SEQUENCING AND EXPRESSION OF THE REGULATORY GENE *tfdR* OF PLASMID pJP4

By

UDAYAKUMAR MATRUBUTHAM

Bachelor of Science
Tamilnadu Agricultural University
Coimbatore, India
1982

Master of Science
Tamilnadu Agricultural University
Coimbatore, India
1984

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1993

OKLAHOMA STATE UNIVERSITY

SEQUENCING AND EXPRESSION OF THE
REGULATORY GENE *tfdR* OF
PLASMID pJP4

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express my appreciation to my adviser, Dr. Alan R. Harker for his constant guidance and encouragement throughout the course of this study. I would like to extend my sincere thanks to Dr. Ulrich Melcher for his timely suggestions and comments as a member of my advisory committee. I also extend my appreciation to my other committee members Drs. Kim Burnham and James Harmon. I wish to thank all my lab mates for their help and friendship.

I wish to express my appreciation to my friends in this Department for their enjoyable companionship.

My sincere appreciation goes to my wife, Sowmya, who with her love and understanding made possible those final experiments, and the writing of this thesis more endurable. I also extend my sincere appreciation to all the members of my family (Amma, Akka, Ravi, Susi, Mangalam, Jinna, Vimala, Nandu and Poori) for their constant encouragement from India. Finally, I would like to dedicate this thesis to my late dad, Appa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Environmental pollution requires very little introduction as it is a common problem these days. Among the pollutants in the environment, man-made chemicals rank foremost in abundance. Comforts and necessities demanded by society have resulted in the synthesis and release into the environment of many organic and inorganic chemicals. Accelerated dispersal of significant quantities of chemical pollutants has increased the demand for cleaner environment and has prompted numerous research studies. Research on the characteristics of toxic chemical pollutants and their fate in the environment has been given high priority.

Chemicals used as refrigerants, fire retardants, degreasers, herbicides, pesticides, paints, plastics and solvents are major pollutants in the environment and cause health hazards (14). Most are man-made chemicals, foreign to the environment and are referred to as xenobiotics. Xenobiotics either persist or are degraded and mineralized. Degradation of xenobiotics depends on various biotic and abiotic factors. Biotic factors include slow enzymatic attack by microorganisms. Microbial degradation of xenobiotics is relatively slow as microbes depend on the availability of nutrients

and abiotic factors like optimum temperature, pH, oxygen concentration, etc. In many cases, biodegradation by-products are lethal to microbes.

Many xenobiotics have been released into the environment. Today, many microorganisms isolated from polluted environments show adaptation to some of them. Microbial adaptation to xenobiotics refers either to a mixed population or to a single species that have survived after being exposed to a pollutant for an undefined period of time. Adaptation has been proposed as the source of resistance to mercuric derivatives found among bacterial species isolated from mercury contaminated waters and soils (87). How did these bacterial species overcome the lethal effects of mercury? What processes occur during adaptation? These questions remain largely unresolved. Many metal resistances have been found in bacteria isolated from other polluted systems. In the last several decades, bacterial species capable of the biodegradation of xenobiotics have also been isolated from both aquatic and soil biota. These species often included members from genera *Arthrobacter*, *Acinetobacter*, *Alcaligenes*, *Bacillus*, *Corynebacterium*, *Flavobacterium* and *Pseudomonas* (see Table I of Chapter II)

Commercial use of bacterial strains capable of biodegradation have made environmental research very competitive and captured the public interest. The application in bioremediation of polluted environments is typified by the Alaskan oil spill in Prince Williams Sound. "Oil eating" bacteria were sprayed on contaminated beaches in an effort to clean up the spill. The particular "oil eating" *Pseudomonas* sp was isolated and commercially patented in the early

1980's. Today, isolation and screening of bacteria with wider biodegradative potentials have almost become common. Attempts to construct novel pathways using molecular techniques are also on the rise (92). Consequently, molecular information on the genetics of biodegradative pathways are essential to better understand the degradation pathways and their regulatory mechanisms.

Molecular investigations have disclosed inconsistencies in the physical arrangement of biodegradative genes and their respective functions among the bacterial species isolated from identically polluted environments. The structure, organization, and regulation of these genes have shown significant genetic variability. Variable gene copy number has been noted in similar pathways expressed in two different bacterial isolates. For example, *Pseudomonas putida* has been shown to carry only a single set of genes that encode degradation of 3-oxoadipate, an intermediate metabolite, formed in both catechol and protocatechuate degradation pathways (58). Whereas, *Acinetobacter calcoaceticus* has two sets of equivalent genes that encode isofunctional enzymes to degrade the intermediate from these two pathways (55). Why does this bacterium require an additional set of genes? Does this represent increased sophistication over the others? Very little is known about the significance of these variations. Incomplete duplication of catechol gene clusters has been reported in bacteria under selective pressure. Chlorocatechol gene clusters, for example, in *Alcaligenes eutrophus* and *Pseudomonas* sp when grown on 3-chlorobenzoate exhibit such a duplication (32, 34). Is this how *A. calcoaceticus* acquired an additional set of genes? Further speculations have raised more questions than answers. Is it

possible that the duplicated gene cluster represents a mechanism inherent in the evolutionary process? Is it likely that the duplicated gene cluster will evolve to provide a degradation pathway for an entirely new compound? To answer these questions more knowledge on the molecular mechanisms and evolution of degradative pathways is essential.

Discrepancies in the organization of degradative genes have also been noted between various bacterial isolates. Often, genes in any two similar or related degradation pathways have shown varied physical configurations. For example, catechol degradation pathway genes (*cat* genes) in *A. calcoaceticus* ADP1 are clustered into operonic units and located on a plasmid (59). In *P. putida* they exist unlinked on the chromosome (58). Reasons for such differences may reflect the varying necessity of bacterial species to transfer or conserve biodegradative genes. Genes located on plasmids are usually more highly transmissible than those on the chromosome and therefore disseminate biodegradative genes relatively rapidly. But, chromosomal genes may also provide more opportunities for complex recombinational events that may result in the evolution of new genes for the efficient degradation of future xenobiotics.

Contemporary studies on the regulatory mechanisms of biodegradative pathways have only provided rudimentary understanding. Most of the known biodegradative pathways have been found to contain regulatory elements. These elements are usually genes that encode proteins which either repress or activate the expression of metabolic genes. The regulatory genes *nahR*, *catR*, *catM*, *tcbR* and *tfdS* are some the control elements identified in

various degradative pathways (73, 68, 56, 90, 42). These genes encode similar sized proteins that belong to the *lysR* family (40). It has been observed that most of these regulatory proteins require specific metabolites for their activity (42, 54). Commonly, these metabolites are by-products of the regulated pathway. No consensus has yet been achieved concerning the nature of the required metabolite. Such in-depth knowledge of all molecular mechanisms is essential to successfully and genetically manipulate bacteria, construct novel pathways, and improve the efficient application of degradation potential.

This thesis describes the study of the molecular structure, expression, and function of the regulatory gene *tfdR* of plasmid pJP4 of *Alcaligenes eutrophus* JMP134. The plasmid pJP4 carries genes which encode degradation of the herbicide 2,4-dichlorophenoxy-acetic acid.

Goals and Strategies of Research

* The first goal was to determine the complete nucleotide sequence of the regulatory gene *tfdR* of pJP4.

** The essential regions on *Bam*HI-E and *Bam*HI-F fragments of pJP4 were sequenced and analysed for possible open reading frames.

* The second goal was to express *tfdR* gene.

** On identifing the open reading frame of *tfdR* gene, appropriate DNA regions were cloned into expression vectors, protein

was expressed, and selectively labeled with radioactive methionine.

* The third goal was to investigate the function of TfdR protein.

** After over-expressing TfdR protein, the protein was checked for DNA binding ability.

* The final goal was to determine the relationship of the *tfdR* gene with other genes in available databases.

** Sequence alignments of *tfdR* DNA and protein sequences were accomplished using available computer software.

Purpose of this Research

The main purpose of this study is to characterize the regulatory gene *tfdR*. This research deals exclusively with the regulatory gene, its organization and expression. The information from this study and the use of genetic engineering will help construction of improved bacterial strains with enhanced degradative ability and broader substrate utilization potentials. Eventually all these efforts will help efficient bioremediation of polluted environments.

CHAPTER II

LITERATURE REVIEW

Overview of the Degradation of Halogenated Organics

Halogenated Organic Compounds

Halogenated organics are used as herbicides, plastics, solvents and degreasers. Halogenation is often implicated as a reason for the persistence of these compounds in the environment (57). For example, the highly publicized environmental problems with the herbicide DDT (1,1,1-trichloro-2,2-bis (p-chlorophenyl) ethane) were due primarily to its recalcitrant nature (64). The pesticide effect requires chlorines, which in turn cause the compound to be more persistent. DDT residues persist on plants, plant products and bioaccumulate in the food-chain. Persistance and accumulation resulted in lethality to many species. Chlorinated hydrocarbons are the most extensively studied of the halogenated organics. They serve as a basis for most of the available information on the

biotransformation of synthetic organics or xenobiotics. Most of the studies have been focused on the aerobic nature of degradation. This is largely because aerobic cultures are easier to work with than anaerobic cultures. But recently, there have been reports on the anaerobic degradation of halogenated hydrocarbons also (see references in 14). Chlorinated hydrocarbons degraded by microorganisms are divided into three groups. They are (1) aliphatic, (2) polycyclic and (3) aromatic (14).

Bacteria Degrading Chlorinated Aromatics

Chlorinated aromatics like phenoxyacetates have been released into the environment over the past several decades. They are considered major pollutants as they are released in substantial quantities, are toxic, and accumulate in sediment and biota. Like many other xenobiotics, these are degraded relatively slowly by soil and aquatic microorganisms. Many soil and aquatic bacteria have been isolated and studied for their capacity to dissimilate chlorinated aromatics like chlorobenzoates (10, 11, 12, 53), chlorobenzenes (36, 75), chlorotoluenes (14), chlorophenols (74), chloroacetamides and chlorophenoxyacetates (25, 43, 48, 86). Table I on the following page gives the names of different bacteria with the ability to degrade an array of chlorinated aromatics.

TABLE I

BIODEGRADATION OF SOME CHLORINATED AROMATIC COMPOUNDS

Compound	Microorganism	Plasmids	Reference
1,2-dichlorobenzene	*Pseudomonas* sp.		36
1,4-dichlorobenzene	*Alcaligenes* sp		75
2,6-dichlorotoluene	*P. cepacia* HCV		88
3-chlorobenzoate	*A. eutrophus*	+	10
	Acinetobacter calcoaceticus		94
	Pseudomonas sp. strain B13	+	21
	P. putida	+	34
	Flavobacterium sp.	+	13
4-chlorobenzoate	*Arthrobacter* sp.		53
	Pseudomonas sp. strain CBS3		71
	A. denitrificans NTB-1		91
	Corynebacterium sepedonicum		95
4-chlorophenol	*Pseudomonas* sp.		74
4-chlorophenylacetate	*Pseudomonas sp.* strain CBS3		46
2,4-D	*A. eutrophus*		61
	Pseudomonas sp.	+	45
	Flavobacterium sp.	+	13
2,4,5-T	*P. cepacia* AC1100		43

"+" indicates plasmid mediated degradation

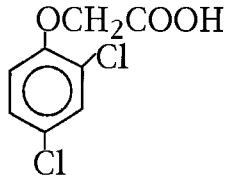
Physiological Pathways for the Degradation of Chlorinated Aromatics

Many different pathways have been described for aerobic bacterial degradation of mono- and dichlorinated aromatic compounds. General mechanisms include breakage of the carbon-halogen bond as the first step by hydrolytic, oxygenolytic or reductive mechanisms, as well as modifications of the gentisate pathway (67). They also include *meta* cleavage of chloroprotocatechuate and the *ortho* cleavage of halosubstituted catechols (67), benzoates, phenols and phenoxyacetates (7, 19, 25, 26, 48, 86). Therefore the degradation of chlorinated aromatics involves modified versions of either *meta* or *ortho* cleavage pathways. An appropriate example to this study, a modified *ortho* cleavage pathway of 2,4-dichlorophenoxyacetate (TFD), is given in Figure 1.

Genetics of Chlorinated Aromatics Degradation

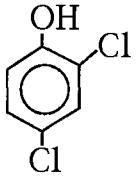
Many bacterial species tested for the capacity to degrade chlorinated aromatics were found to carry the pathway genes on plasmids. Table II contains the names of various plasmids that are involved in the degradation of chlorinated aromatics. Plasmid pRC10 have been show to be homologus with that of pJP4 (13). In this research the plasmid pJP4 of *Alcaligenes eutrophus* JMP134, that encodes the degradation of TFD, 3-chlorobenzoate (3CB), 2-methyl-4-

Figure 1. Proposed *Ortho* Cleavage Pathway for Microbial Degradation of 2,4-D. The names of *tfd* gene and enzyme product are given for each step.



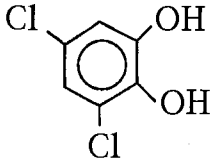
2,4-dichlorophenoxyacetate
(TFD)

⇩ TFD monooxygenase
(*tfdA*)



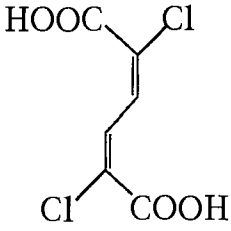
2,4-dichlorophenol

⇩ 2,4-dichlorophenol hydroxylase
(*tfdB*)



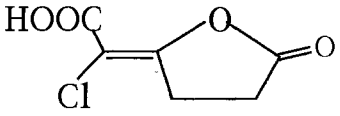
3,5-dichlorocatechol

⇩ Chlorocatechol 1,2-dioxygenase
(*tfdC*)



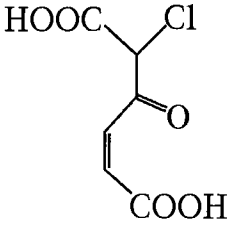
2,4-dichloromuconate

⇩ Chloromuconate cycloisomerase
(*tfdD*)



cis,2-chlorodiene lactone

⇩ 2-chlorodiene lactone hydrolase
(*tfdE*)



Chloromaleylacetate

⟹ Succinate + Acetyl coA

TABLE II

SOME DESIGNATED PLASMIDS ENCODING DEGRADATION OF CHLORINATED HYDROCARBONS

Plasmid	Compound	Size	Reference
pUU204	2-Monochloropropionate	53 kb	37
pKF	4-Chlorobiphenyl	82 kb	30
pSS50	4-Chlorobiphenyl	53 kb	77
pAC21	1,4-Dichlorobiphenyl	65 MDa	12
pJP4	3-Chlorobenzoate (3CBA)		
	2,4-Dichlorophenoxyacetate		
	2-methyl,4-chloroxyacetate	80 kb	21
pAC27	3CBA	110 kb	11
pAC31	3,5-Dichlorobenzoate	105 kb	11
pRC10	3CBA		
	2,4-Dichlorophenoxyacetate		
	2-methyl,4-chloroxyacetate	45 kb	13
pUO1	Flouroacetate	44 MDa	44
pUO11	Flouruacetate	40 MDa	44

chlorophenoxyacetic acid, and trichloroethylene (TCE), has been used.

Regulatory Genes of Chloroaromatic Degradation Pathways

Several regulatory genes regulating catabolic pathways have been identified, such as *catR*, *catM*, *tfdS*, *tcbR* and *clcR* (68, 56, 42, 90, 28). Excepting *tfdS* all these genes have been found to regulate catechol degradation pathways. All these genes have sizes of approximately 1 kb, are transcribed divergently from the target operons or gene, and are separated from these by approximately 170-200 bp. Interestingly, most of these genes encode transcriptional activator proteins, except *catM*, which encodes a repressor. These genes require effector molecules for their activity. The effector of *catR* and *catM* has been found to be *cis,cis*-muconate (59, 58). Effectors have not been identified for other genes. All these regulatory proteins belong to LysR family of activators (82, 40). Although the N-terminal parts of these proteins have been observed to contain similar helix-turn-helix motifs, presumably involved in DNA binding, the C-terminal regions are dissimilar (40). This suggests that perhaps each protein in this family arose by fusion of at least two different domains, one of which was shared by all members.

Degradation of TFD by the pJP4 Encoded Pathway

2,4-Dichlorophenoxyacetate

The structure of TFD is given in Figure 1. TFD is a chlorinated aromatic hydrocarbon belonging to the group of halogenated phenoxyacetates. Its chemical formula is $C_8H_6Cl_2O_3$ and it has a molecular weight of 221.04. TFD is used as an herbicide to kill broad leaf plants. TFD acts like an auxin, the plant growth hormone, and hastens the proliferation of vascular tissues. The enlarged tissues block further transport of nutrients and result in the death of the plants. Monocots with scattered vascular bundles escape the blockage effects caused by tissue proliferation (51).

Degradation of TFD

Unlike many other chlorophenoxyacetates, TFD is degraded by soil microorganisms. TFD has a half life of six weeks on soil surfaces. A number of bacterial strains have been isolated which are capable of TFD degradation. The compound is degraded via a modified *ortho* cleavage pathway (25, 67). TFD is converted into 2,4-dichlorophenol and then to 1,2-dichlorocatechol (refer to Fig.1). Then chlorocatechol is converted to *cis, cis*-chloromuconate, which is transformed into

chlorodienelactone. This is converted to chloromaleylacetate. Chloromaleylacetate is then converted to chlorosuccinate and acetyl coenzyme A via an NADH-dependent reduction or to β-ketoadipate with the consumption of NADH and liberation of chloride (24, 49). All conversions from chloromaleylacetate are encoded by the chromosome and not by pJP4 (49). Acetyl coenzyme A and succinate are formed from β-ketoadipate via the conventional β-ketoadipate pathway and enter the tricarboxylic acid cycle.

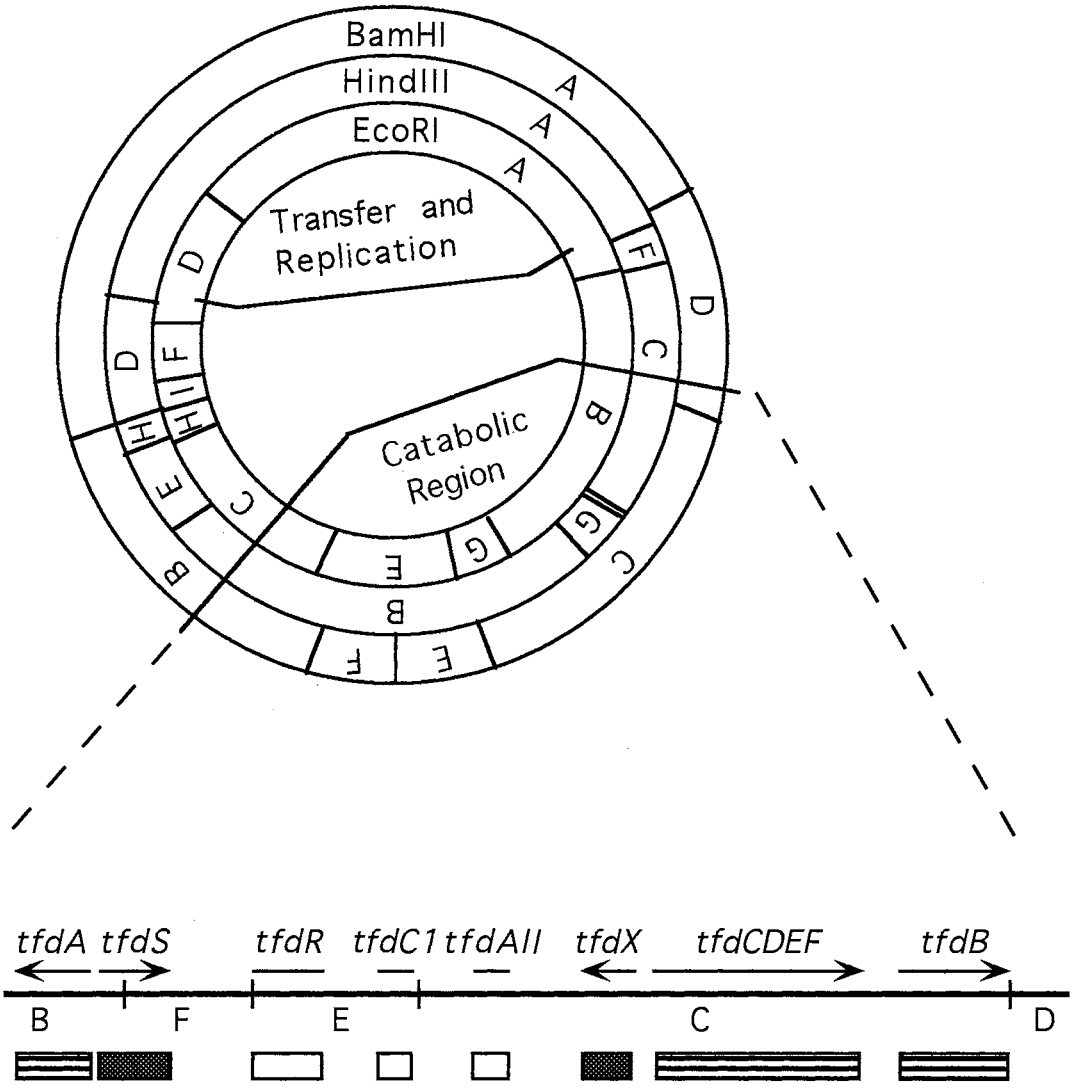
Plasmid pJP4

The plasmid pJP4 was isolated from *Alcaligenes eutrophus* strain JMP134 in Australia (20). It is a conjugative, broad host range and 80 kilobase plasmid. It belongs to the IncP incompatability group of plasmids. It encodes resistance to mercuric chloride and phenyl mercuryl acetate. It also encodes the degradation of TFD, 3CB and TCE (20, 38).

Genes of TFD Pathway

The plasmid pJP4 (Fig. 2) encodes genes for the degradation of TFD. Several genes of this metabolic pathway have been mapped by transposon mutagenesis (22). The genes and their enzyme products are shown in Figure 1, along with the modified *ortho* cleavage

Figure 2. Restriction Map of Plasmid pJP4. The portions on the plasmid mediating various functions are labeled. The regions containing *tfd* genes on *Bam*HI fragments are extrapolated and the positions of *tfd* genes are shown. The arrows indicate the direction of transcription. The short lines indicate unknown direction of genes. Symbols:  completely sequenced;  partially sequenced; and  not sequenced genes.


BamHI
HindIII
EcoRI
Transfer and Replication
Catabolic Region
A
A
A
D
F
D
C
B
G
G
C
B
E
F
E
C
H
H
F
D
B
tfdA
tfdS
tfdR
tfdC1
tfdAII
tfdX
tfdCDEF
tfdB
B
F
E
C
D

pathway (21). The restriction maps of pJP4 and the position of TFD genes are shown in Figure 2. Isofunctional of copies *tfdA* and *tfdC* genes have also been located on pJP4 (63, 34). They are named *tfdAII* and *tfdC1* respectively.

Regulation of TFD Pathway

The expressions of TFD genes are controlled by regulatory elements (39, 41, 42). It has been shown that a 1.2kb region on the *Bam*HI-E fragment was sufficient to negatively regulate the TFD monoxygenase gene *tfdA*, and the chlorocatechol degradation encoding operon *tfdCDEF* (39, 41). The 1.2kb region has been proposed to contain the regulatory gene *tfdR*. It has also been proposed that pJP4 carries another regulatory region comprising small portions of both *Bam*HI-B and *Bam*HI-F (83, 40). This region has been shown to contain the regulatory gene *tfdS* that represses and activates the TFP hydroxylase gene *tfdB*, in the presence or absence of specific metabolites (42). It has already been known that there is an incomplete open reading frame (ORF) near the promoter region of *tfdA* gene (40). This ORF has been proposed as part of *tfdS* and was deduced to encode a LysR homolog protein. It has been strongly suggested that the *tfdS* gene encodes a transcriptional activator protein. The relative positions of genes *tfdR* and *tfdS* on pJP4 are indicated in Figure 2.

CHAPTER III

MATERIALS AND METHODS

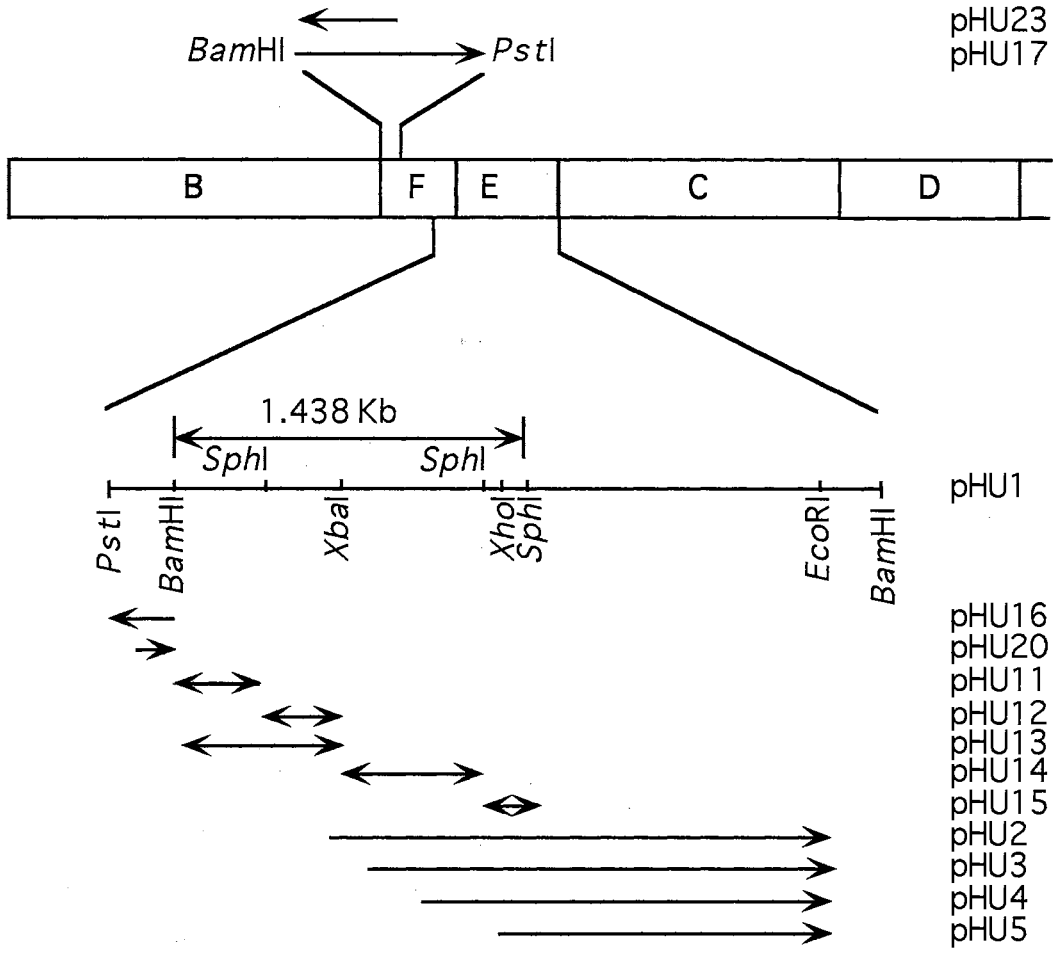
Sequencing of *tfdR* Gene

Essential regions of *Bam*HI-E and *Bam*HI-F of pJP4 were sequenced to determine complete nucleotide sequence of *tfdR*. Sequencing was performed using the method of Sanger et al. (70).

Construction of Sequencing Clones

Different regions of the *Bam*HI-E and *Bam*HI-F fragments of pJP4 were cloned into the sequencing vector pGEM 7zf(+) purchased from Promega. A schematic representation of these constructs is given in Figure 3. The construct pHU1 was linearized with *Bam*HI and used in *Bal*31 digestion. A typical *Bal*31 reaction contained approximately 30μg of linearized DNA in a buffer containing 0.2M NaCl, 20mM Tris.Cl (pH 8.0), 12mM $MgCl_2$, 12mM $CaCl_2$, 2mM EDTA and 0.25μg/μl of BSA. The reaction was performed at 30°C by adding 1 unit of the enzyme. At 1 minute intervals aliquots of the reaction

Figure 3. Construction of Sequencing Clones. Regions of *Bam*HI-E and F of pJP4 were sub cloned into pGEM 7zf(+). The restriction sites used in cloning are indicated. Clones pHU2 to 5, pHU20 and pHU23 were created by *Bal*31 digestion of clones pHU1, pHU16 and pHU17 respectively. The arrows indicate directions of sequencing.


pHU23
pHU17
BamHI
PstI
B
F
E
C
D
1.438 Kb
SphI
SphI
pHU1
PstI
BamHI
XbaI
XhoI
SphI
EcoRI
BamHI
pHU16
pHU20
pHU11
pHU12
pHU13
pHU14
pHU15
pHU2
pHU3
pHU4
pHU5

mixture were drawn into separate tubes containing 200mM EGTA and placed at room temperature. DNA was precipitated by the addition of 2 volumes of ethanol. It was dissolved in 20µl of *Eco*RI buffer and

digested with *Eco*RI to separate the insert from the vector. The digested DNA samples were electrophoresed on a 1.0% low melting agarose gel (SeaPlaque, FMC). The DNA was visualized under UV light after ethidium bromide staining. DNA samples from different time intervals were cut from the gel. The gel slices were melted and directly used in separate ligations. The vector which was used in the ligation had *Eco*RI and *Sma*I cut ends. Construction of other sequencing clones involving cohesive end ligations were performed after digesting the vector and pHU1 with the same enzymes. In these cases the insert DNA after digestion was isolated on an agarose gel as explained previously.

Aliquots of ligated DNA for each time interval were transformed into competent JM109 *E. coli* cells (93). Transformation was accomplished by heat shocking the mixture of cells and DNA at 42°C for 2 minutes. The cells were transferred to LB medium for about 1-2 hours at 37°C. They were then plated on LB plates containing ampicillin (50µg/ml), 0.5mM IPTG and X-Gal (40µg/ml) to select and screen for transformants. White recombinant colonies were picked from each time interval and checked for the presence of insert DNA in the vector plasmid by alkaline mini-lysis method of Birnboim-Doly (6).

Ligations

A ligation reaction was set up in a buffer containing 30mM Tris-HCl (pH 7.8), 10mM $MgCl_2$, 10mM DDT and 1mM ATP. In each reaction approximately 100ng of the vector, 25ng of insert and 1 Weiss unit of T4 DNA ligase were used. Cohesive end ligations were performed at 15°C for 8 hours and blunt end ligations at room temperature for 4-6 hours.

Preparation of Competent Cells

Escherichia coli competent cells were prepared according to the method of Dagert and Ehrlich (16). A single colony of JM109 was inoculated into 5 ml LB medium, and grown overnight at 37°C with shaking. Four ml of this culture were transferred to 400 ml of fresh LB medium and grown at 37°C with shaking. When the culture reached an OD_{590} of 0.4, it was dispensed into prechilled centrifuge tubes and left on ice for 10 minutes. The culture was centrifuged for 7 minutes at 3000 rpm at 4°C to pellet the cells. The cell pellet was suspended in ice-cold, sterile $CaCl_2$ solution (60mM $CaCl_2$, 15% v/v glycerol, 10mM PIPES, pH 7.0). The cells were centrifuged at 2500 rpm at 4°C and resuspended in the same solution. They were left on ice for 30 minutes. The cells were centrifuged again as explained in the earlier step. They were finally resuspended in a very small

volume of $CaCl_2$ solution. Aliquots (250µl) of competent cells were dispensed into prechilled, sterile tubes and stored at -70°C until needed.

Preparation of Plasmid DNA Templates

Plasmid DNA was extracted by the modified Birnboim-Doly alkaline-SDS method (6). Any required sequencing clone was grown in 250 ml LB medium containing ampicillin (50µg/ml) for 8-10 hours. The cells were pelleted by centrifugation. The pellet was resuspended in a solution containing 2M glucose, 0.5M EDTA, 1M Tris-HCl (pH 8.0) and 0.2% lysozyme, and incubated for 5 minutes at room temperature. Cells were denatured by adding alkaline-SDS solution (0.2N NaOH and 1% w/v SDS) and incubation on ice for 20 minutes after gentle mixing. The denatured mixture was then neutralized with 5M potassium acetate (pH 4.8), mixed and incubated on ice for 10 minutes. The mixture was centrifuged at 12,000 rpm for 30 minutes at 4°C. The supernatant was transferred to a clean tube and plasmid DNA was precipitated overnight on ice by adding 0.313 volumes of 42% polyethylene glycol. Plasmid DNA was collected by centrifugation. It was dissolved in sterile TE buffer and layered on cesium chloride-ethidium bromide gradient and centrifuged at 100,000 rpm for 4 hours at 20°C in a table top ultra-centrifuge (Optima TL, Beckman). After completing two such gradient centrifugations, supercoiled plasmid DNA was extracted and ethanol precipitated. It was dissolved in sterile TE buffer and was

quantified by measurment of A_{260} using UV spectrophotometer (UV-160A, Shimadzu).

Denaturation of Template

Approximately 1μg of plasmid DNA was denatured with a solution containing 0.2N NaOH and 0.2mM EDTA and incubated for 5 minutes at room temperature. It was neutralized with a reagent containing 2M ammonium ion and 5M acetate ion. The DNA was precipitated by adding 2.5 volumes of ethanol and storing at -20°C overnight. After pelleting, DNA was washed with 70% ethanol and vaccum dried. It was dissolved in double distilled water for immediate use. Otherwise, the DNA was stored as a dry pellet at -20°C until needed.

Annealing of Template to Primer

Both M13 reverse and forward primers (Promega) were used in DNA sequencing. A typical 25μl annealing reaction mixture contained a primer and plasmid template in an approximate 1:1 molar ratio. It also contained 5μl of *Taq* DNA polymerase buffer (50mM of Tris-HCl pH 9.0, 10mM $MgCl_2$), 2μl of extension/labeling mix (7.5μM each of dGTP, dTTP, dCTP) and the required amount of sterile distilled water to attain the final volume. Annealing was done

by incubating the mixture at 37°C for 10 minutes.

Extension and Labeling

After annealing, 2µl of [α-^{35}S] dATP (1000Ci/ml, 10µCi/µl, NEN) and 1.6µl of sequencing grade *Taq* DNA polymerase (2.5u/µl) were added to the annealing mixture and the new mixture incubated at 37°C for 5 minutes.

Termination

While the annealing reaction was incubating, nucleotide mix for termination reactions was prepared. For each set of sequencing reactions, 4 microcentrifuge tubes were labeled (G, A, T, C) and 1µl of respective deoxy/dideoxy (d/ddNTP) nucleotide triphosphate mix provided by the manufacturer (Promega) was added to each tube. The tubes were capped and set on ice until needed. When the extension and labeling reaction was complete, 6µl of the labeled reaction mixture were aliquoted to each of the 4 tubes containing d/ddNTP mix. The contents were mixed briefly by pipetting. The tubes were incubated at 70°C for 10 minutes. Then, 4µl of stop solution (10mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) were added to each tube and stored on ice. Before loading on a sequencing gel, the tubes were heated at 90°C for

5 minutes. Any excess mixture was stored at -70°C until needed.

Denaturing Gel Electrophoresis

A standard sequencing gel contained 7M urea and 5-6% polyacrylamide with a 19:1 Acrylamide/Bis ratio in 1x TBE (0.89mM Tris, 0.89mM Borate, 2mM EDTA). The gel apparatus was (Sequi-Gen Nucleic acid Sequencing Cell) purchased from Bio-Rad. All gels were run according to electrolyte-gradient method of Sheen and Seed (78). On polymerization, a gel was preelectophoresed using 0.5x TBE in the top and 1x TBE in the bottom chamber. After an hour, samples were loaded and the electrophoresis was continued at 60W constant power. When bromophenol blue in the samples ran off the bottom, 3M sodium acetate was added to bottom chamber buffer to give a final concentration of 1M. This increased the salt concentration and generated an effective gel gradient. After electrophoresis, the gel was processed in a fixer solution (5% acetic acid/5% methanol) for 15-20 minutes. It was blotted on a filter paper and dried on a preheated gel dryer at 80°C for 20-30 minutes. The dried gel was directly exposed to X-ray film (Kodak XAR-5) at room temperature. After sufficient exposure, the X-ray film was removed and processed according to the manufacturer's instructions.

Analysis of Nucleotide Sequences

Nucleotide sequences were stored as MacVector files and analysed for open reading frames using MacVector 3.5 software of International Biotechnologies, Inc., California. The sequences were compared for similarity with other DNA sequences in different data banks. The deduced amino acid sequences of the open reading frames were also compared for similarities with other deduced amino acid sequences in data bases. Comparisons were performed using the Sequence Analysis Software Package of Genetics Computer Group, Inc., Madison, Wisconsin. The BLASTP program (4) that was used to compare amino acid sequences showed alignment of amino acid sequences and gave maximal segment pair (MSP) score and also smallest poisson value. The smaller the poisson value the closer was the similarity between the query and the subject sequence. The program determined the order of identity between various sequences in the data banks with that of TfdR and TfdDII deduced sequences.

Expression of Proteins Using T7 RNA Polymerase/Promoter System

TfdR protein was expressed and labeled according to the method developed by Tabor and Richardson (84) and using the

vectors constructed by S. Tabor.

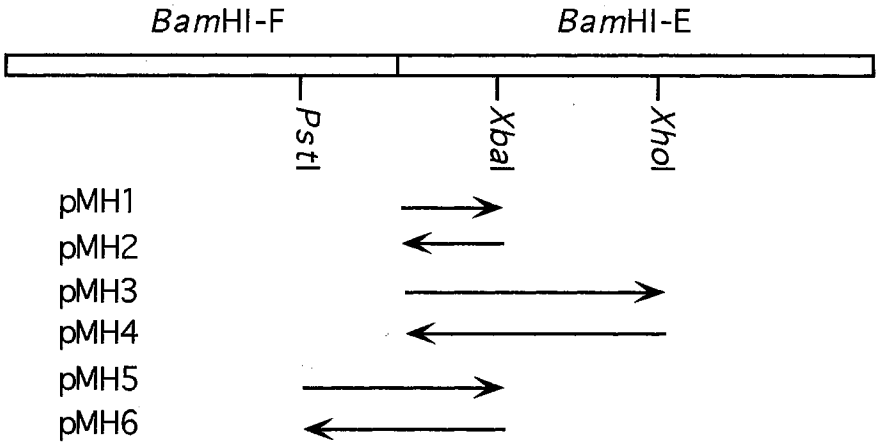
Construction of Expression Clones

Figure 4 shows the construction of expression plasmids. Vectors pT75 and pT76 contain a T7 promoter upstream of a polylinker sequence and an ampicillin resistance gene. Plasmid constructs were transformed into competent *E. coli* strain K38 already containing plasmid pGP1-2. Plasmid pGP1-2 contains T7 RNA polymerase gene under the control of the λ_{P_L}promoter that is repressed by a temperature sensitive repressor (cI857). The plasmid also carries a kanamycin resistance gene. Ligations, tranformation and preparation of competent cells were performed in the same way as explained in the section on sequencing of *tfdR*.

Selective Labeling of Plasmid Encoded Proteins

From each construct, two *E. coli* K38 transformants were used to selectively label proteins. Individual colonies were picked with sterile tooth picks and inoculated into 1ml M9 medium containing 25μg of ampicillin and 25μg of kanamycin. In addition, M9 medium was supplemented with 0.005% of the 18 amino acid mixture (18 amino acid mixture contained 0.1% w/v of all the amino acids except methionine and cystine). After an overnight growth at 30°C, a

Figure 4. Construction of Expression Clones. Plasmids pMH1, pMH3 and pMH6 were constructed by inserting the respective fragments into pT7-5 as shown in the figure. The others were cloned into pT7-6. The arrows indicate the direction of T7 promoter on the vector with respect to the inserted fragment.


BamHI-F
BamHI-E
PstI
XbaI
XhoI
pMH1
pMH2
pMH3
pMH4
pMH5
pMH6

culture was diluted 1:40 with fresh M9 medium (supplemented with antibiotics and amino acids) and grown until it reached an OD_{590} of 0.4. It was shifted to 42°C for 20 minutes to induce expression of T7 RNA polymerase. Rifampicin (200mg/ml) was added to the culture and incubation at 42°C was continued for 10 more minutes (rifampicin inhibits host RNA polymerase activity). The culture was then moved into a 30°C water bath shaker for 20 minutes. Newly synthesized proteins were labeled with 10mCi of [^{35}S]methionine added to 0.5 ml of the culture and after incubating for 5 more minutes at 30°C. The cells were pelleted by spinning in a microfuge for 20 seconds (14,000 x g). They were resuspended in SDS/sample buffer and boiled at 100°C for 5 minutes before loading on a gel. The SDS/sample buffer was composed of 0.125M Tris-phosphate, pH 6.8, 2% SDS, 5% 2-mercaptoethanol, and 10% glycerol.

One-Dimensional Denaturing Electrophoresis

Polyacrylamide gel electrophoresis was performed according to Laemmli (50). The slab gel with 12% separating and 4% stacking polyacrylamide gels was cast containing 0.1% w/v SDS. The separating gel was prepared with 0.375M Tris-HCl (pH 8.8) and the stacking gel with 0.125M Tris-HCl (pH 6.8). Electrophoresis buffer contained 0.025M Trizma base, 0.2M glycine and 0.1% w/v SDS. Samples were run at 10mA constant current until the bromophenol dye entered the separating gel. Then the current was increased to 15mA. The gel was processed in 5% methanol/5% glycerol (w/v)

drying solution after completion of electrophoresis. It was blotted on a filter paper and dried on a prewarmed gel dryer at 80°C for 15 minutes. The dried gel was exposed to X-ray film (Kodak XAR-5) at room temperature. After sufficient exposure the film was processed according to the manufacturer's instructions.

Southern Blotting and Hybridization

In order to determine the presence of additional copies of *tfdR* gene on pJP4, a Southern hybridization was performed on pJP4 with a *tfdR* DNA probe, according to the method by Southern (79).

Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed as described by Maniatis et al. (52). Restriction digests of plasmid pJP4 (*Bam*HI, *Eco*RI, *Hin*DIII) were separated on a 1.0% agarose gel. The agarose gel was prepared and electrophoresed in 1x Tris-acetate-EDTA buffer (pH 7.4, 40mM Tris-acetate, 2mM EDTA).

Gel Pretreatment and Transfer to Nylon Membrane

After electrophoresis, DNA was depurinated by soaking the gel in 0.25M HCl twice for 15 minutes each. It was denatured by soaking the gel twice for 15 minutes each in a solution containing 0.5M NaOH and 1.5M NaCl. The DNA was neutralized in 0.5M Trizma base and 1.5M NaCl (pH 7.0) twice again for 15 minutes each. The DNA was transferred to a Nylon membrane (Zeta probe GT, BioRad) with a vaccum blotter (MilliBlot-V, Millipore) using as transfer solution 20x SSC (3.0M NaCl and 0.3M sodium citrate).

Hybridization

After transfer, the membrane was briefly rinsed in 2x SSC and air dried. It was then dried in a vaccum-oven at 80°C for 30 minutes. Prehybridization was started by sealing the membrane into a heat sealable plastic bag containing 0.25M Na_2HPO_4 (pH 7.2) and 7% w/v SDS hybridization solution and incubating the bag at 65°C for 5 minutes. One corner of the bag was cut and the solution replaced with fresh hybridization solution. Heat denatured DNA probe was then added into the bag and all air bubbles were removed before resealing it. Hybridization was conducted by incubating the bag at 65°C for 16 hours with agitation. The membrane was removed out of the bag and washed in 20mM Na_2HPO_4 (pH 7.2) and 5% w/v SDS at 65°C twice for 30 minutes each. It was washed again twice in 20mM

Na_2HPO_4 (pH 7.2) and 1% w/v SDS at 65°C for 45 minutes each. After washing, the wet membrane was exposed to X-ray film with a plastic wrap in between. The film was processed according to instructions by the manufacturer (Kodak).

Preparation of DNA Probe

The DNA probe used in Southern hybridization consisted of 645 bp *Bam*HI-*Xba*I region on *Bam*HI-E. This region included the first 542 bp of *tfdR* gene and its upstream sequences. The DNA was labeled by nick translation.

Nick Translation

Nick translation was performed with a nick translation kit purchased from Promega. A typical reaction mixture contained 1μg of the DNA, 10μl of nucleotide mix (5μM each of cold dCTP, dGTP, dTTP), 5μl of nick translation buffer (50mM Tris-HCl, pH 7.2 and 10mM $MgSO_4$), 7μl of [α-^{32}P] dATP (400Ci/mM at 10mCi/ml), 5μl of enzyme mix (DNA polymerase I at 1U/μl and DNase I at 0.2ng/μl) and sterile water to attain a final volume of 50μl. The reaction mixture was incubated at 15°C for 1 hour. The reaction was stopped by adding 5μl of 0.2M EDTA (pH 8.0). Unincorporated label was removed by selectively precipitating the DNA with ammonium

acetate and ethanol. The DNA was pelleted after a 70% ethanol wash and dissolved in TE (1M Tris-HCl, 0.5M EDTA, pH 8.0) buffer. The radiolabeled probe was stored at -70°C.

Protein-DNA Interactions

TfdR protein from a cell crude extract was checked for its interaction with the DNA region upstream of the *tfdA* gene on *Bam*HI-B. The same region is present upstream of the *tfdR* gene on *Bam*HI-E as well. This region shows similarity to promoter sequences of *catA*. Protein-DNA binding was demonstrated by mobility shift polyacrylamide gel electrophoresis method of Fried and Crothers (29).

Preparation of Cell Crude Extract

A 500ml culture of the *E. coli* K38(pGP1-2) strain containing plasmid pMH6 was grown in LB with ampicillin and kanamycin (each at 50μg/ml) at 30°C for several hours. When the culture reached an OD_{590} of 0.6, it was heat-induced at 42°C for 30 minutes. It was then transferred to 37°C for 90 minutes to synthesize plasmid encoded proteins. The cells were harvested and washed with 0.9% w/v NaCl solution. They were resuspended in 1ml of a solution containing 10mM Tris-Cl (pH 8.0), 1mM of EDTA and 0.1mM of phenylmethyl-

sulfonyl fluoride. Cells were disrupted using a sonicator (Sonifier Cell Disruptor 350, Branson Sonic Power Co.). The sonicated suspension was centrifuged to remove cell debris and insoluble fractions. Supernatant containing soluble proteins was saved at -70°C until needed.

Quantification of Protein

The concentration of protein in the cell extract was determined by the method described by Bradford (8). Dye reagent was obtained from Bio-Rad and used according to their instructions. Bovine serum albumin was used as the standard and the colorimetric readings were recorded at 595nm with a spectrophotometer (UV-160A, Shimadzu).

End Labeling of DNA

A 315bp *Xba*I-*Sph*I region was used in the binding reaction. It was end-labeled using Klenow polymerase. The end-labeling reaction was performed in the same buffer condition after plasmid pHU12 was digested by *Sph*I and *Xba*I. Klenow polymerase was added along with 5mM each of cold dGTP, dCTP and dTTP and 10μCi of [α-^{32}P]dATP. The contents were mixed and incubated at room

temperature for 15-30 minutes. After removing unincorporated label by repeatedly precipitating DNA by ammonium acetate and ethanol, the end-labeled DNA was washed in 70% ethanol and dried under vaccum. It was dissolved in sterile water and stored at -70°C until needed. The amount of radiolabel incorporated was measured in a scintillation counter (LS 6000SC, Beckman).

DNA-Protein Binding Reaction

The following conditions were used in binding reactions. The binding assay was performed in a total volume of 15µl of 10mM HEPES buffer (pH 7.9) containing 10% glycerol, 100mM KCl, 4mM spermicidine, 0.1mM EDTA, 0.25mM dithiothreitol, 2mM $MgCl_2$, 1.5µg of bovine serum albumin, and 1µg of salmon sperm DNA. Typically between 0.1 to 1.0µg of crude cell extract protein was used in the assay. Approximately 15,000 cpm of the end labeled DNA fragment was used in each assay. Binding reactions were performed at 30°C for 15 minutes. The samples were immediately electrophoresed. For negative control, cell crude extract prepared from *E. coli* K38 clone containing only the vector pT7-5 without insert, was used in separate binding reactions.

Mobility Shift Gel Electrophoresis

A high ionic strength polyacrylamide gel was used as described by Staudt et al., (81). A 5% polyacrylamide gel with 2.5% glycerol was prepared in Tris-glycine-EDTA buffer (50mM tris, 0.38M glycine and 2mM EDTA). The gel was preelectrophoresed for 60 minutes. The samples were loaded and electrophoresis was continued at 100V constant, until bromophenol blue dye from the marker lane reached the bottom of the gel. The gel was transferred to a filter paper and dried on a prewarmed gel dryer at 80°C for 15 minutes. The dried gel was exposed to X-ray film and processed according to the manufacturer's (Kodak) instructions after sufficient exposure.

RNA Blotting and Northern Hybridization

Transcription of *tfdR* was determined by northern hybridization of RNA isolated from *Pseudomonas* PAO1c clones containing derivatives of pJP4. RNA were isolated according to the method of Reddy et al. (66) and hybridization was performed using the method of Thomas (85).

Preparation of RNA

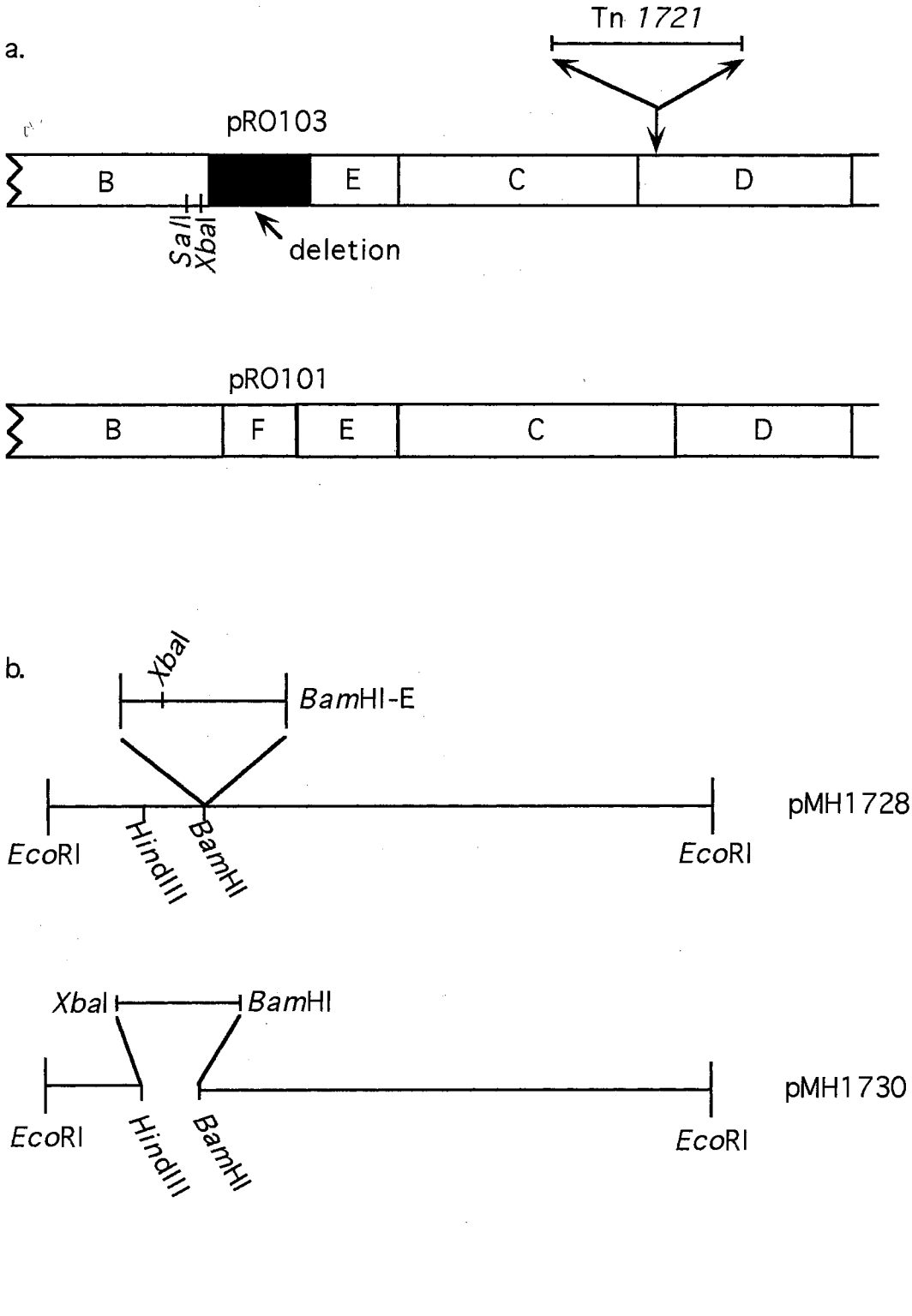
Pseudomonas aeruginosa PAO1c clones with pRO101, pRO103, pHM1728, pRO103/pHM1728 and pRO103/pHM1730 plasmid constructs were used for RNA isolation. Plasmids pRO101 and pRO103 were derived from pJP4 by Harker et al. (1989). The plasmid pRO103 a spontaneous mutant of pRO101 constitutively expresses *tfdA* gene. Plasmid pRO103 has a deletion of 3.6 kb DNA. The deleted DNA includes the entire *Bam*HI-F and portions of *Bam*HI-E and B fragments on pJP4. Partial physical maps of the two plasmids are given in Figure 5a.

RNA was also isolated from PAO1c transformants containing pRO103/pHM1728 and pRO103/pHM1730. Plasmid pHM1728 was constructed by inserting the *Bam*HI-E fragment of pJP4 into the *Bam*HI site of plasmid pRO1727 (refer Fig. 5b). The double plasmid construct restored inducibility of *tfdA* gene expression in pRO103. The construction of plasmid pHM1730 is explained in Figure 5b.

All PAO1c strains were grown in 100 ml of MMO containing either TFD (0.05%) or casamino acids (0.3%) as sole carbon source, at 37°C. When the cells reached log phase of growth 0.05 volumes of stop buffer (200mM Tris-HCl, pH 8.0, 20mM EDTA, 20mM sodium azide) were added to culture and placed on ice. The cells were harvested and suspended in 2 ml of lysis buffer (8% sucrose, 5% Triton X-100, 50mM EDTA and 50mM Tris-HCl pH 7.0) and 100µl of vanadyl-ribonucleoside complex (VRB, GIBCO/BRL). The lysed suspension was transfered to polypropylene tubes containing 1 ml of

Figure 5a. Partial Physical Maps of Plasmids pRO101 and pRO103. The *Bam*HI fragments representing the catabolic region on pJP4 are shown. The inserted position of *Tn1721* is also shown. The deletion on pRO103 is indicated.

5b. Construction of *Bam*HI-E Sub-Clones. Plasmid pMH1728 was constructed by insertion of *Bam*HI-E of pJP4 into pRO1727. Plasmid pMH1730 was created by deletion of *Hin*DIII-*Xba*I region of pMH1728, as indicated.


a.
Tn 1721
pRO103
B
E
C
D
SalI
XbaI
deletion
pRO101
B
F
E
C
D
b.
XbaI
BamHI-E
pMH1728
EcoRI
HindIII
BamHI
EcoRI
XbaI
BamHI
pMH1730
EcoRI
HindIII
BamHI
EcoRI

buffered phenol. The contents were vortexed vigorously for 1 minute. One milliliter of chloroform was added and the suspension revortexed for another minute. The suspension was centrifuged at 8500 rpm (10,000 x g) in a SS-34 for 10 minutes at 4°C. Upper aqueous phase was removed and nucleic acids were pelleted by centrifuging for 10 minutes at 8500 rpm (10,000 x g) in a SS-35 after adding 0.1 volumes of 3M sodium acetate and 2 volumes of 100% ethanol. The pellet was resuspended in 2 ml of 10mM VRB. Phenol/chloroform (1:1) extraction was repeated and nucleic acids were pelleted again as earlier. This pellet was suspended in 2 ml of diethyl pyrocarbonate (DEPC) treated water. One gram of cesium chloride was added to it. The contents were layered on 0.75 ml CsCl cushion (5.7M CsCl in 100mM EDTA, pH 7.0) in a polycarbonate tube. It was centrifuged at 80,000 rpm (280,000 x g) for 1 hour at 20°C using a TLA-100.3 rotor in a table top ultracentrifuge. The RNA pellet was recovered by carefully removing the DNA interface and other supernatant. The RNA was resuspended in DEPC-treated water and precipitated with sodium acetate and ethanol. The precipitated pellet was washed with 70% ethanol and dissolved in DEPC-treated water. RNA from each preparation was quantified spectrophotometrically. Samples were stored at -70°C.

Formaldehyde Gel Electrophoresis of RNA

A 1.2% agarose gel containing 2.2M formaldehyde in RNA borate buffer (3.2mM tetraborate, 30mM borate and 0.2mM

Na_2EDTA, pH 8.0) was used as described by Diamond (18). RNA samples contained 2 volumes of RNA loading buffer (65% formamide, 58% formaldehyde, RNA borate buffer, 0.6µg ethidium bromide, 0.6µg xylene cyanol and 0.6µg bromophenol blue). Samples were heated at 65°C, cooled on ice for 5 minutes before loading on the gel and electrophoresed until xylene cyanol moved to the middle of the gel.

RNA Transfer to Membrane

After electrophoresis the RNA was transferred to a nylon membrane in 2x SSC using a vaccuum blotter. On completion of transfer RNA was UV cross-linked to the membrane using UV-cross linker (Stratagene).

Hybridization

Prehybridization and hybridization were conducted in the same way as described for Southern hybridization. RNA probe was hybridized with RNA on the membrane for 16 hours at high stringency (65°C). The membrane was washed in the same manner as described in Southern hybridization. The air dried membrane was exposed to X-ray film. After enough exposure the film was processed according to the manufacturer's instructions (Kodak).

Preparation of RNA Probe

An RNA probe was made using a riboprobe kit purchased from Promega. Plasmid pHU13 containing the *Bam*HI-*Xba*I region of the *Bam*HI-E fragment was used as template DNA for transcribing the riboprobe. The region contained the first 542 bp of *tfdR* and its upstream sequences. Plasmid pHU13 was linearized with *Bam*HI before use in transcription. A standard transcription reaction contained approximately 0.4 mg/ml linearized template DNA, 10mM DTT, 20 units ribonuclease inhibitor, 2.5mM NTP mix (ATP, GTP, TTP), 12μM CTP and 50μCi of [α-^{32}P]CTP suspended in the transcription buffer (40mM Tris-HCl, pH 7.5, 6mM $MgCl_2$, 2mM spermidine, 10mM NaCl). Transcription was performed by adding 15-20 units of SP6 RNA polymerase and incubating at 40°C for 60 minutes. Labeled RNA transcripts were separated from unincorporated radioactive CTP by repeated ammonium acetate/ethanol precipitations. Precipitated riboprobe was resuspended in nuclease-free water. The probe was complementary to *tfdR* gene transcript. The probe was heat denatured and ice cooled before use in hybridization.

CHAPTER IV

RESULTS

DNA Sequencing

Nucleotide Sequences

The complete nucleotide sequence of the regulatory gene *tfdR* of plasmid pJP4 was determined. It is presented in Figure 6 along with the deduced amino acid sequence of the protein. The gene is 888 bp long and encodes a protein composed of 295 amino acids.

The complete gene was not present in the 1.2 kb *Bam*HI-*Xho*I region on the *Bam*HI-E fragment of pJP4. Analysis of nucleotide sequences revealed two incomplete open reading frames (ORFs) within this region. These ORFs were oriented divergently. Figure 7 gives a physical map of restriction fragments *Bam*HI-E, *Bam*HI-F and *Bam*HI-B on pJP4 and also orientations of these two ORFs within the 1.2 kb region. It also gives respective locations of genes *tfdA* and *tfdS* on *Bam*HI-B.

Preliminary comparison of the 1.2 kb DNA sequence showed 100% identity of the 645 bp *Bam*HI-*Xba*I region on *Bam*HI-E with a

Figure 6. Nucleotide Sequence of *tfdR* and Surrounding Regions of pJP4. The sequences shown include the region between *Xba*I site on *Bam*HI-E to the end of *tfdR* on *Bam*HI-F (refer Fig.7). The codons and deduced amino acid sequence of *tfdR* gene are shown. The putative ribosome binding site is shown in bold.

XbaI

TCTAGAAATGCCGGGCCATATGGTCTAATA

CCTTCCATCAGCGCGAGCGATACGGGATCGGTATGGCTCTCCGGCAGCCGGAGCCAGCCACC

A**GGAG**TGAAGA ATG GAG TTT CGA CAG CTT CGC TAT TTC GTT GCT GCC GCG

M E F R Q L R Y F V A A A

GAG GAG GGC AAC GTC GGT GCC GCC GCG CGG CGG CTG CAT ATT TCC CAG

E E G N V G A A A R R L H I S Q

CCC CCG GTC ACG CGA CAG ATT CAC GCG CTC GAA CAG CAT CTG GGC GTG

P P V T R Q I H A L E Q H L G V

TTG TTG TTC GAG CGC AGC GCG CGC GGC GTG CAG CTC ACG CCC GCC GGG

L L F E R S A R G V Q L T P A G

GCC GCG TTT CTC GAA GAT GCA CGG CGC ATG CTC GAA CTG GGT CGG ACT

A A F L E D A R R M L E L G R T

TCC GTG GAC CGG TCC CGC GCC GCC AGC CGG GGC GAG ATC GGC CAA CTC

S V D R S R A A S R G E I G Q L

GAT ATC GGC TAC CTC GGC ACG GCG ATC TAC CAG ACC GTC CCG GCA TTG

D I G Y L G T A I Y Q T V P A L

CTC CAT GCG TTC ACG CAG GCG GTC CCG GGG GCG ACG CTG TCT CTG GCC

L H A F T Q A V P G A T L S L A

CTG ATG CCC AAG GTG CGG CAG ATC GAG GCC CTG CGT GCC GGC ACC ATC

L M P K V R Q I E A L R A G T I

CAT CTC GGT GTC GGC CGC TTC TAC CCC CAG GAG CCT GGA ATC ACG GTG

H L G V G R F Y P Q E P G I T V

GAG CAC CTG CAC TAC GAA CGG CTG TAT ATC GCA GCG GGT TCG AGC ATT

E H L H Y E R L Y I A A G S S I

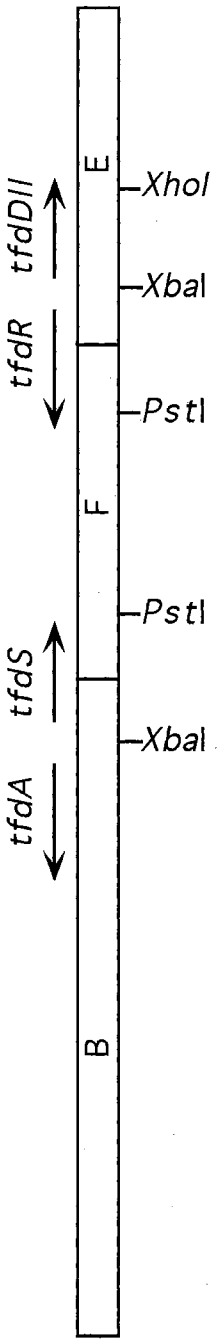
GCG CGC CAG CTG AGA CAG GAT CCG ACG CTG CTG CGG CTC AAG AGC GAG

A R Q L R Q D P T L L R L K S E

Figure 6. Continued

```
TCC CTC GTT CTT TTC CCC AAG GAG GGG AGG CCG AGT TTC GCT GAC GAA
 S   L   V   L   F   P   K   E   G   R   P   S   F   A   D   E
GTG ATC GCC TTG ATG CGC CGG GCC GGG GTC GAG CCG CGC GTG ACG GCG
 V   I   A   L   M   R   R   A   G   V   E   P   R   V   T   A
ATT GTC GAA GAT GTC AAC GCG GCC CTC GGG CTC GTC GCG GCC GGC GCC
 I   V   E   D   V   N   A   A   L   G   L   V   A   A   G   A
GGC GTC ACG CTG GTC CCG GCC TCG GTG GCC GCG ATT CGG CGG CCC TTC
 G   V   T   L   V   P   A   S   V   A   A   I   R   R   P   F
GTC CGG ACG ATG GAG ATG GCC GAT GCG AGC GAC AAG GTG CCG GTC AGC
 V   R   T   M   E   M   A   D   A   S   D   K   V   P   V   S
CTG ACT TAC CTG ACC GAC TCT CGC GTA CCC GTG CTT CGC GCA TTT CTC
 L   T   Y   L   T   D   S   R   V   P   V   L   R   A   F   L
GAT GTC GCA AGA CGC GGG AAA GGA CAG AAA TAG CGC GCT
 D   V   A   R   R   G   K   G   Q   K   *
```

Figure 7. Organization of Restriction Fragments *Bam*HI-B, F and E of pJP4. The positions of genes *tfdR*, *tfdDII*, *tfdS* and *tfdA* are shown. The arrows indicate the direction of open reading frames.


E
F
B
XhoI
XbaI
PstI
PstI
XbaI
tfdDII
tfdR
tfdS
tfdA

region in *Bam*HI-B of pJP4. Published documents have proposed that this region of *Bam*HI-B contains a portion of gene *tfdS*. As shown in Figure 7, *tfdS* is encoded divergently from the *tfdA* gene on *Bam*HI-B. Therefore, the incomplete ORF located within 645 bp region on *Bam*HI-E fragment may represent a portion of the *tfdR* gene to encode a regulatory protein like *tfdS*. This speculation was confirmed when the remaining sequenced nucleotides on the 1.2 kb region were analysed (see below).

DNA sequence between *Xba*I and *Xho*I sites on *Bam*HI-E contains another incomplete ORF in the orientation opposite to the one assigned to *tfdR*. This ORF shows similarity with that of *tfdD* gene encoding chloromuconate cycloisomerase in the *tfdCDEF* operon of pJP4. The similarity suggests that *Bam*HI-E fragment contains an isofunctional *tfdD* gene. Therefore, the ORF on *Xba*I-*Xho*I region has been designated as part of *tfdDII* gene. The position of *tfdDII* ORF on *Bam*HI-E is indicated in Figure 7. The DNA and deduced protein sequences of the incomplete *tfdDII* are shown in Figure 8.

To determine the entire nucleotide sequence of genes *tfdR* and *tfdS*, both ends of the *Bam*HI-F fragment of pJP4 were separately cloned and sequenced (refer Fig. 3). On combining the sequences on *Bam*HI-F adjacent to *Bam*HI-E/F junction with that of the upstream sequences on the 645 bp *Bam*HI-*Xba*I region of *Bam*HI-E fragment the entire nucleotide sequence of *tfdR* gene was obtained. Unpublished portions of *tfdS* gene on *Bam*HI-F were determined from the end sequences on *Bam*HI-F adjacent to the *Bam*HI-B/F junction on pJP4. It was found that both *tfdR* and *tfdS* are 100% identical in DNA sequence. The genes are positioned as inverted

Figure 8. Nucleotide Sequence of *tfdDII* on *Xba*I-*Sph*I region of *Bam*HI-E of pJP4. The non coding strand of *tfdDII* gene and its deduced amino acid sequence are shown. Putative ribosome binding site is indicated in bold.

XbaI

AGATCTGATGGCGGTACGAGTGTCTTTTTCGGTAA

CGGCTGTCGGGGTTGCCGCCGCTGCGCTAGCGTTCCGCCGCGTCTAGCTTCGCTAG**CTTT**GC

CAC TAG CAG CTA GAC GGC GAC GCG GCG TAG GTC GTC AAG CGG GCG GAC

V I V D L P L R R I Q Q F A R L

CCG CGG TTC GTA GTC TCG TCG CAC GAG TAG GCA GAG GTA TGC TTC CCG

G A K H Q S S V L I R L H T K G

CCG TAG CAG CCC TAG CCA CTC AGG TAG TGC GGG ACG CCA GGG ACC ACC

G I V G I G E S I T P C G P W W

TCG CCG CTG TCG CAG CTT CGG TAA GTT CGG TGC TAG TTG GTG ATG GAC

S G D S V E A I Q A T I N H Y L

CCC GGA GAG CAG CAG CCG CTT GGC CGC GAG CTA CGC AGC GCG TAG TAC

G P L V V G E P A L D A S R I M

CGC TTG AGG TAC CGG CGC ACC GGC CGT TAC GGA AGC TGT TCC GGC CGT

A N S M A A W P A M P S T R P A

AGC TCT ACC GTG ACG ACC TGC GCC AGC CGT TCT AGC AGC TAC GGG GCT

S R W H C W T R S A R S S M P R

AGG TGC AGA ACG ACC CTC CTG CGA AAG CGC TGG CCG AGT CGC AGC GTA

S T S C W E D A F A T G S A S H

CCG GCG AGC TGG TGC CCG CTG CAC TTG GTC CTC CAG CTG CTC CGC AAG

G R S T T G D V N Q E V D E A F

GCG TAC GAC CTC CGT CCG TTC CGG CCC CGG AAG TTT GAC TTC TAC CCG

R M L E A G K A G A F K L K M G

CGC GAC GGC GAC CGC GTT CTG GAC GCC GCA CGT AAC CGC TAG CGC TTC

A L P L A Q D L R R A L A I A K

CTT TGA GCT CCC GTT CTG CTC GGA CGC GGC AGC TAG GGT TGC TTC GCA

E T R G Q D E P A P S I P T K R

CCC TGC TCG GCT GGT GCT ACG CCG TGA CGC GGG GAC CTC CGC CGG CGC

G T S R P R C G T A P L E A A A

ACC TTT AAG TAG CTC GTC GGG CAG CGG GCG ACC TTA GAG CTA CGC TAC

W K F I E Q P V A R W N L D A M

CGG GCG TAG GTG GCC GTT CGC GCT TCG TAC G

A R I H R Q A R S M>

repeats on plasmid pJP4 and separated by about 2 kb of DNA.

The DNA sequences reported in this thesis have been deposited in GenBank, under the accession number M98445.

Similarity of *tfdR* gene to *lysR* Family Members

A computational database search was done with both DNA and deduced amino acid sequences of *tfdR* gene. Comparison of the deduced amino acid sequence of TfdR using BLASTP revealed clear similarities with members of *lysR* family. A total of 83,441 sequences and 23,430,555 residues were searched for similarity. Assuming that the *tfdR* gene starts at position 104, of the sequences shown in Figure 6, the similarity was highest in the N-terminal helix-turn-helix motif which is presumed to be the DNA binding region of LysR protein. More identity was observed with amino acid sequences deduced from regulatory genes involved in aromatic hydrocarbon metabolism, such as those for *tcbR* of *Pseudomonas* sp strain P51 (52.6% identity over 291 amino acids), *catM* of *Acinetobacter calcoaceticus* (36.3% identity over 248 amino acids), *catR* of *Pseudomonas putida* (35.8% identity over 288 amino acids) and *nahR* of *P.putida* (20.2% identity over 248 amino acids). The proteins TcbR, CatM, CatR and NahR recorded maximal segment pair (MSP) scores of 748, 144, 149 and 48 respectively in the BLASTP search. The highest score in the search was for TfdS protein (932)

having 100% identity to TfdR. The cut off MSP score in the search was 43.

The computational analysis also revealed similarities with amino acid sequences translated from portions of open reading frames present upstream of *clcABD* gene clusters (80% identity over 72 amino acid overlap) and the *tfdCDEF* operon (65% identity over 47 amino acids) of *E. coli* strain JM103 and *Alcaligenes eutrophus* JMP134 respectively. The two putative ORFs are called *clcR* and *tfdX* respectively. In the BLASTP search the maximal segment pair score for the incomplete amino acid sequence of TfdX was 169. Figure 9 illustrates homology between TfdR and the deduced amino acid residues of ClcR and TfdX. The alignment shown in the figure is from the out put file of BLASTP search.

At the DNA level, the *tfdR* gene showed 63% (over 507 bp) identity with that of *tcbR* gene and 72% identity (over 215 bp) with the partial ORF upstream of *E.coli clcABD* gene clusters.

Similarity of *tfdDII* to Chlorocatechol Degrading Genes

The 794 bp sequence extending between *Xba*I-*Sph*I on *Bam*HI-E fragment, containing the incomplete *tfdDII* ORF, was analysed for similarity. The deduced 234 amino acid sequence of *tfdDII* ORF showed similarity with other deduced amino acid sequences from protein data bank when compared using BLASTP

Figure 9. BLASTP Comparison of TfdR, ClcR and TfdX Deduced Amino Acid Residues. The ClcR and TfdX sequences are partial sequences translated from the nucleotide sequences upstream of the *clcABD* and *tfdCDEF* operons respectively. The identical residues between TfdR and ClcR or TfdX are shown in middle line of the respective pair. Symbols: "+" refers to conserved positive changes; "*" refers to strongly conserved helix-turn-helix motif residues of the proteins. The numbers indicate the positions of the amino acid residues.

```
                          ** ************* ****
TfdR       1 MEFRQLRYFVAAAEEGNVGAAARRLHISQPPVTRQIHALEQHLGVLL 47
             MEFRQLRYF+A+AEEGN+GAAARRLHISQPP+TRQI+ALEQ+LGV+L
ClcR         MEFRQLRYFIAVAEEGNIGAAARRLHISQPPITRQIQALEQDLGVVL

TfdR         FERSARGVQLTPAGAAFLEDARRMLELGR          76
             FER+ RGV+LT+AG++FLEDARR+L++
ClcR         FERTHRGVELTAAGTTFLEDARRLLHVTE          76


                          *  * ** * ****** ****
TfdR       1 MEFRQLRYFVAAAEEGNVGAAARRLHISQPPVTRQIHALEQHLGVLL 47
             ME RQL+YFVA AE G  G+AA+R+HISQPP+TRQI+ALE+ +G  L
TfdX       1 MEIRQLKYFVAVAEAGGFGTAAQRMHISQPPLTRQIQALERDIGAKL 47
```

alignment. A total of 87,423 sequences and 24,167,874 residues were searched. The highest MSP score obatined was 121 and the cut off score was 43. Assuming that the *tfdDII* gene starts at position 98 of the sequences shown in Figure 8, the identity was high with *tcbD* of *tcbCDEF* gene clusters of *Pseudomonas* sp P51 (28.6% identity over 231 amino acids, MSP of 121), *catB* of *P.putida* (27.8% identity over 237 amino acids, MSP of 112), *tfdD* of *tfdCDEF* operon in *A. eutrophus* (26.6% identity over 229 amino acid, MSP of 90), *clcB* of *clcABD* operon in *P.putida* (26.6% identity over 231 amino acids, MSP of 106), *catB* of *cat* operon of *A. calcoaceticus* (26.1% identity over 230 amino acids, MSP of 92). Genes *tcbD*, *tfdD*, *clcB* and *catB* (*A. calcoaceticus*) encode cycloisomerases (E.C-number 5.5.1.7) and gene *catB* of *P. putida* encodes *cic,cis*-muconate lactonizing enzyme (E.C-number 5.5.1.1). FASTA alignment of *tfdDII*, *tcbD* and *tfdD* cycloisomerases is given in Figure 10. The pair-wise alignment of TfdDII and TfdD or TcbD amino acid sequences obtained from the FASTA search were arranged into a mutiple sequence format as shown in the figure.

At the DNA level *tfdDII* gene shared homology with that of *clcB*, *tfdD* and *catB* (*P.putida*) genes having 57.1%, 57.6% and 56.8% identity over 778 bp, 738 bp and 746 bp respectively.

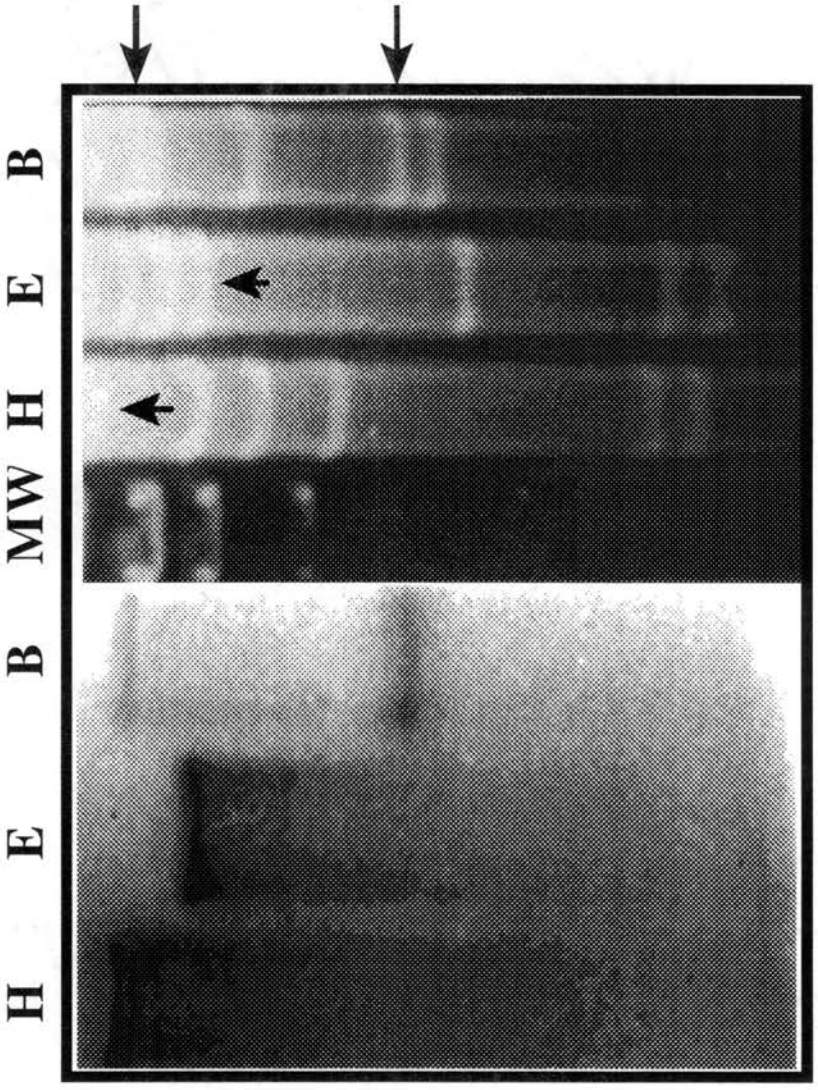
Southern Hybridization

Plasmid pJP4 was checked for any other copies of genes *tfdR* or *tfdS*. A Southern blot of three different restriction digests of pJP4

Figure 10. Alignment of Deduced Amino Acid Sequences of TfdDII and Other Chloromuconate Cycloisomerase. The alignment was based on TFASTA. Symbols: "I" identical amino acid residues; ':" conserved changes and "-" gaps for optimal alignment.

```
TfdDII          VIVDLPLRRIQQFARLGAKHQSSVLIRLHTKGGIVGIGESITPCGPWWSG
                 |||:| ||  |::  ::::|| |:::::: ||:|||||: :  || |::
TcbD    MKIEAISTTIVDVPTRRPLQMSFTIVHKQSYVIVQVKA-GGLVGIGEGGSVGGPIWGS
                ||||:| :|  |::  ::::|| |::|::: :|:||:||: :  || ||:
TfdD    VKIDAIEAVIVDVPTKRPIQMSITIVHQQSYVIVRVYS-EGLVGVGEGGSVGGPVWSA

TfdDII  DSVEAIQATINHYLGPLVVG--EPALDASRIMANSMAAWPAMPSTRPASRWHCWIRSA
        :|:|:|:::|::||:||:||  :::|:::|:: :: ::  :  |:::|      : :|
TcbD    ESAETIKVIIDNYLAPLLVGKDASNLSQARVLMDRAVT--GNLSAKAAIDIALHDLKA
        :::|:|: ::::||:| ::|: |:::|  :::   |  :: :|:::|      : :|
TfdD    ECAETIKIIVERYLAPHLLGTDAFNVSGALQIMARAV__TGNASAKAAVEMALLDLKA

TfdDII  RSSMPRSTSCWEDAFATGSASHGRSTIGDVNQEVDEAFRMLEAGKAGAFKLKMGALPL
        |:   : ::  :::: |::      ::||:::::|:|: |:|: : : ||:|:|| :
TcbD    RALNLSIADLIGGIMRTSIPIAWILASGDTARDIDSALEMIETRRHNRFKVKLGARTP
        |:   : ::  :::: :::      ::||:::::|:|  |:|  : : ||:|:|  :
TfdD    RALGVSIAELLGGPLRSAIPIAWILASGDTKRDLDSAVEMIERRRHNRFKVKLGFRSP

TfdDII  AQDLRRALAIAKETRGQDEPAPSIPIKRGTSRPRCGTAPLEAAAWKFIEQPVARWNLD
        |||| :  :|:|:: ::::   ::   :::  ::  :::||:|: :::|||||:| |::
TcbD    AQDLEHIRSIVKAVGDRASVRVDVNQGWDEQTASIWIPRLEEAGVELVEQPVPRANFG
        ::|| :  |:::: :::    ::   :::  ::  :: ||| : ::|||||:| | :
TfdD    QDDLIHMEALSNSLGSKAYLRVDVNQAWDEQVASVYIPELEALGVELIEQPVGRENIQ

TfdDII  AMARIHRQARSM.....>........>...........   236
        |: |:  |
TcbD    ALRRLTEQNGVAILADESLSSLSSAFELARDHAVDAFS
        |: |:  ::|
TfdD    ALRRLSDNNRVAIMADESLSTLASAFDLARDRSVDVFS
```

Figure 11. Agarose Gel Electrophoresis and Southern Hybridization of pJP4 Restriction Fragments. Panel on the right shows the agarose gel electrophoresis of *Hin*DIII, *Eco*RI and *Bam*HI restriction fragments. Panel on the left shows Southern hybridization of *Hin*DIII-B, *Eco*RI-E, *Bam*HI-B and E fragments with a *tfdR* gene probe. The arrows indicate the position of the four fragments on the agarose gel. MW refers to molecular weight standard.


H E B MW H E B

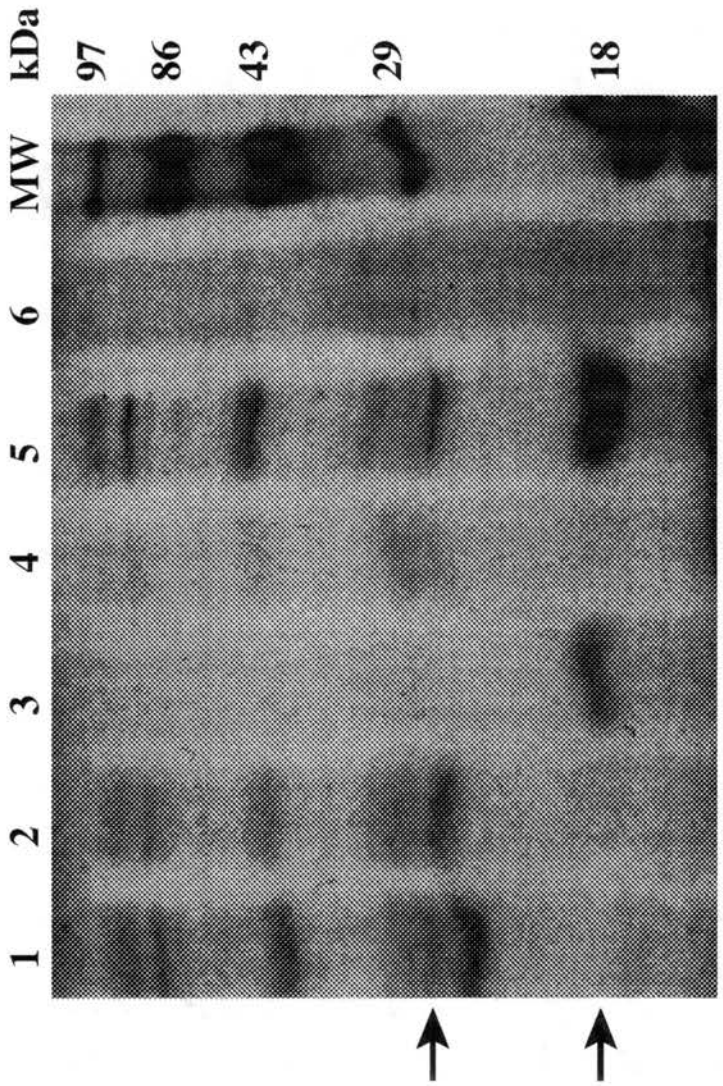
hybridized with a *tfdR* gene probe is shown in Figure 11. The probe consisted of a 645 bp region (*Bam*HI-*Xba*I) on *Bam*HI-E. This region carried the first 542 bp of *tfdR* gene and 103 bp of its upstream sequences. As shown in the figure, fragments *Bam*HI-B, *Bam*HI-E, *Eco*RI-E and *Hin*dIII-B hybridized with the probe. It is already known that fragments *Bam*HI-B and *Bam*HI-E carry portions of *tfdS* and *tfdR* respectively (refer Fig. 7). Fragments *Eco*RI-E and *Hin*dIII-B carry entire *tfdR* and *tfdS* genes. The Southern hybridization was performed at high stringency and therefore eliminates non-specific hybridization. It confirms that no other regions on pJP4 contain copies of these two genes.

Expression and Selective Labeling of TfdR

E.coli k38(pGP1-2) clones containing different DNA constructs of *tfdR* gene were expressed by heat induction using T7 Promoter/T7 RNA polymerase system. All plasmid encoded proteins were labeled with amino acid ^{35}S-methionine. Autoradiographs of labeled proteins electrophoresed on SDS polyacrylamide gels are shown in Figures 12 and 13.

As shown in Figure 12, expression clones containing either plasmid pMH4 or pMH2 expressed a polypeptide of 18,400 Daltons (lanes 5 and 3 respectively). Plasmid pMH4 contained a 1.2 kb (*Bam*HI-*Xho*I) insert DNA carrying only the first 542 bp of *tfdR*.

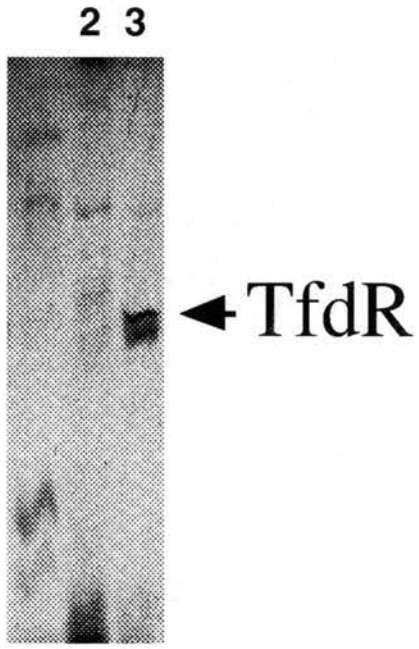
Figure 12. Expression of the *tfdR* and *tfdDII* Genes Sub-Cloned in *E. coli* K38(pGP1-2). Cell extracts of *E. coli* K38(pGP1-2) clones containing various expression plasmids were separated on SDS-polyacrylamide gel. Clones carried respectively pMH3 (lane 1); pT7-5 (lane 2); 3, pMH2 (lane 3); pMH1 (lane 4): pMH4 (lane 5) and pT7-6 (lane 6). Lane 7 contained molecular weight markers (in kilodaltons). The arrows indicate 31 kDa truncated *tfdDII* gene product in lane 1 and truncated 18.4 kDa *tfdR* gene products in lanes 3 and 5.


kDa
97
86
43
29
18
MW
6
5
4
3
2
1

Whereas plasmid pMH2 contained only a 645 bp (*Bam*HI-*Xba*I) DNA insert that carried the first 542 bp of *tfdR* gene. The size of the polypeptide expressed from these two plasmids were equal to the size predicted from the DNA sequence of the truncated ORF. The expression clone containing plasmid pMH3 expressed a protein of approximately 31,000 Daltons (lane 1). This plasmid contained the same DNA insert as in plasmid pMH4 but in the opposite orientation with respect to the T7 promoter on the vector. The 31,000 Daltons protein was expressed from the incomplete *tfdDII* gene present within the DNA insert (refer Fig. 4). The size of this truncated protein did not compare with the predicted size of incomplete *tfdDII* ORF, as the polypeptide also included translated portions of vector sequences. Unfortunately, the vector does not have stop codons behind the multiple cloning site. Therefore expression of any gene product must be terminated from its own stop codon present within the insert DNA. As seen in the Fig. 12, no proteins were expressed from clones carrying only plasmid vectors without insert DNA (lanes 2 and 6).

The DNA sequence revealed the complete *tfdR* gene located across the *Bam*HI-E/F junctions of pJP4. New expression clones were constructed with entire *tfdR* gene and expressed for the complete protein. As shown in Figure 13, the *E. coli* K38(pGP1-2) clone containing plasmid pMH6 expressed a protein with an approximate size of 32,000 Daltons (lane 3). The plasmid pMH6 contained a DNA insert of 1.3 kb (*Xba*I-*Pst*I) consisting the entire *tfdR* gene. The size of this expressed protein was in agreement with the size predicted from the DNA sequence. The clone containing the same insert in the

Figure 13. Expression of Complete TfdR Protein. Cell extracts of expression clones carrying complete *tfdR* gene were separated on SDS-polyacrylamide gel. Lanes: 1, molecular weight markers in kilodaltons; 2, cell extracts of clone with pMH6; 3, cell extracts of clone with pMH5. The arrow indicates the 32 kDa *tfdR* gene product in lane 3.

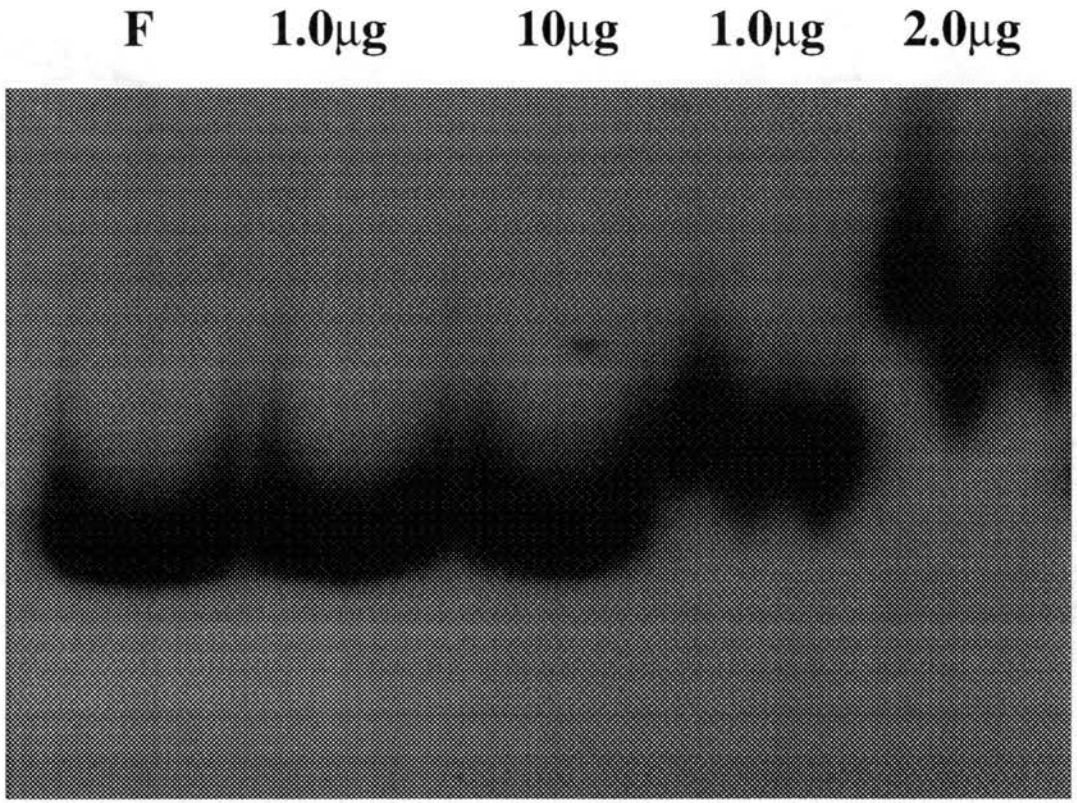

2 3
TfdR

reverse orientation (plasmid pMH5) did not express any protein (lane 2).

Protein-DNA Interactions

To determine whether regulation of *tfdA* gene expression was due to direct interaction of the *tfdR* protein with the *tfdA* upstream DNA region, a set of DNA-protein binding reactions were performed. TfdR protein was overexpressed by heat induction in the *E.coli* k38(pGP1-2) clone harboring plasmid pMH6. Cell extracts containing overexpressed *tfdR* were tested for their ability to bind with regions of DNA sequences upstream of the *tfdA* gene. The DNA was end labeled. An autoradiograph showing electrophoretic mobility shift of DNA sequences after interacting with cell extracts containing the TfdR protein is shown in Figure 14. As shown in the figure, the gel lane containing free labeled DNA did not show any shift in mobility (lane 1, F). Whereas, lanes containing labeled DNA that was interacted with increasing amounts of the cell extracts showed retardation in the mobility of DNA (lanes 4 and 5 with 1.0 to 2.0μg proteins respectively). No mobility of DNA was observed in the lanes were the labeled DNA was interacted with increasing amounts of crude extract from cells containing only the vector pT7-5 (lanes 2 and 3). DNA fragment from other regions of pJP4 when used as negative control did not show any retardation (data not shown). Mobility shift assay concluded that TfdR proteins bind to certain regions on the upstream sequences of *tfdA* gene.

Figure 14. Mobility Shift Assay of Promoter DNA With Cell Extracts Containing TfdR. Lanes: 1, Free DNA; 2, DNA plus 1.0 μg of crude extract from cells with vector alone; 3, DNA plus 10.0 μg of crude extract from cells with vector alone; 4, DNA plus 1.0 μg of cell extract containing TfdR; 5, DNA plus 2.0 μg of cell extract containing TfdR.

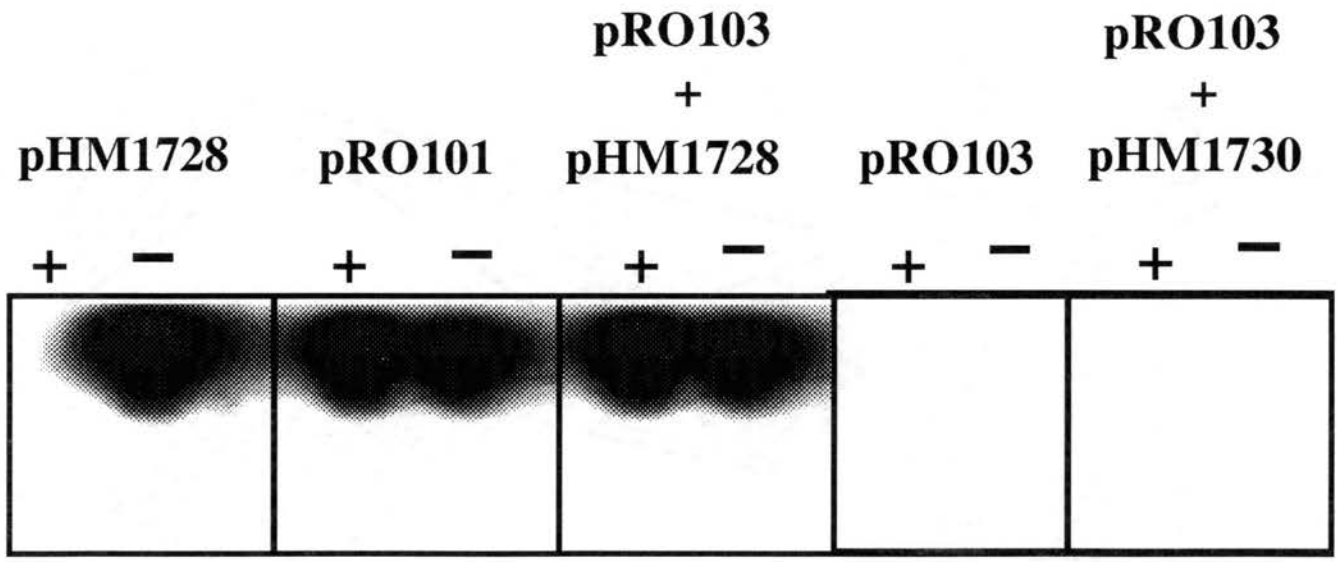

F
1.0μg
10μg
1.0μg
2.0μg

Northern Hybridization

A northern blot of RNA hybridized to a riboprobe is shown in Figure 15 on pages 75 and 76. The RNAs were isolated from *Pseudomonas aeruginosa* PAO1c strains containing derivatives of pJP4 and different plasmid constructs of *tfdR* gene. The riboprobe contained sequences identical to the coding strand of *tfdR* and *tfdS* genes, which hybridized to the messenger RNA transcribed from it.

Northern hybridization indicated that plasmid pRO101 produced messenger RNA from both or either one of*tfdR* and *tfdS* genes when induced or uninduced for the TFD degradation pathway. It confirmed that pRO103 did not contain functional copies of the two genes, as the clone containing pRO103 did not produce any transcript that would hybridize with the probe (lane 4). Northern hybridization with transcripts made from clones carrying two plasmids (pRO103/pHM1728) confirmed that the truncated *tfdR* gene cloned in pHM1728 produced transcripts under uninduced or induced conditions. The plasmid pHM1728 by itself did not make any transcripts when induced by TFD (lane 1), as the *Pseudomonas* strain does not have any *tfd* metabolic genes on the chromosome. No transcript was made from pHM1730 under TFD induced or uninduced conditions (lane 5). This plasmid has a deletion of the truncated *tfdR* gene (refer Fig. 5b).

Figure 15. Northern Hybridization of Transcripts With *tfdR* Riboprobe. RNAs were extracted from *Pseudomonas aeruginosa* PAO1c strains carrying pRO101, pRO103, pHM1728 or pHM1730 and hybridized to a *tfdR* riboprobe after electrophoresis. The lanes with RNA are shown by the name of the plasmids the strains carried. The hybridized transcript had an approximate size of 1200 nucleotides (RNA markers are not shown).


pHM1728
pRO101
pRO103
+
pHM1728
pRO103
pRO103
+
pHM1730
+ −
+ −
+ −
+ −
+ −

CHAPTER V

DISCUSSION

Preface

In this research, the complete 888 bp DNA sequence of the regulatory gene *tfdR* and the terminal 346 bp of the regulatory gene *tfdS*, in the TFD degradation pathway of plasmid pJP4 of *Alcaligenes eutrophus* strain JMP134, were determined. These genes were found to be 100% identical in the region sequenced and to encode a LysR family protein. This is the first instance where two copies of a regulatory gene have been identified in a xenobiotic degradation pathway carried on a plasmid.

Nucleotide sequences revealed that 542 bp of *tfdR* are located on *Bam*HI-E and the remaining 346 bp on *Bam*HI-F. When *Bam*HI-E was first identified as encoding an element which negatively regulates the expression of *tfdA* gene, it was believed to contain a complete regulatory gene (39). It is evident that only the first 542 bp of *tfdR* are essential for negative regulation. These are also the exact 542 bp of *tfdS*, present on the small portion of *Bam*HI-B, on the upstream region of *tfdA* gene.

Sequence analysis of the *Bam*HI-E fragment also revealed the presence of a putative gene *tfdDII*, encoding an isofunctional chloromuconate cycloisomerase enzyme. This identification is supported by the protein expression data. It provids evidence that *tfdDII* gene is functional and directs the synthesis of a protein. The exact enzyme activity of the expressed protein must be determined in the future.

Southern hybridization conducted in this research concluded that pJP4 carries no additional copies of *tfdR* or *tfdS* genes and indicated that no DNA homology exists between *tfdR* and the putative regulatory gene *tfdX*. However, it was determined through data base search that TfdX shared significant similarity to TfdR in the amino acid sequence deduced from the incomplete *tfdX* ORF.

Northern hybridization indicated complete deletion of both *tfdR* and *tfdS* genes in pRO103, as no transcripts were found to bind with the *tfdR* riboprobe. As no accurate mapping of the deletion in pRO103 is available it could not be elucidated if pRO103 contains small portions of either *tfdR* or *tfdS* open reading frame. If present, these portions will be short inverted repeats. The configuration may hinder synthesis of transcripts from both of the genes. Or even if they were synthesised, the short transcripts must have been very unstable and hence no transcripts were found to hybridize with the riboprobe.

DNA Sequences, Proteins and Regulation

Results from this research provided complete structure of regulatory genes *tfdR* and *tfdS*. They helped to dispose the theory that the two genes are different and control varied *tfd* genes. The results also prompted critical analysis of earlier publications. On analysis, it was found that the conclusions drawn on the regulation of *tfd* genes had been based on inadequately generated data.

The small deleted segment on *Bam*HI-B of pRO103 must have repressor activity as *Bam*HI-E. The segment must have caused the same effect like *Bam*HI-E fragment, when placed *in trans* with pRO103 (39). Because the deleted segment contains the same sequence as in *Bam*HI-*Xba*I of *Bam*HI-E fragment. Had the small segment been determined for repressor activity *in trans* with pRO103, *tfdR* would have identified upstream of *tfdA* on *Bam*HI-B instead on *Bam*HI-E fragment.

The cloned DNA fragments that presumably carried *tfdCDEF* and *tfdB* genes, do not represent precise operon regions of pJP4 (41,42). The fragments contained stretches of upstream or downstream sequences as well. It was observed both in this research and by others (90) that the upstream sequence of the *tfdCDEF* operon contains the putative regulatory gene *tfdX* which shares similarity to TfdR in the deduced amino acid sequence. Therefore, results obtained in complementation studies involving genes *tfdR* and *tfdS* placed *in trans* to genes *tfdCDEF* and *tfdB* respectively, must be determined again. It is possible the results

respectively, must be determined again. It is possible the results may have provided artifactual information due to cross reactivity of regulatory proteins TfdR or TfdS with TfdX.

It was also observed that the gene copy number of the regulatory and the metabolic gene has not been equally maintained in complementation studies (41, 42). The regulatory gene has been on multiple copy vectors placed *in trans* with low copy plasmids like pRO101 or pRO103. The copy numbers for *tfdR in trans* to pRO103, carrying *tfdCDEF* and *tfdX*, has been in an approximate ratio of 150:2. The difference in the protein quantities made by *tfdR* and *tfdX* genes could have been a major factor in the expression of their activities. The higher quantity TfdR protein could have cross reacted with the *tfdCDEF* promoter and totally eliminated interactions of TfdX protein with the promoter.

This speculation is supported by interesting observations made by others with TcbR protein. TcbR, a TfdR homologous protein, activates the expression of *tcbCDEF* operon in Pseudomonas sp. P51. The operon which encodes the degradation of chlorocatechol has high homology to the *tfdCDEF* operon of pJP4. It had been demonstrated that TcbR cross reacted with the promoter of *tfdCDEF* (90). If a homologous regulatory protein from a different bacterium has the ability to interact with *tfdCDEF* promoter, it is highly possible that TfdR or TfdS present in the same bacterium and homologous to TfdX could also react to *tfdCDEF* promoter. Therefore the actual regulatory protein TfdX produced in lower quantities might not have contributed to the results at all. Importantly, the location of *tfdX*, upstream and divergent of *tfdCDEF*, is identical to the configurations

found in the catechol and chlorocatechol genes of other bacterial species (2, 28, 31, 33, 55, 89). They all have divergently oriented regulatory genes encoding a LysR family protein. None of the known chlorocatechol genes are controlled by a distantly located regulatory gene like *tfdR*. Hence, it is very possible the *tfdX* gene regulates *tfdCDEF* operon.

Variable functions that were observed for genes *tfdR* and *tfdS* may have been due to different sizes of proteins made from the respective clones (41, 42). Any *tfdR* clone contained a 1.2 kb *Bam*HI-*Xho*I region present on *Bam*HI-E. This region carried only the first 542 bp of *tfdR* gene which was not known then (41). However, *tfdS* clone contained the entire gene located on *Bam*HI-B and F (42). The two clones produced 180 amino acid truncated TfdR and 295 amino acid complete TfdS proteins, respectively. The truncated TfdR contained a carboxy terminus, which was same as the central region of the complete TfdS. Therefore, these partially identical proteins may have behaved variably.

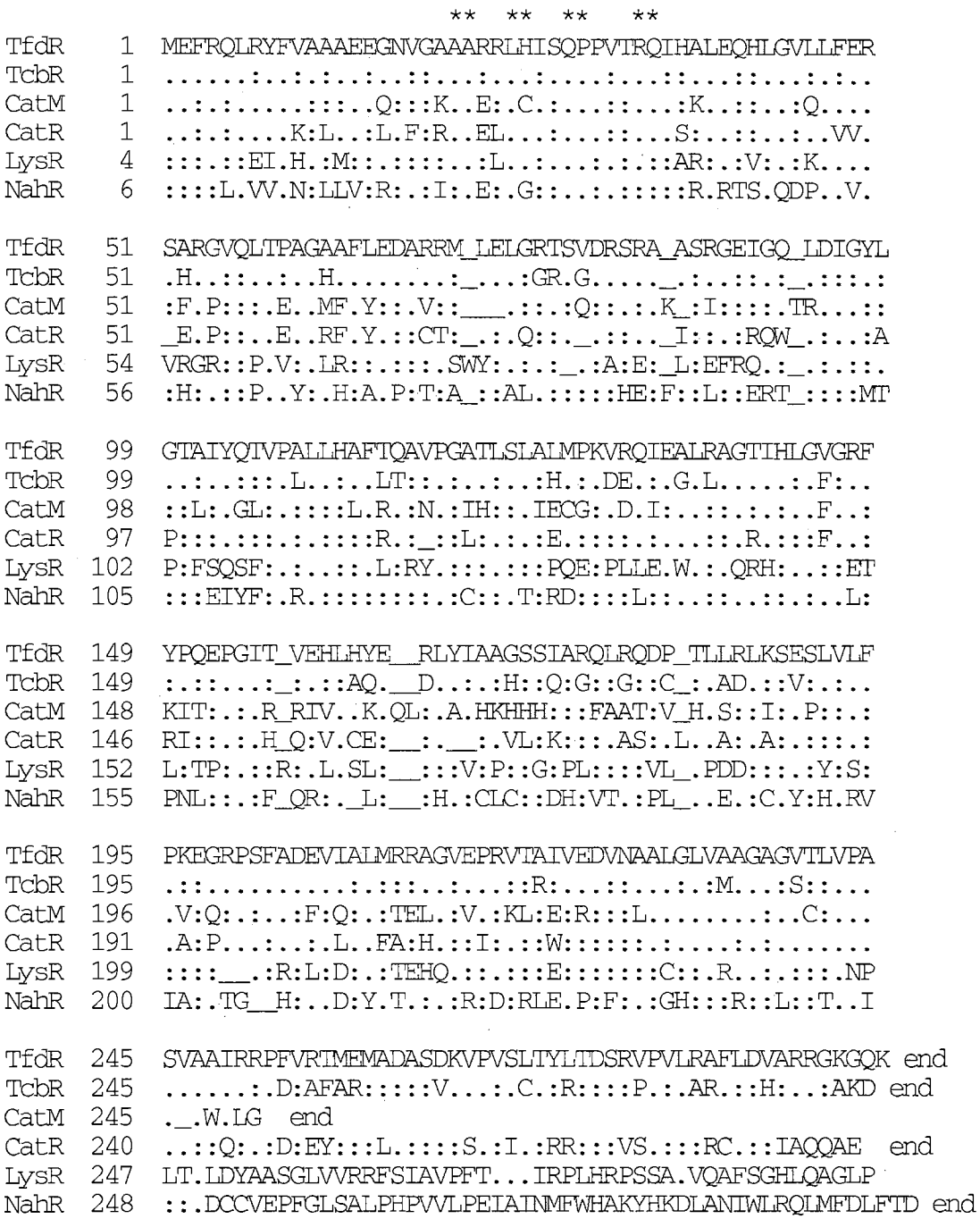
LysR family proteins have been proposed to contain a DNA binding motif in their first 70 amino acids (40). Hence the two proteins which shared 180 identical residues contained the same DNA binding motif. LysR family proteins have also been observed to contain non-homologous central region and carboxy terminus that are proposed to confer specific binding sites for effector metabolites (40, 73, 82). The truncated TfdR unlike the complete TfdS, probably contained no interaction sites for metabolites mediating the regulation of *tfdB* gene. It was probably why *tfdS* and not *tfdR* gene had been proposed to regulate the expression of *tfdB*.

The truncated *tfdR* gene on *Bam*HI-E fragment had been reported to derepress the expression of *tfdA* and *tfdCDEF* genes in the presence of TFD or 4-chlorocatechol (4-CC) (41). Truncated TfdR must have interacted with these two metabolites or their downstream products. Therefore its carboxy terminus, which is the same as the central region of the complete TfdS, probably contains interaction sites for these metabolites. Alternately, the complete protein may not interact with these two metabolites due to conformational hinderance. This could only be proven by performing derepression study with the entire gene.

Published data have shown that *Bam*HI-E when placed *in trans* negatively regulates *tfdA* gene on pRO103 (39). In this case, truncated TfdR must have repressed the gene. Observations made in this research suggested that *tfdR* may also act as an activator. The *tfdR* gene specific transcripts were found to be produced under derepressed conditions and also all known LysR family members are transcriptional activators with the exception of CatM (40). When the amino acid sequences of activator proteins such as TcbR, CatR, LysR, NahR were compared with that of TfdR, it strongly suggested that TfdR may also have activator functions.

As shown in Figure 16, homology exists among the amino terminus and central domain of TfdR, TcbR, and CatR proteins. If the homology can imply identical functions, TfdR is also an activator protein. CatR and CatM have been shown to require *cis,cis*-muconate for their function (59, 56). Then, TfdR may also require muconate for activation. Chloromuconate a downstream by-product of the TFD pathway is a potential canditate. Incidentally, the same has been

Figure 16. Alignment of Deduced Amino Acid Sequences of TfdR and Other Regulatory Proteins. The pair-wise alignment of TfdR with each of the proteins obtained from the FASTA output file were converted into multiple sequence format as shown in the figure. Symbols: "." identical amino acid residues; ":" conserved changes; "-" gaps needed for optimal alignment; " * " strongly conserved residues involved in the helix-turn-helix N-terminal parts of the proteins. Numbers indicate the positions of the amino acid residues in the total sequence of the deduced proteins.


** ** ** **
TfdR 1 MEFRQLRYFVAAAEEGNVGAAARRLHISQPPVTRQIHALEQHLGVLLFER
TcbR 1
CatM 1
CatR 1
LysR 4
NahR 6
TfdR 51 SARGVQLTPAGAAFLEDARRM_LELGRTSVDRSRA_ASRGEIGQ_LDIGYL
TcbR 51
CatM 51
CatR 51
LysR 54
NahR 56
TfdR 99 GTAIYQIVPALLHAFTQAVPGATLSLALMPKVRQIEALRAGTIHLGVGRF
TcbR 99
CatM 98
CatR 97
LysR 102
NahR 105
TfdR 149 YPQEPGIT_VEHLHYE__RLYIAAGSSIARQLRQDP_TLLRLKSESLVLF
TcbR 149
CatM 148
CatR 146
LysR 152
NahR 155
TfdR 195 PKEGRPSFADEVIALMRRAGVEPRVTAIVEDVNAALGLVAAGAGVTLVPA
TcbR 195
CatM 196
CatR 191
LysR 199
NahR 200
TfdR 245 SVAAIRRPFVRIMEMADASDKVPVSLTYLIDSRVPVLRAFLDVARRGKGQK end
TcbR 245 end
CatM 245 end
CatR 240 end
LysR 247
NahR 248 end

proposed as the effector for TcbR (90).

It was observed that the published data on the activities of *tfd* metabolic genes had not taken into account the activities of recently identified isofunctional genes (63, 34). Data on the monooxygenase activity of *tfdA* gene has not excluded the activity of *tfdAII* gene. Similarly the activities of genes *tfdDII* or the *tfdC1* have not been separately determined. The isofunctional genes probably have significant impact on the total enzyme activity. They may also have a role in the global control of the TFD pathway by providing effector by-products.

Analyses of the published data and the findings from this research suggest that the regulatory mechanisms of all *tfd* genes must be redetermined. Alternately, it could be speculated that the *tfd* metabolic genes are regulated by regulatory genes located closest to them. For instance, genes *tfdR*, *tfdS* and *tfdX* may regulate genes *tfdDII*, *tfdA* and *tfdCDEF*, respectively. And gene *tfdB* may be regulated by some downstream element as its promoter shows no homology to other *tfd* promoters (see below).

Promoters and Regulatory Sites

Multiple DNA sequence alignment of various promoters is shown in Figure 17. The alignments contain intervening DNA sequences between translational start sites of the regulatory gene and the metabolic gene or genes found in pairs. Except for *tfd* genes,

Figure 17. Multiple Sequence Alignment of Promoters. The intervening DNA sequences between the regulatory and metabolic gene pair were aligned with that of the sequences between *tfdR* and *tfdDII* genes. The sequences were taken from GenBank data base. Gaps introduced for optimal alignment have been denoted by ".". The underlined bases refer to the transcription start sites of regulatory genes. The conserved nucleotides in Domains 1 and 2 are shown in upper case.

```
                                                                  Domain 1
tfdR-tfdDII                caTtctTCAcTCcTgGTgGCTGGctccGGctgCCGGAgAGCCATACCGATCCCGTATCGcTcGCGCTGAtggAA..
tfdS-tfdA                  caTtctTCAcTCcTgGTgGCTGGctccGGctgCCGGAgAGCCATACCGATCCCGTATCGcTcGCGCTGAtggAA..
tfdX-tfdCDEF            gtgagcaaGcaCctgCggTcTcGC.........aactagGGAaAGaCATACCGAaaCCGTATgGtctagGCgGAgAaAA..
clcR-clcA                catttaaGacCtcTggTTtTccCTaGgggtGGacaCCGctaAGCCATACCGATCCCGTATtGcaaagGCTaAAAaAA..
tcbR-tcbCDEF                       CAtggcgcGctttccGagttGGtgatgtGcctatatTACgcAaaCCGTAaCGaTgGCtgactAAttt..
catR-catBC                                                        CATcagacctCCagggtatggtGgGagAttcAttc
tfdB                     gaggTgGcTCggattTTcTttgcGcgaagtGgctCatttttaCttTgCgcgcaaCaaAggctcttCagccttActAtt

                           Domain 2
tfdR-tfdDII         GGTATTaGACCaTATGGcCCGGCAtTTCTAGACTACCgCCaTGctcAcagAaAAagcCattgcCGaCAgCCccaACggcggC.G.
tfdS-tfdA           GGTATTaGACCaTATGGcCCGGCAtTTCTAGACTACCgCCaTGatAAaactcggctgCtctCtCGTCtgCtGgaACATcttCAG.
tfdX-tfdCDEF        GGTATTGGACgGTATGGgattGCcg.TCTctAtTATgaCtgTGcgggGCaATtgTCgCCcaCGtGTtcAtgttttCATgAcg...
clcR-clcA           GGTATTGGACCGcATGacaCGcgAa.TCTTAGCatTCatgTTTgAAgcAcCaAcTCATCggTGttTCAACCatcAgATcttgaAa
tcbR-tcbCDEF        GGTATTGGACgGcATaGgCCGGCgc.TgcctAgcATtcaCTcGtccCGCgtcAATCcggacCGgGcaAACCGacAaAggAgaccG
catR-catBC          GaTATTGGACgGCtat..CaGGg...TCTcgcgcAatcCtTgaacAAGCaAgtAaaTgcacCaCtaagtgtGctgatTgAaCgta
tfdB                c..AaTGcACCactcGGggCccgcgcgCTgGcagcCCGaaAcaACgGgACCtggCAaCAaTGACccCTGttTCttCTcAcaaAG.

tfdR-tfdDII         acGcGatcgcaAggCggCgCaGAtCgAagCgaTcGAAAcGGtG                                  199
tfdS-tfdA           GcGcGctgagccgtCTttttGaAaCAgtctctTaGAAAaGGaGcaaaaaaagtg                       209
tfdX-tfdCDEF        GaGGcaaAgTg                                                                  162
clcR-clcA           GgaGacgAgTcAtg                                                               175
tcbR-tcbCDEF        GgatG                                                                        156
catR-catBC          tcGaGgcAaTtAttgTg                                                            136
tfdB                agaGagacaTaAcaaTtgaCGGAggAAgaCatg
```

all others have been shown controlled by the divergently located regulatory gene.

Of the aligned promoters, the *tcbR-tcbCDEF* and *catR-catBC* pairs have only been characterized (90, 1). The upstream sequences of *tfdB* gene have also been included as it has been proposed to be regulated by *tfdS*. Its promoter has also not been characterized, and is not known to contain any divergent regulatory gene. As shown in the Figure 17, excluding *tfdB* gene, very high conservation of nucleotides was found in two different domains of the promoter regions. Based on the characterized *tcbR-tcbCDEF* promoter, "domain 2" represent the -35 region of both genes in a pair (the -35 regions of *tcbR-tcbCDEF* have been shown to overlap and share the same base pairs on either strand). In *catR-catBC* the -10 region of one gene overlapped the -35 region of the divergent gene and vice versa. This is because it has a very short (132 bp) intervening sequence (1). All other pairs were found to contain approximately 170-200 intervening base pairs.

The fact that *tfdB* gene promoter shares no homology with others suggests a different type of regulation. It is probably not regulated by *tfdS* or *tfdX* gene.

Similarities and conserved nucleotides among promoters of *tfdA*, *tfdCDEF* and *tfdDII* genes indicate close identity in the mechanisms of their regulation. Interestingly, a possible σ^{70}-binding site TTGGAC-18 bp-CTATCCT present in a constitutive mutant of *catB* promoter was found to be nearly identical to a similar TTGGAC-18 bp-CTATTAT site in *tfdCDEF* promoter (3, 17). This is worthy of note because it is known that *tfdR* repressed constitutive expression of

tfdCDEF. A similar TTAGAC-18 bp-TAGACTA site found both on *tfdA* and *tfdDII* promoters is nearly identical to TTGGAC-18 bp-CTAGCAT and TTGGAC-18 bp-TAGCATT σ^{70}-binding sites on *tcbCDEF* and *clcA* promoters respectively. The expression of *tcbCDEF* operon has been reported to be activated by *tcbR* (90). If the presence of similar nucleotides in the σ^{70}-binding sites of both *tfdA* and *tfdDII* suggest activation, then they are probably activated by *tfdS* and *tfdR* genes, respectively.

Recently, two different *CatR* binding sites have been identified in *catR-catBC* promoter (59). They are a repression binding site (RBS) and an activation binding site (ABS). CatR protects RBS in the presence or absence of *cis, cis*-muconate and ABS in the presence of *cis, cis*-muconate (59). Comparison of these two sites with the aligned promoter sequences reveals "domain 1" having high homology with RBS and "domain 2" with ABS. In both RBS and ABS few "G" residues have been identified by methylation interference that are critical for CatR binding. Interestingly, in "domain 2" the "G" residue is conserved at the same position as in the ABS. In the RBS, of the five critical "G" residues, 3 of them are located on one strand of the DNA and separated by 4 bp from the other two (GG) located on the opposite strand. In "domain 1" which showed homology to RBS, "TAT" ("TAA" in *tcbR-tcbCDEF*) were present at the same position instead of the 3 Gs and was separated by 4 bp from a "AT" ("TT" in *tcbR-tcbCDEF* and *tfdX-tfdCDEF*) on the other strand. The variation noticed in the type of nucleotide base pairs in "domain 1" probably reflects the difference in net charge of the regulatory proteins interacting with it. CatR has an estimated pI of only 6.1 (68),

whereas TfdR and TcbR have a pI of 10 (90). Therefore, the switch in the type of base pairs may represent the appropriate steric and spatial interaction between the protein and the promoter. It is perhaps premature to speculate. It can only be confirmed with high resolution X-ray crystallography performed with highly purified proteins and promoter DNA.

The promoters of *tcbR* and *nahR* genes of *Pseudomonas putida* have been shown to share some conserved nucleotide residues (89). These residues on *nahR* promoter are protected by NahR protein on a foot-printing assay (72). NahR is a LysR member and activator of the *nahS* gene encoded in *Pseudomonas putida* (73). These residues on *tcbR* promoter are located at the corresponding positions as on *nahR* (90). Therefore both TcbR and NahR may have similar DNA binding domains on promoter. Interestingly, none of these protected nucleotide residues except a "G" is conserved in other promoters. The observation suggests that TfdR and TcbR may have different DNA binding domains, even though they shared similarity in amino acid and promoter DNA sequences. This also reflects possible differences in the mechanisms of regulation mediated by the two genes.

Evolutionary Conservation

The homology of related regulatory proteins has revealed evolutionary linkages and better insight into evolutionary

mechanisms. As shown in Figure 16, proteins TfdR, TfdX, TcbR, ClcR, CatR and CatM shared less identity with LysR but significant similarity among themselves. The BLASTP search also provided closer identity of amino acid sequences of TfdR, TcbR, CatR, CatM and TfdX. The maximal segment pair scores ranged from 748 to 144 and the smallest poisson numbers were 1.8e-130 to 1.8e-16 for these proteins. The LysR protein had a MSP of 88 with a poisson number of 0.00015. The similarity of functionally related proteins opens the possibility for a sub family within the main LysR family. The respective regulatory genes encoding these proteins also showed significant conservation of DNA sequences. Therefore, all these regulatory genes must have evolved from a parental gene distantly related to *lysR* gene. The parental gene, due to environmental pressure, probably evolved into these closely related genes. It is interesting to note that NahR encoded by *Pseudomonas putida* shares the least similarity (21%) with TfdR encoded by *Alcaligenes eutrophus*. *Alcaligenes eutrophus* has been classified as a pseudomonad until very recently (80). Whereas, the protein ClcR encoded by *E. coli* shares significant similarity with TfdR in its deduced amino acid sequence from the reported incomplete ORF and the also promoter region of *clcR* reveals significant conservation of nucleotides similar to *tfdR* promoter. Therefore it is likely that *nahR* may have evolved much earlier than *tfdR* or from a different parental gene. And *tfdR* and *clcR* may have closer evolutionary relationship than *tfdR* and *nahR*.

During the evolution of degradative pathways, there appears to have been a central degradation system like the chlorocatechol

degradation genes around which peripheral genes have evolved to degrade substituted aromatic compounds. In the TFD pathway of pJP4, the monooxygenase gene (*tfdA*) and the phenol hydroxylase gene (*tfdB*) may have evolved only after the evolution of chlorocatechol degrading *tfdCDEF* genes (88). Because these two peripheral genes probably were recruited after the spread of TFD in the environment. This observation is supported by the identification of homologus chlorocatechol genes in *Pseudomonas* and *Acinetobacter* strains that carry no monoxygenease or hydroxylase genes (1, 89, 55).

Another interesting observation is that the clustered chlorocatechol or catechol genes have also been identified to carry divergent regulatory genes as in the *tfdCDEF-tfdX* pair (1, 55, 90). The regulatory gene is usually located on the upstream region and is transcribed divergently. It appears that related pathways have acquired identical control mechanisms during evolution. Alternately, the control mechanism has been as well conserved as the central degradation system. In pJP4, the regulatory gene *tfdX* paired to the central system probably gave rise to the identical genes *tfdR* and *tfdS*.

The presence of an ORF in *Bam*HI-E fragment which may encode chloromuconate cycloisomerase suggests the possibility of an isofunctional *tfdDII* gene in plasmid pJP4. Earlier report has indicated the presence of isofunctional *tfdC1* on *Bam*HI-E fragment that is homologous to *tfdC* (34). Gene *tfdC1* has been located near the *Bam*HI-E/C junction, closer to *Bam*HI-C (refer Fig. 18). As incomplete DNA sequence of *tfdDII* and none of *tfdC1* were determined, it could

not be predicted whether *tfdDII* and *tfdC1* are in an operon or cluster with *tfdDII* preceding *tfdC1* in gene order. If they are as predicted, it will not be the first case where the order of isofunctional genes has been found to be altered. The phenomenon has been reported in the *pca* genes of *Acinetobacter calcoaceticus* and *Pseudomonas putida* (23, 58) and in the *cat* genes of *A. calcoaceticus* (76). It has been proposed that such altered gene arrangements occur due to evolutionary divergence(58).

The rearrangement in pJP4, will be more interesting if any other isofunctional *tfd* genes are also found on *Bam*HI-C, linked with the two on *Bam*HI-E. A little further down on *Bam*HI-C, the isofunctional *tfdAII* was identified, and whose sequence and orientation are not known (63). It is therefore likely that more *tfd* genes are present between *tfdDII* and *tfdAII* constituting a new operon with altered order. The isofunctional operon may be the primitive sequence from which other *tfd* genes have evolved.

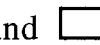
The reasons for the presence of altered isofunctional genes are not known. It is probably advantageous to the organism because it may provide global control of the degradation pathway. For instance, if a by-product of TFD pathway inhibited the activity of a gene in the regular operon, it would arrest further degradation. The accumulated by-product could turn on the isofunctional operon with the altered gene order. The first gene in the altered operon will then enzymatically degrade the inhibiting by-product and simultaneously switch on the main pathway again. Or it could be much more complex. Future research of *Bam*HI-E and adjacent *Bam*HI-C on pJP4 will provide answers about the presence of isofunctional genes and

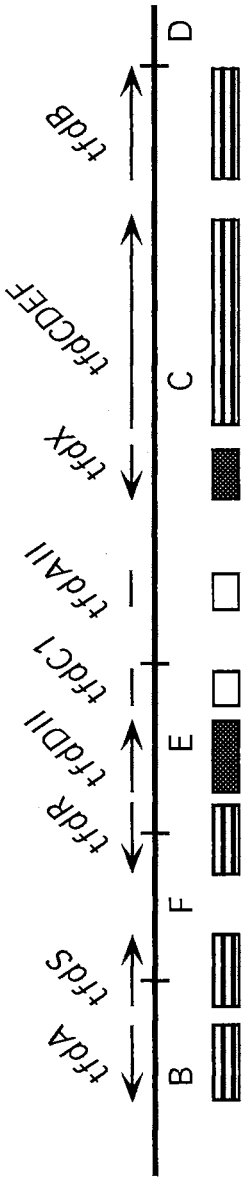
their functions in the fine-tuning of pathway regulation.

Conclusions

The regulation of TFD pathway of pJP4 remains unresolved. The data collected from this research have provided new insight into the structure of *tfd* regulatory genes and their mechanisms. A comprehensive illutsration of all known *tfd* genes is given on Figure 18. Future research on each of these regulatory genes performed separately and jointly with all or individual metabolic genes will help to learn better the control of TFD pathway.

The knowledge on the regulatory gene *tfdR* has indicated other possible functions of the gene in the expression of various catabolic genes. If the TfdR protein interacts with different metabolites as speculated in this research, the *tfdR* gene would be a prospective candidate in the construction of novel degradation pathways. A novel pathway with wider substrate utilization ability when constructed with *tfdR* gene, the dynamic nature of TfdR protein will offer control significantly in the efficient degradation of aromatic hydrocarbons.

Figure 18. Comprehensive Illustration of *tfd* Genes on pJP4. The *Bam*HI-B, F, E, C and D fragments of pJP4 and locations of known *tfd* genes are shown. The arrows indicate the direction of transcription. The short lines indicate unknown direction of genes. Symbols:  completely sequenced;  partially sequenced; and  not sequenced genes.


tfdA
tfdS
tfdR
tfdDII
tfdC1
tfdAII
tfdX
tfdCDEF
tfdB
B
F
E
C
D

BIBLIOGRAPHY

1. Aldrich, T. L., and A. M. Chakrabarty. 1988. Transcriptional regulation, nucleotide sequence, and localization of the promoter of the *catBC* operon in *Pseudomonas putida*. J. Bacteriol. 170:1297-1304.

2. Aldrich, T. L., B. Frantz, J. F. Gill, J. J. Kilbane, and A. M. Chakrabarty. 1987. Cloning and complete nucleotide sequence determination of the *catB* gene encoding *cis, cis*-muconate lactonizing enzyme. Gene 52:185-195.

3. Aldrich, T.L., R. K. Rothmel, and A. M. Chakrabarty. 1989. Identification of nucleotides critical for activity of the *Pseudomonas putida catBC* promoter. Mol. Gen. Genet. 218:266-271.

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

5. Beck, C. F., and R. A. J. Warren. 1988. Divergent promoters, a common form of gene organization. Microbiol. Rev. 52:318-326.

6. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7:1513-1523.

7. Bollag, J.-M., G. G. Briggs, J. E. Dawson, and M. Alexander. 1968. 2,4-D metabolism. Enzymatic degradation of chlorocatechols. J. Agric. Food Chem. 16:829-833.

8. Bradford, M. M. 1976. A rapid and sensitive method for the determination of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

9. Brennan, R. G., and B. W. Matthews. 1989. The helix-turn-helix DNA binding motif. J. Biol. Chem. 264:1903-1906.

10. Chattterjee, D. K., and A. M. Chakrabarty. 1981. Plasmids in the biodegradation of PCBs and chlororbenzeoates. FEMS Symp. 12:213-219.

11. Chattterjee, D. K., and A. M. Chakrabarty. 1982. Genetic rearrangements in plasmids specifying total degradation of chlorinated benzoic acids. Mol. Gen. Genet. 188:279-285.

12. Chatterjee, D. K., S. T. Kellogs, S. Hamada, and A. M. Chakrabarty. 1981. Plasmid specifying total degradation of 3-chlorobenzoate by a modified *ortho* pathway. J. Bacteriol. 146:639-646.

13. Chaudhry, G. R., and G. H. Huang. 1988. Isolation and characterization of a new plasmid from *Flavobacterium* sp. which carries the genes for degradation of 2,4-dichlorophenoxyacetate. J. Bacteriol. 170:3897-3902.

14. Chaudhry, G. R., and S. Chapalamadugu. 1991. Biodegradation of halogenated organic compounds. Microbiol. Reviews 55:59-79.

15. Cuskey, S. M., V. Pecoraro, and R. H. Olsen. 1987. Intial catabolism of aromatic biogenic amines by *Pseudomonas aeruginosa* PAO: pathway description, mapping of mutations, and cloning of essential genes. J. Bacteriol. 169:2398-2404.

16 Dagert, M., and S. D. Ehrlich. 1974. Prolonged incubation in calcium chloride improves competence of *Escherichia coli* cells. Gene 6:23-28.

17. Deretic, V., W. M. Konyescsni, C. D. Mohr, D. W. Martin, and N. S. Hibler. 1989. Common denominators of promoter control in *Pseudomonas* and other bacteria. Bio/Technology 7:1249-1254.

18. Diamond, L. 1992. Alternative northern Hybridization analysis. In, The red book bulletin, sup. 17. Current protocols in molecular biology. John Wiley & Sons, Inc. N. Y.

19. Don, R. H., M. Hellwig, W. Reineke, and H.-J. Knackmuss. 1974.

Isolation and characterization of a 3-chlorobenzoate degrading pseudomonad. Arch. Microbiol. 99:61-70.

20. Don, R. H., and J. M. Pemberton. 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. 145:681-686.

21. Don, R. H., and J. M. Pemberton. 1985. Genetic and physical map of the 2,4-dichlorophenoxyacetic acid degradative plasmid pJP4. J. Bacteriol. 161:466-468.

22. Don, R. H., A. J. Weightman, H.-J. Knackmuss, and K. N. Timmis. 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134 (pJP4). J. Bacteriol. 161:85-90.

23. Doten, R. C., K.-L. Ngai, D. J. Mitchell, and L. N. Ornston. 1987. Cloning and genetic organization of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. 143:834-840.

24. Duxbury, J. M., J. M. Tiedje, M. Alexander, and J. E. Dawson. 1970. 2,4-D metabolism: enzymatic conversion of chloromaleylacetic acid to succinic acid. J. Agric. Food Chem. 18:199-201.

25. Evans, W. C., B. S. W. Smith, H. N. Fernley, and J. I. Davies. 1971. Bacterial metabolism of 2,4-dichlorophenoxyacetate. Biochem. J 122:543-551.

26. Evans, W. C., B. S. W. Smith, P. Moss, and H. N. Fernley. 1971. Bacterial metabolism of 4-chlorophenoxyacetate. Biochem. J. 122:509-517.

27. Frantz, B., and A. M. Chakrabarty. 1986. Degradative plasmids in *Pseudomonas*, p. 295-323. In J. R. Sokatch and L. N. Ornston (ed), The Bacteria, vol. 10. The Biology of *Pseudomonas*. Academic Press, Inc., Orlando, Fla.

28. Frantz, B., and A. M. Chakrabarty. 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA

84:4460-4464.

29. Fried, M., and D. M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. 9:6505-6525.

30. Furukawa, K., and A. M. Chakrabarty. 1982. Involvement of plasmids in total degradation of chlorinated biphenyl. Appl. Environ. Microbiol. 44:619-626.

31. Ghosal, D., and I.-S. You. 1988. Nucleotide homology and organization of chlorocatechol oxidation genes of plasmids pJP4 and pAC27. Mol. Gen. Genet. 211:113-120.

32 Ghosal, D., and I.-S. You. 1988. Gene duplication in haloaromatic degradative plasmids pJP4 and pJP2. Can. J. Microbiol. 34:709-715.

33. Ghosal, D., and I.-S. You. 1989. Operon structure and nucleotide homology of the chlorocatechol oxidation genes of plasmids pJP4 and pAC27. Gene 83:225-232.

34. Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty. 1985. Genes specifying degradation of 3-chlorobenzoic acid in plasmids pAC27 and pJP4. Proc. Natl. Acad. Sci. USA 82:1638-1642.

35. Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty. 1985. Microbial degradation of halogenated compounds. Science 228:135-142.

36. Haigler, B. E., S. F. Nishino, and J. C. Spain. 1988. Degradation of 1,2-dichlorobenzene by a *Pseudomonas* sp. Appl. Environ. Microbiol. 54:294-301.

37. Hardman, D. J., P. C. Gowland, and J. H. Slater. 1986. Plasmids from soil bacteria enriched on halogenated alkanoic acids. Appl. Environ. Microbiol. 51:44-51.

38. Harker, A. R., and Y. Kim. 1990. Trichloroethylene degradation by two independent aromatic degrading pathways in *Alcaligenes eutrophus* JMP134. J. Bacteriol. 171:314-320.

39 Harker, A. R., R. H. Olsen, and R. J. Seidler. 1989. Phenoxyacetic acid degradation by the 2,4-dichlorophenoxyacetic acid (TFD) pathway of plasmid pJP4: mapping and characterization of the TFD regulatory gene *tfdR*. J. Bacteriol. 171:314-320.

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

41. Kaphammer, B., J. J. Kukor, and R. H. Olsen. 1990. Regulation of the *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradative plasmid pJP4. J. Bacteriol. 172:2280-2286.

42. Kaphammer, B., and R. H. Olsen. 1990. Cloning and characterization of *tfdS*, the repressor-activator gene of the *tfdB*, from the 2,4-dichlorophenoxyacetic acid catabolic plasmid pJP4. J. Bacteriol.172:5856-5862.

43. Karns, J. S., J. J. Kilbane, S. Duttagupta, and A. M. Chakrabarty. 1983. Metabolism of halophenols by 2,4,5-trichlorophenoxyacetic acid-degrading *Pseudomonas cepacia.* Appl. Environ. Microbiol. 46:1176-1181.

44. Kawasaki, H., H. Yahara, and K. Tonomura. 1981. Isolation and characterization of plasmid P4-01 mediating dehalogenation of haloacetate and mercury resistance in *Moraxella* sp. Agric. Biol. Chem. 45:1477-1482.

45. Kilpi, S., V. Backstrom, and M. Korhola. 1980. Degradation of 2-methyl-4-chlorophenoxyacetic acid (MCPA), 2,4-dichlorophenoxyacetic acid (2,4-D), benzoic acid and salicylic acid by *Pseudomonas* sp. HV3. FEMS Microbiol. Lett. 8:177-182.

46. Klages, U., A. Markus, and F. Lingens. 1981. Degradation of 4-chlorophenylacetic acid by a *Pseudomonas* species. J. Bacteriol. 146:64-68.

47. Knackmuss, H.-J. 1981. Degradation of halogenated and sulfonated hydrocarbons, p. 189-212. In T. Leisinger, A. M. Cook, R. Hutter, and J. Nuesch (ed.), Microbial degradation of xenobiotics and recalcitrant compounds. Academic Press, Inc. (London), Ltd., London.

48. Knackmuss, H.-J., and M. Hellwig. 1978. Utilization and cooxidation of chlorinated phenols by Pseudomonas sp. B13. Arch. Microbiol.117:1-7.

49. Kukor, J. J., R. H. Olsen, and J.-S. Siak. 1989. Recruitment of a chromosomally encoded maleylacetate reductase for degradation of 2,4-dichlorophenoxyacetic acid by plasmid pJP4. J. Bacteriol.171:3385-3390.

50. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature (London) 227:680-685.

51. Loos, A. M. 1975. Phenoxyalkanoic acids, p1-101. In P. C. Kearney and D. D. Kaufman (ed.) Herbicides: Chemistry, Degradation and Mode of action. Marcel-Dekker, Inc., NY.

52. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

53. Marks, T. S., A. R. W. Smith, and A. V. Quirk. 1984. Degradation of 4-chlorobenzoic acid by *Arthrobacter* sp. Appl. Environ. Microbiol. 48:1020-1025.

54. Masse, R., F. Messier, L. Peloquin, C. Ayotte, and M. Sylvestre. 1984. Microbial degradation of 4-chlorophenyl, a model compound of chlorinated biphenyls. Appl. Environ. Microbiol. 47:947-951.

55. Neidle, E. L., C. Hartnett, S. Bonitz, and L. N. Ornston. 1988. DNA sequence of *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene *catA*: evidence for evolutionary divergence of intradiol dioxygenases by acquisition of DNA sequence repetitions. J. Bacteriol. 170:4874-4880.

56. Neidle, E., C. Hartnett, and L. N. Ornston. 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. 171:5410-5421.

57. Neilson, A. H., A.-S. Allard, and M. Remberger. 1985.

Biodegradation and transformation of recalcitrant compounds, p.29-86. In O. Hutzinger, (ed.) The handbook of environmental chemistry, vol. 2, part C. Springer-Verlag KG, Berlin.

58. Ornston, L. N., J. E. Houghton, E. Neidle, and L. Gregg. 1990. Subtle selection and novel mutation during evolutionary divergence of the β-ketoadipate pathway, p. 207-225. In S. Silver, A. M. Chakrabarty, B. Iglewski, and S. Kaplan (ed), *Pseudomonas*: Biotransformations, pathogenesis, and evolving biotechnology. ASM Washington, D. C.

59. Parsek, M. R., D. L. Shinabarger, R. K. Rothmel, and A. M. Chakrabarty. 1992. Roles of CatR and *cis, cis*-muconate in activation of the *catBC* operon, which is involved in benzoate degradation in *Pseudomonas putida*. J. Bacteriol. 174:7798-7806.

60. Pemberton, J. M., B. Corney, and R. H. Don. 1979. Evolution and spread of pesticide degrading ability among soil microorganisms, p. 287-299. In K. N. Timmis and A. Puhler (ed.), Plasmids of medical, environmental and commercial importance. Elsvier/North-Holland Biomedical Press, Amsterdam.

61. Pemberton, J. M., and P. R. Fisher. 1977. 2,4-D Plasmids and persistence. Science 268:732-733.

62. Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin. 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172:2351-2359.

63. Perkins, E. J., and P. F. Lurquin. 1988. Duplication of a 2,4-dichlorophenoxyacetic acid monooxygenase gene in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. 170:5669-5672.

64. Peyton, T. O. 1984. Biological disposal of hazardous waste. Enzyme Microb. Technol. 6:146-154.

65. Pierce, G. E., T. J. Facklam, and J. M. Rice. 1981. Isolation and characterization of plasmids from environmental strains of bacteria capable of degrading the herbicide 2,4-D. Dev. Ind.

Microbiol. 22:401-408.

66. Reddy, K. J., R. Webb, and L. A. sherman. 1990. Bacterial RNA isolation with one hour centrifugation in a table-top ultracentrifuge. BioTechniques 8:250-251.

67. Reineke, W., and H.-J. Knackmuss. 1988. Microbial degradation of haloaromatics. Annu. Rev. Microbiol. 42:263-287.

68. Rothmel, R. K., T. L. Aldrich, J. E. Houghton, W. M. Coco, L. N. Ornston, and A. M. Chakrabarty. 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the *lysR* family. J. Bacteriol. 172:922-931.

69. Rothmel, R. K., D. L. Shinabarger, M. R. Parsek, T. L. Aldrich, and A. M. Chakrabarty. 1991. Functional analysis of the *Pseudomonas putida* regulatory protein CatR: Transcriptional studies and determination of the CatR DNA-binding site by hydroxy-radical footprinting. J. Bacteriol. 173:4717-4724.

70. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.

71. Savard, P., L. Peloquin, and M. Sylvestre. 1986. Cloning of *Pseudomonas* sp. strain CBS3 genes specifying dehalogenation of 4-chlorobenzoate. J. Bacteriol. 168:81-85.

72. Schell, M. A., and E. F. Poser. 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. J. Bacteriol. 171:837-846.

73. Schell, M. A., and M. Sukordhaman. 1989. Evidence that the transcription activator encoded by the *nahR* gene of *Pseudomonas putida* is evolutionary related to the transcription activators encoded by the *nodD* gene of *Rhizobium*. J. Bacteriol. 171:1952-1959.

74. Schmidt, E., M. Hellwig, and H.-J. Knackmuss. 1983. Degradation of chlorophenols by a defined mixed microbial community. Appl. Environ. Microbiol. 46:1038-1044.

75. Schraa, G., M. L. Boone, M.S. M. Jetten, A. R. W. Neerven, P. J. Colberg, and A. J. B. Zehnder. 1986. Degradation of 1,4-dichlorobenzene by *Alcaligenes* sp. strain A175. Appl. Environ. Microbiol. 52:1374-1381.

76. Shanley, M. S., E. L. Neidle, R. E. Parales, and L. N. Ornston. 1986. Cloning and expression of *Acinetobacter calcoaceticus catBCDE* genes in *Pseudomonas putida* and *Escherichia coli*. J. Bacteriol. 165:557-563.

77. Shields, M. S., S. W. Hooper, and G. S. Sayler. 1985. Plasmid mediated mineralization of 4-chlorobiphenyl. J. Bacteriol. 163:802-889.

78. Sheen, J.-Y., and B. Seed. 1988. Electrolyte gradient gels for DNA sequencing. Biotechniques 6:942-944.

79. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

80. Stanier, R. N., N. Palleroni, and M. Doudoroff. 1966. The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. 43:159-271.

81. Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature 323:640-643.

82. Stragier, P., O. Danas, and J.-C. Patte. 1983. Regulation and diaminopamilate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. J. Mol. Biol.168:2950-2955.

83. Streber, W., K. N. Timmis, and M. H. Zenk. 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxy-acetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. 169:2950-2955.

84. Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA 82:1074-1078.

85. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. 77:5201.

86. Tiedje, J. M., J. M. Duxbury, M. Alexander, and J. E. Dawson. 1969. 2.4-D metabolism: pathway of degradation of chlorocatechols by *Arthrobacter* sp. J. Agric. Food Chem. 17:1021-1026.

87. Vandenbergh, P. A., R. H. Olsen, and J. F. Colaruotolo. 1981. Isolation and genetic characterization of bacteria that degrade chloroaromatic compounds. Appl. Environ. Microbiol. 42:737-739.

88. van der Meer, J. R., W. M. de Vos, S. Harayama, and A. J. B. Zehnder. 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol. Reviews 56:677-694.

89. van der Meer, J. R., R. I. L. Eggen, A. J. Zehnder, and W. M. de Vos. 1991. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: Evidence for specilization of catechol 1.2-dioxygenases for chlorinated substrates. J. Bacteriol. 173:2425-2434.

90. van der Meer, J. R., A. C. Frijiters, J. H. Leveau, R. I. Eggen, A. J. Zehnder, and W. M. de Vos. 1991. Characterization of the *Pseudomonas* sp. strain P51 gene *tcbR*, a *lysR*-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon, and analysis of the regulatory region. J. Bacteriol. 173:3700-3708.

91. van der Tweel, W. J. J., N. ter Burg, J. B. Kok, and J. A. M. de Bont. 1986. Biotransformation of 4-hydroxybenzoate from 4-chlorobenzoate by *Alcaligenes denitrificans* NTB-1. Appl. Microbiol. Biotechnol. 25:289-294.

92. Weightman, A. J., R. H. Don, P. R. Lehrbach, and K. N. Timmis. 1984. The identification of genes encoding haloaromatic catabolic enzymes and the construction of hybrid pathways for substrate mineralization, p. 47-80. In G. S. Omenn and A. Hollaender (ed.), Genetic control of environmental pollutants. Plenum Press, New York.

93. Yanisch-Perron, C., J. Vieira and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.

94. Zaitsev, G. M., and Y. N. Karasevich. 1984. Utilization of 2-chlorobenzoic acid by *Pseudomonas cepacia*. Mikrobiologiya 53:75-80.

95. Zaitsev, G. M., and Y. N. Karasevich. 1985. Preparatory metabolism of 4-chlorobenzoic and 2,4-dichlorobenzoic acids in *Corynebacterium sepedonicum*. Mikrobiologiya 54:356-359.

VITA

Udayakumar Matrubutham

Candidate for the Degree of

Doctor of Philosophy

Thesis: SEQUENCING AND EXPRESSION OF THE REGULATORY GENE *tfdR* OF PLASMID pJP4

Major Field: Microbiology

Biographical:

Personal Data: Born in Coimbatore, Tamilnadu, India, May 31, 1961, the son of Muthuswamy and Renganayaki Matrubutham; married Sowmya Thiagarajan on June 21, 1991.

Education: Graduated from Suburban High School, Coimbatore, India, in May 1977; graduated Pre-University Course from Madras University at Madras, India, in May 1978; received Bachelor of Science degree in Agriculture from Tamilnadu Agricultural University, Coimbatore, India, in May 1982; received Master of Science degree in Microbiology from Tamilnadu Agricultural University, Coimbatore, India, in May 1984; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1993.

Professional Experience: Teaching Assistant, Department of Microbiology and Molecular Genetics, Oklahoma State University, August, 1988 to May 1991; Research Assistant at National Pulses Research Institute, India, June, 1985 to June 1988.

Fellowships: UNDP/FAO/ICAR Junior Research Fellow 1982-84.