (Sasse-Dwight and Gralla 1988). When nitrogen becomes insufficient, a cascade of reactions occur leading to the phosphorylation of enhancer-binding protein NtrC, which then binds to the upstream activating sequence (UAS) and activates transcription (Keener et al., 1988). The activation event occurs via DNA looping (Su et al. 1990) and involves the use of ATP to convert the closed complex to an open complex, which is active for transcription (Sasse-Dwight and Gralla 1988).

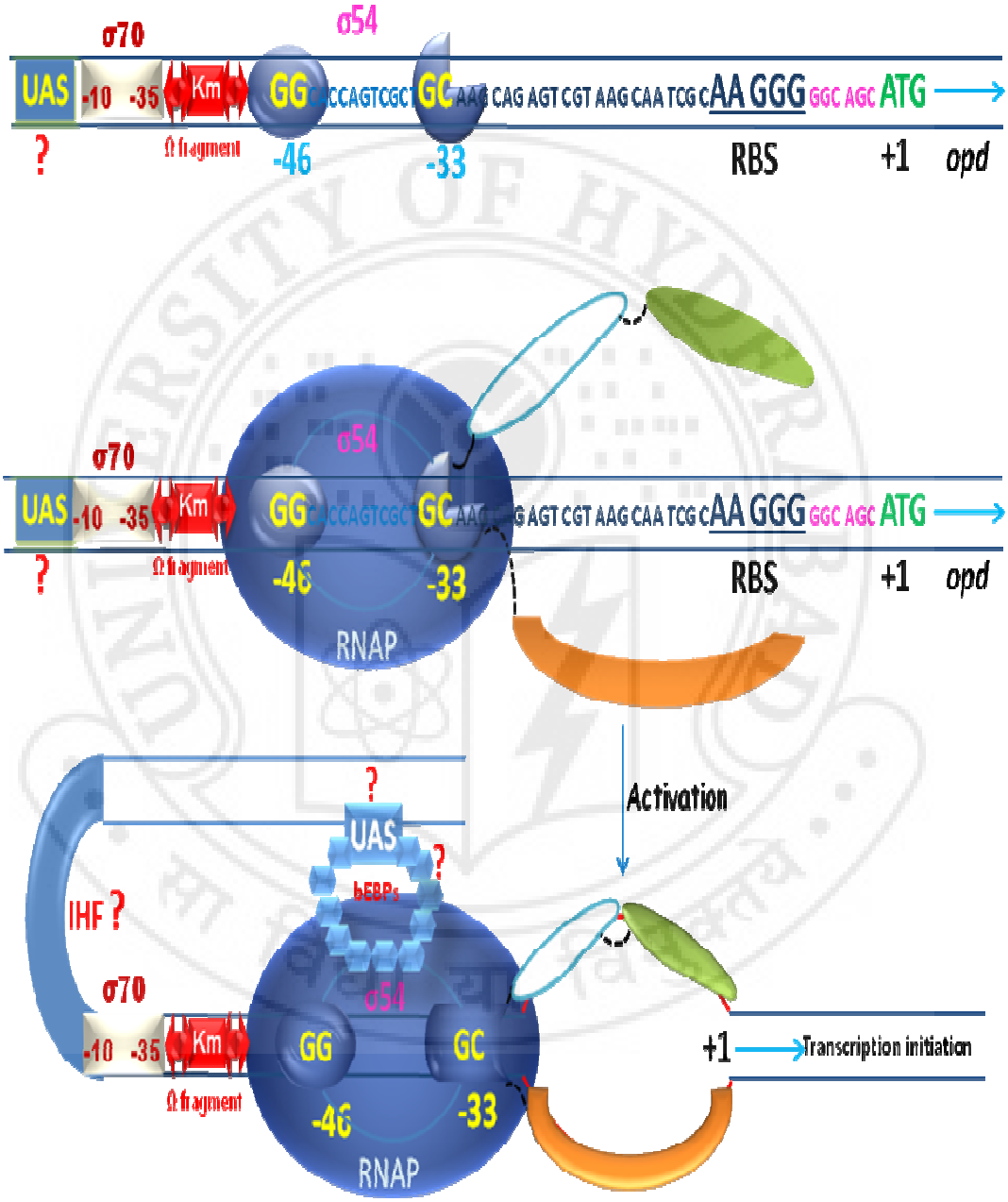
Open complex formation by RNAP-$\sigma^{54}$ requires ATP hydrolysis by activator proteins, which bind to enhancer like sequences located upstream of the promoter site (Sasse-Dwight and Gralla, 1988). The $\sigma^{54}$ activators are therefore also referred to as bacterial enhancer binding

proteins (bEBP). This is functionally analogous to enhancer dependent initiation of eukaryotic RNA polymerase II, which requires an input of energy from ATP hydrolysis provided by TFIIH (transcription factor IIH) (Lin et al., 2005). ATP hydrolysis by bEBPs provides energy for remodeling the $\sigma^{54}$–RNAP closed complex, resulting in further DNA melting and loading of the template strand of DNA into the RNAP active site (Cannon et al., 2003). Energy is transferred to the closed complex through a physical interaction between $\sigma^{54}$ and the AAA+ domain of the bEBP (Bordes et al., 2004; Lin et al., 2005). The process of closed to open promoter complexes is a multistep process (Davis et al., 2005). The bEBPs (bacterial enhancer-binding proteins) are AAA+ (ATPase associated with various cellular activities) family members involved in regulating bacterial gene expression (Ishihama, 2000; Browning and Busby, 2004). The bEBPs are essential for transcription initiation using the major variant $\sigma$ factor, $\sigma^{54}$, commonly involved in tightly regulated bacterial responses to stress (Buck et al., 2000; Bose et al., 2008).

On careful examination, in the putative $\sigma^{54}$ dependent promoter of *opd* gene, the conserved dinucleotides (GG and GC at –46 and –33 regions) are separated with a gap of 11 bp. A very weak promoter activity, obtained for *lacZ* fusion (pSM9) containing only $\sigma^{54}$ dependent promoter alone does not justify the putative $\sigma^{54}$ dependent promoter predicted upstream of *opd* gene is functional. As discussed in earlier sections, almost all consensus $\sigma^{54}$ dependent promoters contain upstream activating sequences (UAS) (Buck, et al. 1986) which participate in converting closed promoter complex into open promoter complex (Buck et al., 1986). Deletion of UAS from $\sigma^{54}$ dependent promoters significantly lowers the transcriptional activity (Reitzer and Magasanik, 1986).

The *lacZ* fusion pSM9, as shown in Fig. 3. 2, is constructed by taking just basic promoter element. It does not contain much of the upstream region of *opd* gene which probably contain UAS. Therefore, the weak β-galactosidase activity cannot be taken as basis to dismiss the claim on existence of a $\sigma^{54}$ dependent promoter upstream of *opd* gene. The UAS is found in all most all reported cases only upstream of the promoter motifs. They are located anywhere from 300 bp to 1 kb away from the promoter motif (Buck et al., 1986). In order to identify if such UAS like motif exists upstream of *opd* promoter, a *lacZ* fusion was constructed by cloning the entire upstream region of *opd* gene. Such *lcaZ* fusion (pSM11) includes both putative $\sigma^{70}$ and $\sigma^{54}$ dependent promoters. It is therefore difficult to predict actual promoter responsible for transcription activation of *opd* gene. Insertion of omega fragment between these two promoters terminates transcription initiated from upstreamly located $\sigma^{70}$ dependent promoter and hence the *lacZ* activation as monitored by β-galactosidase activity will be exclusively from $\sigma^{54}$ dependent promoter. The obtained β-galactosidase activity for this construct (514.44) suggests that $\sigma^{54}$ dependent promoter is functional in *B. diminuta.* Such activity is not seen in the construct that was generated by including only $\sigma^{54}$ dependent promoter (pSM9). This may be due to non availability of UAS in this construct. The omega fragment present in pSM11 only prevents transcription from $\sigma^{70}$ dependent promoter and does not prevent possible interactions between $\sigma^{54}$ dependent promoter and putative UAS by inducing a loop in the DNA region available between $\sigma^{54}$ dependent promoter and UAS. A schematic model showing $\sigma^{54}$ dependent transcriptional activation of *opd* gene is shown in fig. 3. 12.

**Fig. 3. 12**




A schematic representation showing proposed model of $\sigma^{54}$ dependent transcriptional activation of *opd* gene.

The $\sigma^{54}$ dependent promoters are inducible promoter. They respond to a number of physiological conditions including that of nitrogen limiting condition. A phosphorylated transcription activator after bending to UAS influence the transcription activation of $\sigma^{54}$ dependent promoters. Phosphorylation of transcription activator protein is only done under defined physiological conditions. In the absence of phosphorylation transcription activator protein becomes inactive and only basal level of expression will be seen from $\sigma^{54}$ dependent promoter. In the present study, we have grown *B. diminuta* cells in rich medium. There is no stress for the cells to show elevated *lacZ* activation from $\sigma^{54}$dependent promoter. In fact it is not really known under which circumstances the *opd* gene is induced. As this study is mainly focused on membrane targeting of OPH no effort is made to identify a physiological situation that contributes for the induction of *opd* gene in *B. diminuta.*

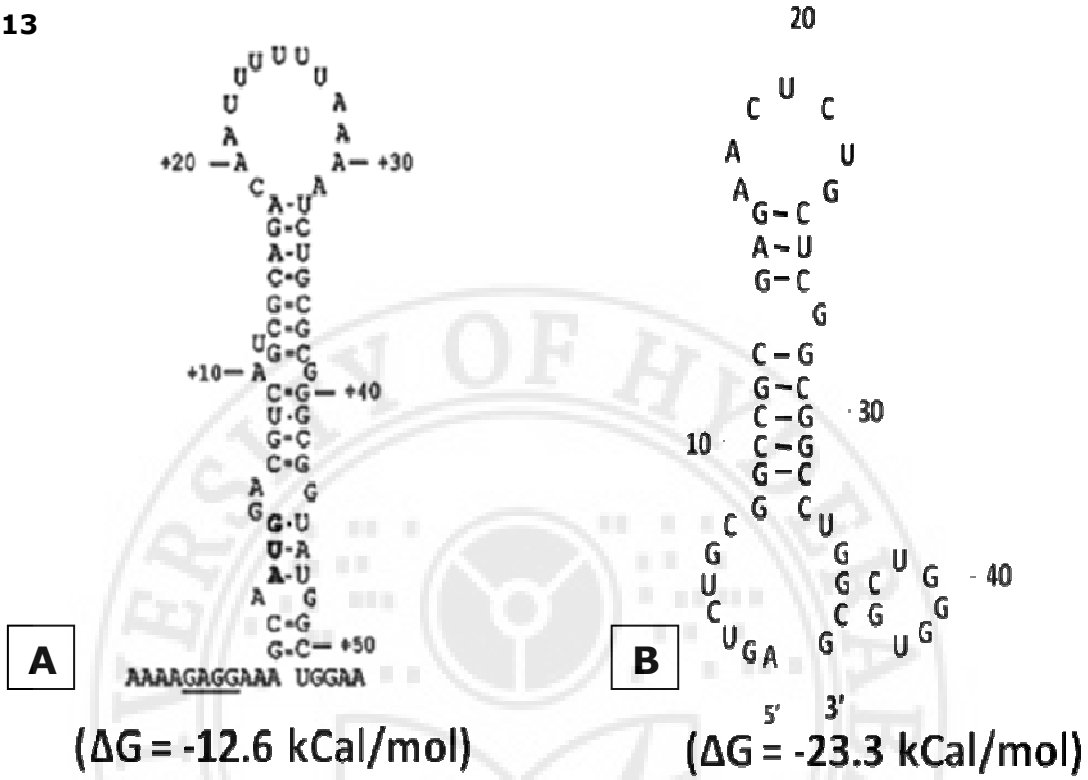
Existence of dual promoters upstream of genes or operons is not uncommon in prokaryotic world (Keener and Kustu, 1988). A classical example to cite for existence of both $\sigma^{70}$ and $\sigma^{54}$ dependent promoters is *ntr* operon (Keener and Kustu, 1988). In *ntr* regulon, the $\sigma^{70}$ is responsible for basal level expression, whereas $\sigma^{54}$ dependent promoter is induced in response to nitrogen limiting conditions (Hirschman et al., 1985). If β-galactosidase activities obtained from pSM7 and pSM11 are examined, it is a clear indication to suggest that both $\sigma^{70}$ and $\sigma^{54}$ dependent promoters found upstream of *opd* gene are functional. A dual promoter system appears to be controlling the expression of *opd* gene in *B. diminuta.*

As stated in the earlier sections of this chapter, the study was prompted by an unusual observation made during the heterologous expression of *opd* gene. The construct made to express mature form of OPH (mOPH) produced 2.5 fold more protein when compared to the construct

generated for coding precursor form of OPH (preOPH). If this result is seen together with β-galactosidase activity levels obtained for *opd-lacZ* fusions pSM8 and pSM10 (Fig. 3. 3), it clearly indicates existence of a *'cis'* element modulating the expression of *opd* gene. Inverted repeats have been shown to play crucial roles in various cellular activities. Their involvement in regulation of DNA replication, transcriptional regulation, as well as translational control of gene expression is well documented (Scortecci et al. 1999). In fact, the secondary structures of leader sequences have also been identified as both transcriptional and translational down-regulators of gene expression (Wood et al. 1996; Curie and McCormick 1997; Hemmings-Mieszczak et al. 1998; Scortecci et al. 1999). Initially Real Time PCR was done to assess if the IR sequence is acting as transcription terminator. Though the *opd* variants produced variable amounts of OPH they all have identical vector back bone and are controlled by identical transcriptional and translational signals. One of the possible reasons for pHYS00 to produce less amount of OPH, is due to termination of transcript at IR sequence, initiated from the $T_7$ promoter of the vector. Such termination of transcript should lead to variation in production of *opd* specific mRNA from these expression plasmids. The reduced transcription level of mRNA might be contributing for the low level of OPH synthesis in cells having pHYS400. In order to gain better insights on this hypothesis, mRNA produced from these plasmids pHYS400, pHNS400 and pHIR400 were quantified by Real Time PCR. As shown in Fig. 3. 11 there was no difference in the concentration of *opd* specific mRNA made by these three constructs. In the event of finding identical *opd* specific mRNA concentration, for the reduced OPH concentration in cells having pHYS400, the event of transcription is not responsible. Therefore IR mediated regulatory event may be associated with translation process.

Predictions of mRNA secondary structure mediated modulation of gene expression is very common both in prokaryotic and eukaryotic world. Instead of taking very unsimilar situation for discussion, here mRNA secondary structure translation modulation of Tat-dependent Formate Dehydrogenase N (FDH-N) is taken for drawing comparison with the regulation of OPH. Like OPH, FDH-N is a metalloenzyme. It is associated with membrane (Punginelli et al., 2004). If *E. coli* cells are grown under anaerobic condition, FDH-N represents about 10% of membrane protein. The FDH-N, like OPH, is a Tat substrate and depended on Twin arginine transport machinery for membrane targeting. It is also associated to the inner membrane facing periplasm (Punginelli et al., 2004). The FDH-N is coded by an operon designated as *fdnGHI*. The high-resolution crystal structure of FDH-N has shown that it adopts an $(\alpha\beta\gamma)_3$ "trimer-of-trimers" architecture, with the active site of the enzyme facing periplasm (Jormakka et al., 2002). Out of these three polypeptides only FdnG has signal peptide, which has twin arginine motif. The rest of the polypeptides do not contain any signal peptides. The FdnG leads them to the membrane in a prefolded form (Stanley et al., 2002). While experiments were conducted to show that FDH-N is a Tat substrate, mutations were generated by changing one of the invariant arginines (R) of Tat motif to lysine. This mutation has contributed for dramatic increase in FdnG synthesis (Punginelli et al., 2004). Further, it also contributed for the increase of partner subunits synthesis contributing for elevated synthesis of FDH-N (Punginelli et al., 2004). Punginelli and associates have done a number of experiments to show that *fdnG* mRNA has a secondary structure which required a free energy $\Delta G$ of -12.6 kcal/mol (Fig. 3. 13). Further they have shown that the secondary structure has modulated the expression of FDH-N in *E. coli.* The authors also suggested that this modulation facilitates interaction of partner subunits promoting the formation of a functional complex to be transported in a prefolded form.

**Fig. 3. 13**




Secondary structure and ΔG value comparisons for *fdnG* (A) and *opd* (B) specific mRNAs

Like FDH-N, OPH is a membrane protein and targets membrane via Tat pathway. It is not yet known if it is part of the multi subunit complex. As described earlier four types of proteins choose Tat route for membrane targeting. OPH does not require large cofactors and therefore it does not belong to those group of proteins that take Tat route along with large cofactors. However, having known that it interacts with Orf306, a product of *orf306* gene located adjacent to the *opd* gene of transposon-like *opd* gene cluster (unpublished results), its existence as multisubunit complex in the membrane of *B. diminuta* cannot be ruled out. If that is true, the mRNA secondary structure mediated modulation of OPH expression gain lot of significance.

# Conclusions

- **A well defined Tat motif that shows high similarity to the consensus Tat motif involved in transport / membrane targeting of prefolded proteins in prokaryotes is discovered in *B. diminuta.***

- **The 29 amino acid long signal peptide is not required for folding of OPH to gain active conformation.**

- **The invariant arginines found in Tat motif are required for membrane targeting of OPH in *B. diminuta.***

- **The Tat motif inserts OPH at the periplasmic face of inner membrane.**

- **Upstream of *opd* gene two putative promoter elements were identified. One of them showed high similarity to the consensus $\sigma^{70}$ dependent promoter. The second promoter element showing similarity to the consensus $\sigma^{54}$ dependent promoter activated promoter less *lacZ* gene, indicating the possibility of existing a second promoter upstream of *opd* gene.**

- **An inverted repeat sequence found in the signal peptide coding region of *opd* gene is probably responsible for down regulation of OPH in *E. coli.***

**Research Publications:**


1. **Purushotham Gorla.,** Pandey, J.P., Sunil, P. Merrick, M.J. and Siddavattam, D. (2009) Organophosphate hydrolase in *Brevundimonas diminuta* is targeted to the periplasmic face of the inner membrane by the twin arginine translocation (Tat) pathway. *J. Bacteriol.* **191 (20):** 6292-6299.

2. Pandey, J.P., **Purushotham Gorla,** Manavathi, B. and Siddavattam. D. (2009) mRNA secondary structure modulates the translation of organophosphate hydrolase (OPH) in *E. coli*. *Mol. Biol. Rep.* **36:** 449-454.

3. Pakala S.B, **Purushotham Gorla**, Pinjari A.B, Ravikumar, K., Rajasekhar B, Yanamandra, Y, Merrick M.J, Siddavattam D (2006) Biodegradation of methyl parathion and pnitrophenol: Evidence for the presence of a *p*-nityrophenol 2-hydroxylase in Gramnegative *Serratia* sp. strain DS001. *App. Microbiol. Biotechnol.* **73:** 1452-1462.

4. Veerana, G., Paul, E.V., **Purushotham Gorla**, Siddavattam, D. and Karegoudar (2006) Complete mineralisation of dimethylformamide by *Ochrobactrum* sp. DGVK1 isolated from the soil samples collected from the coalmine leftovers. *Appl. Microbiol. Biotechnol.* **71:** 369-375.

5. Bramanandam, M., Pakala, S.B., **Purushotham Gorla**, Merrick, M.J. and Siddavattam, D. (2005). Influence of zinc and cobalt on expression and activity of parathion hydrolase from *Flavobacterium* sp. ATCC27551. *Pest.Biochem.Physiol.* **83:** 37-45.