SUBCLONING AND NUCLEOTIDE SEQUENCE OF TWO POSITIVE ACTING
REGULATORY GENES, \textit{xylR} AND \textit{xylS}, FROM THE
\textit{Pseudomonas putida} HS1 TOL PLASMID PDK1

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Teh-Tsai Chang, B. S., M. S.
Denton, Texas
May, 1992
SUBCLONING AND NUCLEOTIDE SEQUENCE OF TWO POSITIVE ACTING
REGULATORY GENES, xylR AND xylS, FROM THE
Pseudomonas putida HS1 TOL PLASMID PDK1

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Teh-Tsai Chang, B. S., M. S.
Denton, Texas
May, 1992
Chang, Teh-Tsai, Subcloning and Nucleotide Sequence of Two Positive Acting Regulatory Genes, xylR and xylS, from the Pseudomonas putida HS1 TOL Plasmid pDK1, Doctor of Philosophy (Molecular Biology), May, 1992, 241 pp., 15 tables, 41 figures, bibliography, 307 titles.

TOL plasmids of Pseudomonas putida encode enzymes for the degradation of toluene and related aromatics. These genes are organized into two operons regulated by the XylR and XylS transcriptional activators. Previous analysis of the TOL pDK1 catechol-2,3-dioxygenase gene (xylE) and a comparison of this gene to xylE from the related TOL plasmid pWW0, revealed the existence of a substantial level of sequence homology (82%). The majority of the observed base substitutions were silent, having no effect on the amino acid sequences of the encoded proteins. The comparison of the pDK1 and pWW0 xylRS region sequences described here reveals a much higher level of homology in this region. Here the homology is in excess of 96% for a 3.3 kb segment extending from 200 bp downstream of xylR to the end of the xylS coding region. The XylR proteins encoded by the two TOL plasmids were also compared. The DNA sequences in the coding regions for XylR were 98% identical, giving a predicted protein homology of 98.4%. In the case of xylS, the predicted amino acid sequences were found to be 95% identical as encoded by
DNA sequences which were 96% identical. This exceedingly high level of nucleotide sequence homology, not observed in comparisons of the meta-cleavage genes on enzymes suggests that the xylRS regions of pDK1 and pWW0 share a common ancestral sequence with a more recent origin than the meta-cleavage operons of the two plasmids.
ACKNOWLEDGEMENTS

I am deeply indebted to the following faculty members for their guidance and friendship. Drs. Kester, Knesek, Wu, Shanley, Donahue and O'Donovan were very helpful and very friendly. Dr. Robert C. Benjamin provided me with an excellent environment in which to complete my experiments. Without Dr. Benjamin's help and encouragement, I would not have been able to obtain the results that I report here. Second, I would like to take this opportunity to thank my fellow graduate students working on the fourth floor of the Biology Building. Finally, I would like to thank my family members, my parents Mr. T.-S. Chang and Mrs. T. H. Chang, my wife Cindy, my son Michael and my daughter Winnie for all the support they have given during this period of time. Because of their sacrifices, my studies here in the United States became possible.
TABLE OF CONTENTS

LIST OF TABLES ......................................................... ix

LIST OF ILLUSTRATIONS ............................................. x

LIST OF ABBREVIATIONS .............................................. xiii

CHAPTERS

I. INTRODUCTION ..................................................... 1

Biodegradation and *Pseudomonas*

Historical Background

Channelling of Aromatics to Dihydric Phenols Either

by Monooxygenases or Dioxygenases

The ortho-Cleavage of the Dihydroxylated Aromatic

Rings

A Review of Known Catabolic Plasmids

The Archetypal TOL Plasmid pWW0 and Its RP4

Cointegrate Plasmid pTN2

TOL Plasmid pWW53 and Its RP4 Cointegrate Plasmid

pWW53-4

TOL Plasmid pDK1 and Its RP4 Cointegrate Plasmid pDKR1

Other TOL Plasmids

Genetic Organization of TOL Plasmids and the

Duplication of Catabolic DNA Sequences
The Upper Pathway Operon
The meta-Cleavage Pathway Operons of TOL Plasmids
Transposons
   Transposons: Tn4651 and Tn4653
Regulation of xyl Gene Expression
TOL Regulatory Proteins
   The Role of the Regulatory Protein XylR
   The Role of the Regulatory Protein XylS
Promoters
   $P_u$ and $P_s$ Belong to a Specific Promoter Class
   $P_r$ and $P_m$ Belong to a Well Known Promoter Class
The Regulation of Lower Pathway Gene Expression
Objectives of This Study
   Subcloning of xylRS Region of the TOL Plasmid pDK1
   Determination of the Nucleotide Sequence of the xylRS Region
   Computer Analysis of DNA and Predicted Amino Acid Sequences

II. MATERIALS AND METHODS ........................................... 58

Isotopes, Chemicals and Restriction Endonucleases
Biological Media and Growth Conditions
Bacterial Strains and Plasmids
Working Cultures and Stocks
Isolation of E. coli Plasmid Vectors
Preparative Scale DNA Cleavage by Restriction Endonucleases
DNA Analysis Using Agarose Gel Electrophoresis
Molecular Weight Markers
Determination of DNA Concentrations
Isolation of Plasmids From Pseudomonas putida
Phenol Extraction and Ethanol Precipitation
Separation of Cloned DNA Fragments From the Vector by
Preparative Vertical Agarose Gel Electrophoresis
Recovery of DNA Fragments from Agarose Gels by
Electroelution
Gel Photography Using Polaroid Positive/Negative Films
Ligation
Preparation of Competent *E. coli* Cells
Transformation
Toothpick Assay and the Verification of Insertional
Inactivation
Alkaline Lysis Plasmid Isolation Method
Physical Characterization of Recombinant Clones
Screening for Restriction Endonuclease Cleavage Sites
DNA Analysis by Acrylamide Gel Electrophoresis
Cleavage of Electroeluted DNA by Restriction
Endonucleases
Subcloning DNA Fragments Into M13mp18
Preparation of Single-Stranded Templates for DNA
Sequencing
Dideoxyribonucleotide DNA Sequencing Reactions Using
Single-Stranded Templates
DNA Sequencing Gel Electrophoresis
Autoradiography of DNA Sequencing Gels Containing
$^{35}$S-labelled Nucleotides
DNA Sequencing Reactions Using Double-Stranded Plasmid
DNA Templates
Safety Precautions

III. RESULTS .........................................................102

Subcloning of the *xylRS* DNA Fragment from the TOL
Plasmid pDKR1
Construction of the Recombinant Plasmid pBK990
Physical Characterization of Recombinant
Clones
Preparative Scale Isolation of the Recombinant Plasmid pBK990
Purification of the Cloned xylRS Fragment
Screening for Restriction Endonuclease Cleavage Sites
Subcloning TOL Fragments into M13mp18
Determination of xylRS Nucleotide Sequence Using the Dideoxyribonucleotide Method
Preparation of Recombinant Single-Stranded Templates for DNA Sequencing Reactions
DNA Sequence Analysis of the Cloned xylRS Region
dNTP Reactions
dITP Reactions
Dideoxy Sequencing with Synthetic Oligonucleotide Primers
Extended Distance DNA Sequencing Analysis
DNA Sequences of the xylRS Region
xylRS DNA Sequence
xylR DNA Sequence and Encoded Amino Acid Sequence
xylS DNA Sequence and Encoded Amino Acid Sequence
Computer Analysis of DNA and Amino Acid Sequences
Computer Analysis of DNA Sequences
Comparison of xylR DNA Sequences Between TOL Plasmids pDK1 and pWW0
Comparison of xylS DNA Sequences Between TOL Plasmids pDK1 and pWW0
Comparison of DNA Sequences of the xylR xylS Intergenic Region
Computer Analysis of Amino Acid Sequences
Comparison of the Amino Acid Sequences of XylR Proteins Encoded by the Two TOL Plasmids, pDK1 and pWW0
Comparison of the Amino Acid Sequences of XylS Proteins Encoded by the Three TOL Plasmids,
pDK1, pWW0 and pWW53

IV. DISCUSSION and CONCLUSIONS.............................. 188

V. BIBLIOGRAPHY................................................. 196
LIST OF TABLES

1. Important scientists in the field of aromatic metabolism.............................................2
2. Plasmids encoding catabolic functions..............................................12
3. Aromatic compounds degraded by *P. putida* mt-2............................................13
4. Selected plasmids and incompatibility groups.......................................................16
5. TOL catabolic plasmids..........................................................................................23
6. Gene organizations of TOL regions from three related TOL plasmids pDK1, pWW0 and pWW53..................................................27
7. The upper pathway operon of TOL plasmid pWW0..............................................29
8. The meta-cleavage pathway operon of pWW0.......................................................35
9. Examples of dioxygenases......................................................................................37
10. TOL region recombinant subclones.........................................................................41
11. Transposons............................................................................................................45
12. Minimal medium for pseudomonads........................................................................60
13. Bacterial strains and plasmids................................................................................62
14. Genotype of *E. coli* host DH5α............................................................................68
15. Comparison of different classes of TOL promoters............................................184
LIST OF ILLUSTRATIONS

1. Benzoate and substituted benzoate degradation pathways..5
2. The restriction endonuclease cleavage map of the 40 kb pDK1 TOL region ........................................... 21
3. Gene organizations of TOL regions from three TOL plasmids ......................................................... 26
4. Meta-cleavage pathway ........................................ 33
5. Transposons Tn4651 and Tn4653................................. 44
6. Regulation of the upper and lower operons...............49
7. Flow chart showing the strategy utilized for the construction of recombinant plasmid pBK990 and the determination of the nucleotide sequence of the 3.9 kb BamHI fragment of this plasmid..................56
8. Genetic map of E. coli strain K-12 DH5α .................... 64
9. Plasmid vector pBR325............................................. 66
10. Genetic map and detailed multiple cloning site of bacteriophage M13mp18 ............................................. 70
11. Construction of the recombinant plasmid pBK990.......104
12. Electrophoretic analysis of restriction endonuclease products from the recombinant plasmid pBK990 .......108
13. Identification of the recombinant subclone pBK990.....111
14. Restriction endonuclease cleavage map showing the
orientation of the xylR and xylS genes within the 12.5 kb recombinant plasmid pBK990

15. Agarose gel analysis of the 3.9 kb BamHI fragment encoding xylRS

16. Screening of the 3.9 kb BamHI fragment for restriction endonuclease cleavage sites

17. The DNA sequencing strategy for the xylR and xylS genes

18. Single-stranded templates prepared from recombinant M13 plaques

19. Example of an autoradiogram from dNTP DNA sequencing reactions

20. Examples of M13mp18/19 DNA sequencing primers

21. Autoradiogram of a DNA sequencing gel showing the radioactive products resulting from two sets of dITP sequencing reactions

22. The restriction endonuclease cleavage map of the xylS region

23. Autoradiogram showing nucleotide sequence analysis using double-stranded templates and oligonucleotide primers

24. Nucleotide sequence of the xylRS region

25. The restriction endonuclease cleavage map of the xylRS region

26. The restriction endonuclease cleavage map of the xylR region
27. Codon usage for xylR ........................................150
28. The nucleotide and predicted amino acid sequences of xylR
from the TOL plasmid pDK1..........................152
29. Codon usage for xylS ........................................156
30. The nucleotide and predicted amino acid sequences of xylS
from the TOL plasmid pDK1..........................158
31. Kyte and Doolittle hydrophobicity profile of XylR......162
32. Kyte and Doolittle hydrophobicity profile of XylS.....164
33. Comparison of xylR nucleotide sequences from pWW0 and
   pDK1..................................................166
34. Comparison of the amino acid sequences of XylR proteins
   from the TOL plasmids pDK1 and pWW0............171
35. Comparison of xylS nucleotide sequences from pWW0 and
   pDK1..................................................173
36. GC content of the xylRS intergenic region.............177
37. σ^54- and σ^70-dependent promoters of the TOL plasmids pWW0
   and pDK1..............................................179
38. Comparison of the putative ribosome binding sites for xyl
   genes of the TOL plasmids pDK1 and pWW0........181
39. Comparison of xylRS operator-promoter regions from pWW0
   and pDK1..............................................183
40. Comparison of XylS amino acid sequences from pWW0,
   pDK1 and pWW53 (S_3) ..............................186
41. Comparison of the XylRXylS region of related TOL
   plasmids ............................................194
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenine</td>
</tr>
<tr>
<td>A&lt;sub&gt;550&lt;/sub&gt;</td>
<td>absorbance at 550 nanometer</td>
</tr>
<tr>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>ampicillin resistance</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>Cm&lt;sup&gt;R&lt;/sup&gt;</td>
<td>chloramphenicol resistance</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleotide</td>
</tr>
<tr>
<td>dITP</td>
<td>deoxyinosine nucleotide triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDaltons</td>
</tr>
<tr>
<td>Km&lt;sup&gt;R&lt;/sup&gt;</td>
<td>kanamycin resistance</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>μg</td>
<td>microgram(s)</td>
</tr>
<tr>
<td>μl</td>
<td>microliter(s)</td>
</tr>
<tr>
<td>MOPS</td>
<td>[3-(N-Morpholino)propanesulfonic acid]</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer(s)</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>RBS</td>
<td>ribosome binding site</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
</tbody>
</table>
T thymine
TcF tetracycline resistance
TAE tris-acetate-EDTA
TBE tris-borate-EDTA
TE tris-EDTA
Tris-HCl tris(hydroxymethyl)aminomethane hydrochloride
V volts
w/v weight per volume
15-mer a 15-nucleotide oligonucleotide
CHAPTER I

INTRODUCTION

Biodegradation and *Pseudomonas*

Aromatic hydrocarbons are widely distributed in the environment and the study of microbial pathways has been the primary focus of a number of research laboratories throughout the world (Table 1). The oxidation of these aromatic hydrocarbons is primarily due to the biodegradative activities of bacteria and fungi. A number of bacterial genera can utilize hydrocarbons as sole sources of carbon and energy, the specific biodegradation pathways involved depending upon the particular substrate as well as the genetic composition of the microorganism.

One reason that *Pseudomonas* sp. are ideal microorganisms for the study of biodegradation is that pseudomonads are widely distributed in soil and water. The ability of *Pseudomonas putida* strains to degrade aromatic compounds has been intensively studied (Burlage et al., 1989; Harayama et al., 1990). Genes coding for the oxidative catabolism of aromatic compounds can be of either chromosomal or plasmid origin. In the case of microorganisms belonging to the genus
Table 1. Important scientists in the field of aromatic metabolism

<table>
<thead>
<tr>
<th>Name</th>
<th>Nationalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>W. O. Tausson</td>
<td>Russia</td>
</tr>
<tr>
<td>Osamu Hayaishi</td>
<td>Japan</td>
</tr>
<tr>
<td>Frank Happold</td>
<td>U. K.</td>
</tr>
<tr>
<td>W. C. Evans</td>
<td>U. K.</td>
</tr>
<tr>
<td>I. C. Gunsalus</td>
<td>U. S. A.</td>
</tr>
<tr>
<td>L. Nicolas Ornston</td>
<td>U. S. A.</td>
</tr>
<tr>
<td>Peter A. Williams</td>
<td>U. K.</td>
</tr>
<tr>
<td>Kenneth N. Timmis</td>
<td>Germany</td>
</tr>
<tr>
<td>Teruko Nakazawa</td>
<td>Japan</td>
</tr>
<tr>
<td>David T. Gibson</td>
<td>U. S. A.</td>
</tr>
<tr>
<td>B. W. Holloway</td>
<td>Australia</td>
</tr>
<tr>
<td>Simon Silver</td>
<td>U. S. A.</td>
</tr>
<tr>
<td>H.-J. Knackmuss</td>
<td>Germany</td>
</tr>
<tr>
<td>Shigeaki Harayama</td>
<td>Japan</td>
</tr>
<tr>
<td>F. Christopher H. Franklin</td>
<td>Sweden</td>
</tr>
<tr>
<td>Sachiye Inouye</td>
<td>Japan</td>
</tr>
<tr>
<td>A. M. Chakrabarty</td>
<td>U. S. A.</td>
</tr>
<tr>
<td>J. L. Ramos</td>
<td>Spain</td>
</tr>
<tr>
<td>P. J. Chapman</td>
<td>U. K.</td>
</tr>
</tbody>
</table>
Pseudomonas, the biodegradation of aromatic hydrocarbons occurs by two major pathways (Fig. 1). The chromosome-encoded ortho-fission pathway is utilized when the substrates are benzoates or phenols and utilizes enzymes of high substrate specificity. The TOL plasmid-encoded meta-fission pathway is expressed when methylated benzoates, xylenes, toluene or benzene are substrates and is comprised of enzymes of relatively lower substrate specificity (Williams and Murray, 1974; Worsey and Williams, 1975).

**Historical Background**

*Pseudomonas arvilla* mt-2 was isolated in Japan during the 1950's. In 1959, Dagley and Stopher demonstrated that there were two pathways in the oxidative metabolism of aromatic compounds by microorganisms, the ortho-cleavage pathway and the meta-cleavage pathway (Dagley and Stopher, 1959). Japanese biochemists Nozaki et al. (Nozaki et al., 1963) used *Pseudomonas arvilla* mt-2 to study the enzyme metapyrocatechase (catechol-2,3-dioxygenase, C230) in the early 1960's. Simultaneous studies of catechol metabolism in other species demonstrated that catechol could be oxidized either through intradiol fission by catechol-1,2-oxygenase (ortho cleavage) or through extradiol fission by catechol-2,3-dioxygenase (meta cleavage). The original Japanese strain was later renamed *P. putida* mt-2. In 1966,
Figure 1. There are two pathways for the degradation of benzoate in *Pseudomonas putida*. The *ortho*-cleavage pathway is chromosomally encoded while the *meta*-cleavage pathway is plasmid encoded.
Benzoate
(Substituted benzoate)
\[ \rightarrow \]
cis-diols
\[ \rightarrow \]
para and unsubstituted derivatives
\[ \rightarrow \]
meta-Pathway
\[ \rightarrow \]
ortho-Pathway
\[ \rightarrow \]
\[ xyIE \]
Catechol-2-3-dioxygenase (C230)
\[ \rightarrow \]
\[ catA \]
Catechol-1,2-dioxygenase (C120)

\[ xyLG \]
2-hydroxymuconic Semialdehyde Dehydrogenase (HMSD)
\[ \rightarrow \]
\[ NADH \]
\[ xylF \]
2-hydroxymuconic Semialdehyde Hydrolase (HMSH)
\[ \rightarrow \]
\[ COOH \]
\[ COOH \]
\[ xyLG \]
4-oxalocrotonate Transacetylase (OT)
\[ \rightarrow \]
\[ COOH \]
\[ COOH \]
\[ xylH \]
4-oxalocrotonate Decarboxylase (OD)
\[ \rightarrow \]
\[ xyL \]
2-oxopenten-4-enoate Hydratase (OEH)
\[ \rightarrow \]
\[ xylK \]
4-hydroxy-2-oxoevalerate Aldolase (HOA)
\[ \rightarrow \]
\[ CH_3CSCOa + HOCCH_2CH_2COOH \]
aetyl CoA succinate

\[ OH \]
\[ xyL \]
2-hydroxymuconic Semialdehyde Hydrolase (HMSH)
\[ \rightarrow \]
\[ COOH \]
\[ COOH \]
\[ catC \]
Muconolactone Isomerase (MI)
\[ \rightarrow \]
\[ COOH \]
\[ COOH \]
\[ catD \]
Hydrolase I (HI)
\[ \rightarrow \]
\[ COOH \]
\[ COOH \]
\[ catE/J \]
\[ \rightarrow \]
\[ catF \]
\[ CH_3CSCOa + HOCCH_2CH_2COOH \]
aetyl CoA succinate
Stanier showed that *Pseudomonas putida* mt-2 decomposed benzoate through the *meta*-cleavage pathway, while most previously studied strains of pseudomonads degraded catechol by the *ortho*-cleavage pathway (Stanier, 1966). In 1973, Japanese microbiologists Nakazawa and Yokota described the differences between the *ortho*-cleavage and *meta*-cleavage pathways and suggested that a plasmid might encode the hydrocarbon catabolic enzyme genes of the *meta*-cleavage pathway.

It was not until the early seventies that scientists realized that *Pseudomonas* genes encoding many enzymes involved in hydrocarbon degradation were in fact plasmid-encoded. Degradative plasmids such as SAL, NAH, CAM, OCT and TOL were discovered during this period of time. In 1972, Chakrabarty of the University of Illinois demonstrated that *Pseudomonas putida* could use salicylate as a sole carbon source and that the genes encoding this ability were located on the salicylate plasmid (SAL plasmid). Soon afterwards, British scientists Williams and Murray (Williams, 1974) demonstrated that the genes for the *meta* pathway enzymes were located on the TOL plasmid pWW0. Since then, pWW0 has become one of the most studied *Pseudomonas* catabolic plasmids.

During the early 1980's, the structural and regulatory *xyl* genes of the 40 kb TOL region on plasmid pWW0 (115 kb) were intensively mapped and studied. This was carried out
primarily by three groups of scientists: K. N. Timmis (Switzerland), T. Nakazawa (Japan) and P. A. Williams (Wales). The degradation of toluene and xylenes by *P. putida* mt-2 has been shown to be due to an inducible enzyme pathway specified by the archetypal TOL plasmid pWW0 (Williams and Murray, 1974). Nakazawa and coworkers have isolated a recombinant plasmid, pTN2, which consists of the entire RP4 genome of 54 kb and a 56 kb segment of the TOL plasmid pWW0 (Nakazawa et al., 1980). The construction of the cointegrate plasmid pTN2 (RP4::TOL) was accomplished by the conjugation of a heat sensitive *P. aeruginosa* with a heat resistant *E. coli* strain (Nakazawa et al., 1980). pTN2 is transmissible to *E. coli*. Later experiments showed that the induction of the TOL pathway enzymes in cells of *P. putida* carrying pTN2 was similar to that of cells carrying the wild-type TOL plasmid. Subsequent work on pWW0 and pTN2 has identified most/all of the genes of the upper and lower pathways. This has been accomplished by a combination of molecular cloning and transposon mutagenesis experiments (Franklin et al., 1981; Inouye et al., 1981; Harayama et al., 1984; Timmis et al., 1985). Franklin and coworkers identified two *Pseudomonas putida* TOL plasmid regulatory genes, xylR and xylS, using Tn5 insertional inactivation and demonstrated the functions of their gene products (1983). Dowing and Broda reported a cleavage map of the TOL plasmid pWW0 (115 kb)
using the restriction endonucleases XhoI and HindIII (Dowing and Broda, 1979). Inouye and coworkers (Inouye et al., 1980) mapped the xyl genes by using molecular cloning techniques and manipulations of recombinant plasmids in E. coli. Franklin and coworkers mapped xyl genes using transposon mutagenesis and molecular cloning in P. putida (Frankling et al., 1985). The structural genes of the upper and the lower operons of the TOL region were mapped using the RP4::TOL recombinant plasmid pTN2.

In the late 1980's, DNA and predicted amino acid sequences of the xylRS genes on the TOL plasmid pWW0 were determined (Inouye et al., 1986; Spooner et al., 1986; Inouye et al., 1988).

In 1988, Tsuda and Iino demonstrated that at least two independent transposable elements were present on the TOL plasmid pWW0 (1988). Tn4651 is a 56 kb transposon and Tn4653 is a 70 kb transposon. Both Tn4651 and Tn4653 belong to the class II transposon family.

Pseudomonas putida strain HS1 carries the TOL plasmid pDK1 (125 kb). The TOL plasmid pDK1 is a close relative of the TOL plasmid pWW0. Strain HS1 was isolated from soil samples by Kunz and Chapman (Kunz and Chapman, 1981). Plasmid pDK1 has all of the degradative genes needed for the degradation of toluene and xylenes. Strain PaW630 (Trp\(^{-}\), Str\(^{r}\)), which carries the RP4 recombinant plasmid pDKR1 (95
kb), was constructed by P. A. Williams. This recombinant plasmid of pDKR1 has a 40 kb TOL segment (Shaw et al., 1988) from the wild-type pDK1 which encodes all of the pDK1 xyl genes required for the degradation of toluene and xylenes.

Channelling of aromatics to dihydric phenols either by mono-oxygenases or dioxygenases

Generally speaking, aromatic compounds are stable because of the difficulty in attacking the conjugated ring structures enzymatically. How can microorganisms such as Pseudomonas species transform aromatic hydrocarbons to central metabolites? A group of dioxygenases and monooxygenases form the core of a series of enzymes which are utilized to carry out this transformation. The hydroxylation reactions require either a mono- or dioxygenase, depending upon whether or not the substrate carries a hydroxyl group on the aromatic ring (Harayama et al., 1989). Examples of the reactions are given below:

monooxygenase

Phenol → Catechol
dioxygenase

Benzoate → Catechol

The ortho-cleavage of the dihydroxylated aromatic
rings

The chromosomally encoded β-ketoacidipate pathway (ortho-fission pathway) is also used widely for the microbial dissimilation of aromatic compounds (Stanier, 1973). Ortho cleavage incorporates two atoms of oxygen into catechol between the 2 hydroxylated carbon atoms, resulting in the formation of cis,cis-muconic acid (Nozaki et al., 1968). The ortho-cleavage pathway has an absolute specificity for catechol and cannot utilize its methylated derivatives (Feist and Hegeman, 1969). Chromosomal genes encoding the degradation of catechol (cat genes) and protocatechuate (pca genes) to succinate and acetyl-CoA via β-ketoacidipate are located in two separate clusters on the chromosomes of Acinetobacter calcoaceticus. Each individual cluster is then organized into two or three operons (Neidle et al., 1987). In addition, biochemical and genetic analysis of the β-ketoacidipate pathway enzymes of P. putida and A. calcoaceticus indicate a substantial evolutionary relatedness between the two. Several studies published by Ornston's group have suggested that the 11 structural genes needed for benzoate degradation via the β-ketoacidipate pathway are organized into associated transcriptional units within a 17 kb segment of the A. calcoaceticus chromosome. Benzoate degradation pathway genes encoded by the A. calcoaceticus chromosome include the benABCD and catABCDEFIJ genes (Neidle
et al., 1987). The benABC region encodes subunits a and b of the dioxygenase and the reductase. Furthermore, these scientists have also suggested that the benABC genes are evolutionally related to the TOL plasmid xylXYZ genes (Neidle et al., 1987; Harayama et al., 1991).

A review of known catabolic plasmids

Both TOL and SAL belong to a large family of catabolic plasmids. A couple of examples are shown in Table 2. Some specific characters of these catabolic plasmids are discussed below:

A. TOL plasmids

The TOL plasmid pWWO of Pseudomonas putida encodes enzymes for the degradation of some twenty aromatic hydrocarbons (including toluene, meta-, para-xylene, meta-ethyltoluene, 1,2,4-trimethylbenzene and their aldehyde, alcohol and acid derivatives) (Table 3). This plasmid is the most extensively characterized of the catabolic plasmids. Several other TOL plasmids, such as pDK1 and pWW53, have a similar genetic organization and are very closely related to the archetypal TOL plasmid pWW0.

B. NAH plasmids

The presence of a NAH plasmid enables its host to utilize naphthalene as a carbon and energy source. Cane and Williams studied one of the NAH plasmids, pWW60, and its
Table 2. Plasmids encoding catabolic functions

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Host</th>
<th>Size</th>
<th>Compounds catabolized</th>
</tr>
</thead>
<tbody>
<tr>
<td>pWW0</td>
<td><em>P. putida</em></td>
<td>115 kb</td>
<td>Toluene and xylenes</td>
</tr>
<tr>
<td>pDK1</td>
<td><em>P. putida</em></td>
<td>125 kb</td>
<td>Toluene and xylenes</td>
</tr>
<tr>
<td>NAH</td>
<td><em>P. putida</em></td>
<td>80 kb</td>
<td>Naphthalene</td>
</tr>
<tr>
<td>SAL</td>
<td><em>P. putida</em></td>
<td>80 kb</td>
<td>Salicylate</td>
</tr>
<tr>
<td>OCT-CAM</td>
<td><em>P. putida</em></td>
<td>&gt;300 kb</td>
<td>Alkanes and campher</td>
</tr>
<tr>
<td>pCS1</td>
<td><em>P. diminuta</em></td>
<td>80 kb</td>
<td>Parathion</td>
</tr>
<tr>
<td>Ti plasmid</td>
<td><em>A. tumorfecius</em></td>
<td>200 kb</td>
<td>Opines</td>
</tr>
<tr>
<td>Sym plasmid</td>
<td><em>Rhizobium sp.</em></td>
<td>150-350 kb</td>
<td>Flavonoids</td>
</tr>
</tbody>
</table>
Table 3. Aromatic compounds degraded by *P. putida* mt-2

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Aldehyde</th>
<th>Alcohol</th>
<th>Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toluene</td>
<td>Benzaldehyde</td>
<td>Benzy alcohol</td>
<td>Benzoate</td>
</tr>
<tr>
<td><em>p</em>-xylene</td>
<td><em>p</em>-methyl-</td>
<td><em>p</em>-methyl-</td>
<td><em>p</em>-toluate</td>
</tr>
<tr>
<td></td>
<td>benzaldehyde</td>
<td>benzy alcohol</td>
<td></td>
</tr>
<tr>
<td><em>m</em>-xylene</td>
<td><em>m</em>-methyl-</td>
<td><em>m</em>-methyl-</td>
<td><em>m</em>-toluate</td>
</tr>
<tr>
<td></td>
<td>benzaldehyde</td>
<td>benzy alcohol</td>
<td></td>
</tr>
<tr>
<td>3-ethyl-toluene</td>
<td>3-ethyl-</td>
<td>3-ethyl-</td>
<td>3-ethyl-</td>
</tr>
<tr>
<td></td>
<td>benzaldehyde</td>
<td>benzy alcohol</td>
<td>benzoate</td>
</tr>
<tr>
<td>1,2,4-trimethyl-benzene</td>
<td>2,4-</td>
<td>2,4-</td>
<td>2,4-</td>
</tr>
<tr>
<td></td>
<td>benzaldehyde</td>
<td>benzy alcohol</td>
<td>benzoate</td>
</tr>
</tbody>
</table>
derivative pWW60-22 (Cane and Williams, 1986). Restriction endonuclease mapping and transposon mutagenesis of pWW60-22 have shown that the early enzymes of the naphthalene catabolic pathway (naphthalene to salicylate) and the meta pathway genes (salicylate to pyruvate) are clustered in two distinct and physically separated operons (Cane and Williams, 1986). The meta-cleavage pathway genes and enzymes share extensive homology with those of TOL plasmids (Cane and Williams, 1986).

C. SAL plasmids

SAL plasmids allow the host organism to utilize salicylate as an energy and carbon sources. SAL plasmids are very closely related to NAH plasmids with regard to the encoded catabolic pathway and share extensive homology at the DNA and enzyme level (Gunsulus et al., 1989).

D. OCT plasmids

OCT plasmids allow the host cell to utilize alkanes as an energy and carbon sources (Chakrabarty, 1972 and 1978).

E. CAM plasmids

CAM plasmids allow the host organism to utilize camphor as an energy and carbon source (Chakrabarty, 1972 and 1978).

F. PAR plasmids

A plasmid encoding parathion degrading enzymes (PAR plasmid), pCMSI, was isolated from P. diminuta MG. The gene encoding the parathion hydrolase has been identified and its
nucleotide sequence determined (Wild et al., 1988; Mulbry et al., 1989).

**The archetypal TOL plasmid pWW0 and its RP4 cointegrate plasmid pTN2**

The TOL plasmid pWW0 (115 kb) specifies a set of enzymes for the catabolism of toluene and related compounds. The genetics and biochemistry of the pWW0 encoded pathway have been extensively investigated (Worsey et al., 1978; Inouye et al., 1983; Franklin et al., 1981, 1983; Harayama et al., 1984). The degradation of toluene and xylenes by *P. putida* has been shown to be due to an inducible enzyme pathway specified by the archetypal TOL plasmid pWW0 (Williams and Murray, 1974). Subsequent molecular cloning and transposon mutagenesis studies of pWW0 and pTN2, an RP4 cointegrate plasmid (Table 4) carrying the pWW0 catabolic genes, have located each of the genes of the upper and lower pathways (Franklin et al., 1981; Inouye et al., 1981; Harayama et al., 1984; Timmis et al., 1985). Franklin and coworkers identified two *Pseudomonas putida* TOL plasmid regulatory genes, *xylR* and *xylS*, using Tn5 insertional mutations, and demonstrated the functions of their gene products (Franklin et al., 1983).

Nakazawa and coworkers have isolated a recombinant plasmid, pTN2, which consists of the 54 kb RP4 replicon and a
Table 4. Selected plasmids and incompatibility groups

<table>
<thead>
<tr>
<th>Incompatibility group</th>
<th>Plasmid</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IncP1</td>
<td>RK2, RP1</td>
<td>60 kb</td>
<td>self-transmissible, broad host range</td>
</tr>
<tr>
<td>IncP2</td>
<td>CAM-OCT</td>
<td>&gt;300 kb</td>
<td>alkane utilization</td>
</tr>
<tr>
<td>IncP9</td>
<td>TOL</td>
<td>115 kb</td>
<td>hydrocarbon degradation</td>
</tr>
</tbody>
</table>
56 kb segment of the pWW0 TOL plasmid (Nakazawa et al., 1980). The construction of the cointegrate plasmid pTN2 (RP4::TOL) was accomplished by conjugating a heat sensitive P. aeruginosa strain with and a heat resistant E. coli strain (Nakazawa et al., 1980). Later experiments showed that the induction of the TOL pathway enzymes in cells of P. putida carrying pTN2 is similar to that of the wild-type TOL plasmid. pTN2 is transmissible to E. coli.

The structural genes of the upper and the lower operons of pTN2 were mapped by using molecular cloning techniques in E. coli. A restriction endonuclease cleavage map of pWW0 was first reported by Dowing and Broda in 1979. This group of scientists described a cleavage map for the restriction endonucleases XhoI and HindIII. Later, further genetic mapping of the cointegrate pWW0 derivative RP4::TOL (110 kb), was described by Inouye et al. (Inouye et al., 1980). Franklin and coworkers also mapped the xyl genes on the TOL plasmid pWW0 using transposition mutagenesis and molecular cloning in P. putida.

TOL plasmid pWW53 and its RP4 cointegrate plasmid pWW53-4

The TOL plasmid pWW53 was isolated in North Wales by P. A. Williams' group. This nonconjugative plasmid is approximately 105 to 110 kb in size. The TOL plasmid pWW53
encodes two divergent but homologous copies of the meta-pathway operon (Williams et al., 1987).

The RP4 conintegrate plasmid pWW53-4, containing about 35 kb of pWW53 DNA, encodes a set of inducible enzymes required for the oxidative degradation of toluene via the meta-cleavage of catechol (Keil et al., 1985; Williams et al., 1988). The genes responsible for this oxidation are organized into an upper pathway operon (toluene to benzoate) and a separate meta-cleavage pathway operon (benzoate to pyruvate). These are controlled respectively by the products of the regulatory genes xylR and xylS (Keil et al., 1985 and 1987). pWW53-4, which contains only one copy of the meta pathway operon, was constructed in vivo by transferring RP4 from P. putida AC34 (RP4) into P. putida MT53 using a filter mating technique. One Tc\textsuperscript{r} Cb\textsuperscript{r} Km\textsuperscript{r} mTol\textsuperscript{+} transconjugant containing both RP4 and pWW53 was then used as a donor in a filter mating experiment with PaW340 (Sm\textsuperscript{r}). Transconjugants able to grow on m-toluate in the presence of streptomycin were selected and from these pWW53-4 was isolated.

TOL plasmid pDK1 and its RP4 conintegrate plasmid pDKR1

The TOL plasmid pDK1 (125 kb), from Pseudomonas putida HS1, encodes the enzymes for the degradation of toluene, meta- and para-xylene to central metabolites. The catabolic genes of pDK1 from P. putida HS1 are also organized into two
operons. The upper pathway operon ($xylCMABN$) encodes enzymes for the degradation of toluene and xylenes ($meta$- and $para$-xylenes) to benzoate and toluates. The $meta$-cleavage pathway operon ($xylXYZLTEGFJQKIH$) encodes enzymes for the degradation of benzoate and toluates to acetaldehyde and pyruvate (Shaw et al., 1988). Map positions for $xyl$ genes within the TOL region of pDK1 (Fig. 2) have been determined by Shaw and Williams (1988) using its RP4 cointegrate pDKR1. Their results have shown that the genetic organization of the TOL plasmid pDK1 is quite similar to that of the archetypal TOL plasmid pWW0. The products of two regulatory genes, $xylR$ and $xylS$, regulate the expression of these two operons. These regulatory genes are located in a region residing between the upper and the lower operons.

The cointegrate plasmid RP4::TOL (pDKR1) carries a 40 kb TOL segment inserted into RP4. This region encodes all of the catabolic pathway genes for toluene degradation from the wild-type TOL plasmid pDK1.

Three steps were required in order to construct the RP4 cointegrate plasmid pDKR1. First, the plasmid RP4 was mated into the nonconjugative $P$. putida strain HS1 (pDK1). Transconjugants were selected as TOL$^+$ coding by RP4 antibiotic markers. Second, the HS1 transconjugant containing RP4 and pDK1 was used as a donor in a mating with a $P$. putida recipient (Sm$^F$). Finally, selection against the
Figure 2. The gene organization of the 40 kb TOL region of the RP4 cointegrate plasmid pDKR1 (100 kb). This figure was adapted from Shaw and Williams, 1988, page 2468.
donor was carried out using streptomycin. PaW630 was able to grow on both m-xylene and m-toluate and contained the smallest of the plasmids examined (pDKR1).

Other TOL plasmids

Other TOL plasmids (Table 5) also carry the catabolic genes required for metabolism of aromatic hydrocarbons.

1. pWW15

*Pseudomonas putida* MT15 contains a large plasmid, pWW15, of about 250 kb. This plasmid also encodes the genes for toluene and xylene catabolism (Keil and Williams, 1985).

2. pWW14 and pWW20

Although both pWW14 and pWW20 are some 40 kb larger than pWW15, the similarities in the DNA fragment distribution from restriction endonuclease digests suggests that these three TOL plasmids share a degree of sequence homology of probably more than 80% (Keil and Williams, 1985).

3. pRJ1 is a naturally occurring conjugative plasmid coding for toluene degradation and resistance to streptomycin and sulfonamides (Yano and Nishi, 1980).

4. *P. putida* strain R5-3 carries a 95 kb TOL plasmid. This strain was isolated from sewage sludge and has not been extensively characterized (Carney et al., 1989).
Table 5. TOL plasmids encoding catabolic functions

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Host</th>
<th>Size of plasmid</th>
<th>Geographic locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>pWW0</td>
<td><em>P. putida</em> mt-2</td>
<td>115 kb</td>
<td>Japan</td>
</tr>
<tr>
<td>pWW53</td>
<td>PaW53</td>
<td>105-110 kb</td>
<td>Wales</td>
</tr>
<tr>
<td>pDKl</td>
<td><em>P. putida</em> PpCl</td>
<td>125 kb</td>
<td>USA</td>
</tr>
<tr>
<td>pWW14</td>
<td>PaW14</td>
<td>300 kb</td>
<td>Wales</td>
</tr>
<tr>
<td>pWW15</td>
<td>PaW15</td>
<td>250 kb</td>
<td>Wales</td>
</tr>
<tr>
<td>pWW20</td>
<td>PaW20</td>
<td>300 kb</td>
<td>Wales</td>
</tr>
</tbody>
</table>
Genetic organization of TOL plasmids and the duplication of catabolic DNA sequences

P. A. Williams' and his associates have been studying the genetic organization of different TOL plasmids for nearly two decades. The genetic organization of the TOL regions from pWW0, pWW53, pDK1 and pWW15 have been described by this group of scientists during this period of time. Comparisons of the genetic organization of the TOL regions of the three plasmids (pWW0, pDK1 and pWW53) (Fig. 3 and Table 6) have revealed a number of similarities, as well as some differences. Keil et al. (1985) described the duplication of catabolic DNA sequences. O'Donnell and Williams subsequently identified a duplicated upper pathway operon on the 250 kb pWW15 TOL plasmid (1986). Osborne (Osborne et al., 1988) have also reported the duplication of catabolic DNA sequence. Williams have suggested the existence of a second xylS gene on TOL plasmid pDK1. Thus, the duplication of TOL catabolic sequences in TOL plasmids appears to be a common occurrence.

The upper pathway operon

*Pseudomonas putida* mt-2 can use twenty aromatic hydrocarbons as sole carbon and energy sources (toluene, *meta*-xylene, *para*-xylene, 3-ethyltoluene, 1,2,4-trimethylbenzene and each of their alcohols, aldehydes and acids). The upper pathway operon specifies enzymes
Figure 3. Gene organization of the TOL regions from three TOL plasmids. This figure was adapted from Shaw and Williams et al., 1988, page 2473.
RP4::TOL plasmid cointegrates

pDK R1

pWW53-4

pWW0

0 10 20 30 40 kbp
Table 6. Gene organizations of TOL regions from three related TOL plasmids pDK1, pWW0 and pWW53

<table>
<thead>
<tr>
<th>TOL</th>
<th>RP4::TOL</th>
<th>Organization of xyl genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDK1</td>
<td>pDKR1</td>
<td>&lt;----- &lt;-- --&gt; -------------- &gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NBAMC P_u R P_x P_s S P_m XYZLTEGFJQKIH</td>
</tr>
<tr>
<td>pWW53</td>
<td>pWW53-4</td>
<td>&lt;----- &lt;-- --&gt; -------------- &gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NBAMC P_u R P_x P_s S HIKQJFGETLZXY P_m</td>
</tr>
<tr>
<td>pWW0</td>
<td>pTN2</td>
<td>&lt;-- --&gt; &lt;-------------- &lt;-----</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R P_x P_s S HIKQJFGETLZXY P_m NBAMC P_u</td>
</tr>
</tbody>
</table>
(Table 7) that convert the aromatic compounds to their corresponding carboxylic acids. This operon is induced by toluene, xylenes and their alcohol derivatives. The induction is mediated by XylR, together with the σ^{54}-RNA polymerase complex. In the presence of substrate inducers, initiation of transcription from the OP1 promoter is activated by the XylR activator (Lorenzo et al., 1991). The methyl group at carbon 1 in the aromatic ring is sequentially oxidized to yield the corresponding carboxylic acids (Abril et al., 1989). In order to precisely map the positions of structural genes of the upper operon, Tn5 insertions located within or near the upper pathway genes were mapped and correlated with the phenotypes mediated by the mutant plasmids. Combined with Harayama's studies using maxicells a xylCMABN order had been determined for this operon (Jaraua, et al., 1986; Harayama et al., 1989).

Xylene monooxygenase (XO) has two subunits XylA and XylM. The activities of this xylene monooxygenase have been described by Harayama (Harayama et al., 1986 and 1989). Suzuki (1990) reported the nucleotide sequences of the two genes (1,110 bp for xylM and 1,053 bp for xylA) and the derived amino acid sequences of the protein products encoded by the xylMA region. Two proteins of 41 kDa (XylM) and 39 kDa (XylA) were suggested. Ramos (1991) demonstrated that
Table 7. The upper pathway operon of TOL plasmid pWW0

<table>
<thead>
<tr>
<th>Gene</th>
<th>Enzyme</th>
<th>Abbreviation</th>
<th>Nucleotide</th>
<th>M. W.</th>
</tr>
</thead>
<tbody>
<tr>
<td>xlyC</td>
<td>Benzaldehyde dehydrogenase</td>
<td>BZDH</td>
<td>1,617 bp</td>
<td>56 kDa</td>
</tr>
<tr>
<td>xylM</td>
<td>xylene monooxygenase</td>
<td>XO</td>
<td>1,110 bp</td>
<td>41 kDa</td>
</tr>
<tr>
<td>xylA</td>
<td>xylene monooxygenase</td>
<td>XO</td>
<td>1,053 bp</td>
<td>39 kDa</td>
</tr>
<tr>
<td>xylB</td>
<td>Benzyl alcohol dehydrogenase</td>
<td>BADH</td>
<td>1,146 bp</td>
<td>43 kDa</td>
</tr>
<tr>
<td>xynN</td>
<td>unknown</td>
<td>N. A.</td>
<td>1,416 bp</td>
<td>52 kDa</td>
</tr>
<tr>
<td>OPI</td>
<td>none</td>
<td>Pu</td>
<td>200 bp</td>
<td>none</td>
</tr>
<tr>
<td>xlyR</td>
<td>XylR</td>
<td>XylR</td>
<td>1,701 bp</td>
<td>64 kDa</td>
</tr>
</tbody>
</table>
xylene monooxygenase also has benzyl alcohol dehydrogenase activity. XO has a broad substrate specificity and can use even nitrobenzenes as substrates (Ramos et al., 1992). Since xylene monooxygenase is able to convert indole to the blue dye indigo, the xylMA region can potentially be used in a reporter function role in cloning vectors.

xylB is 1,146 bp in size and encodes the enzyme benzyl alcohol dehydrogenase (BADH). Chalmers et al. (1991) reported that TOL plasmid encoded benzyl alcohol dehydrogenase consists of a single polypeptide of 381 amino acids (43 kDa). Benzyl alcohol dehydrogenase (BADH) levels have been measured in bacteria containing the plasmid pWWO grown under noninducing (5 mM glucose as sole carbon source) and inducing (various aromatic substrates) conditions (Harayama et al., 1989).

xylC encodes the enzyme benzaldehyde dehydrogenase (BZDH). This protein has 538 amino acids and has a molecular weight of 56.3 kDa. Its function is very similar to the benzaldehyde dehydrogenase isolated from A. calcoaceticus (Chalmers et al., 1991).

xylN is 1,416 bp in size. The encoded polypeptide has a molecular weight of 52 kDa and it is post-translationally processed to a smaller protein of 43 kDa (Harayama et al., 1989). The function of this protein is still unknown.
The meta-cleavage pathway operons and enzymes of TOL plasmids

Harayama and coworkers used transposon mutagenesis to locate a number of the meta-cleavage pathway structural genes (Fig. 4 and Table 8) and the gene order xylD (toluate-1, 2-dioxygenase), xylL (1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase), xylE (catechol-2,3-dioxygenase), xylG (hydroxymuconic semialdehyde dehydrogenase), xylF (hydroxymuconic semialdehyde hydrolase), xylJ (2-oxopent-4-enoate hydratase), xylI (4-oxalocrotonate decarboxylase) and xylH (4-oxalocrotonate tautomerase) was established (Harayama et al., 1984).

The toluate-1,2-dioxygenase complex destabilizes the aromatic ring at the expense of NADH. Toluate-1, 2-dioxygenase, encoded by the TOL plasmid xylXYZ genes (originally designed only xylD as above by Harayama), is a broad substrate range dioxygenase. This aromatic ring dioxygenase is composed of three subunits. The nucleotide sequence of the xylXYZ and the three derived amino acid sequences have been reported (Harayama et al., 1991). This enzyme complex forms cis-diols from many C-3- and C-4-substituted benzoates, some disubstituted benzoates and benzoate itself (Harayama et al., 1986). The toluate-1,2-dioxygenase complex (TO) is responsible for changing benzoate and substituted benzoates into catechol and
Figure 4. Meta-cleavage pathway. This figure was adapted from Harayama et al., 1987.
substituted catechols. The formation of dihydric phenols (aromatic compounds carrying two hydroxyl groups) seems to be an important biochemical strategy of bacteria to destabilize the chemically stable resonant structure of the aromatic ring and thus to facilitate its subsequent opening (Dagley et al., 1986). One way to form dihydric phenols is by dioxygenation of the aromatic ring. This involves the incorporation of two atoms of molecular oxygen onto the ring to produce cis-diols, and is followed by dehydrogenation (Harayama et al., 1989). An example of this mechanism gave this below:

\[
\text{TO} \rightarrow \text{DHCDH} \\
\text{benzoate} \rightarrow \text{cis-dihydrodiols} \rightarrow \text{catechols}
\]

\[\text{xylXYZ} \rightarrow \text{xylL}\]

Toluolate-1,2-dioxygenase activity is readily measured in \textit{P. putida} \textit{mt-2} cells. TO activities were first assayed as the stimulation of oxygen uptake by whole cells following the addition of appropriate substrates to a final concentration of 0.2 mM. Measurements were made at room temperature using an oxygen electrode.

\textit{xylL} encodes 1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase which converts the \textit{cis}-dihydrodiols to catechols or substituted catechols. This \textit{cis}-dihydrodiol dehydrogenase is an NAD-dependent enzyme. \textit{E.}
Table 8. The meta-cleavage pathway operon of pWW0

<table>
<thead>
<tr>
<th>Gene</th>
<th>Enzyme or proteins</th>
<th>Nucleotides</th>
<th>M. W.</th>
</tr>
</thead>
<tbody>
<tr>
<td>xylX</td>
<td>Toluate-1,2-dioxygenase</td>
<td>1,365 bp</td>
<td>57 kDa</td>
</tr>
<tr>
<td>xylY</td>
<td>Toluate-1,2-dioxygenase</td>
<td>489 bp</td>
<td>20 kDa</td>
</tr>
<tr>
<td>xylZ</td>
<td>Toluate-1,2-dioxygenase</td>
<td>1,011 bp</td>
<td>39 kDa</td>
</tr>
<tr>
<td>xylL</td>
<td>1,2-Dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase</td>
<td>810 bp</td>
<td>28 kDa</td>
</tr>
<tr>
<td>xylT</td>
<td>Chloroplast-type ferredoxin</td>
<td>339 bp</td>
<td>12 kDa</td>
</tr>
<tr>
<td>xylE</td>
<td>Catechol-2,3-dioxygenase</td>
<td>921 bp</td>
<td>36 kDa</td>
</tr>
<tr>
<td>xylF</td>
<td>2-Hydroxymuconic semialdehyde hydrogenase</td>
<td>846 bp</td>
<td>30.6 kDa</td>
</tr>
<tr>
<td>xylG</td>
<td>2-Hydroxymuconic semialdehyde dehydrogenase</td>
<td>1,461 bp</td>
<td>51.7 kDa</td>
</tr>
<tr>
<td>xylJ</td>
<td>2-hydroxypent-2,4-dienoate hydratase</td>
<td>669 bp</td>
<td>23.9 kDa</td>
</tr>
<tr>
<td>XylK</td>
<td>2-Oxyhydroxypent-4-enolate aldolase</td>
<td>1,056 bp</td>
<td>39 kDa</td>
</tr>
<tr>
<td>xylI</td>
<td>4-Oxalocrotonate decarboxylase</td>
<td>807 bp</td>
<td>29 kDa</td>
</tr>
<tr>
<td>xylH</td>
<td>4-Oxalocrotonate tautomerase (4-Oxalocrotonate isomerase)</td>
<td>450 bp</td>
<td>3.5 kDa</td>
</tr>
<tr>
<td>xylR</td>
<td>XylR</td>
<td>1,701 bp</td>
<td>64 kDa</td>
</tr>
<tr>
<td>xylS</td>
<td>XylS</td>
<td>966 bp</td>
<td>36.5 kDa</td>
</tr>
</tbody>
</table>
coli K-12 cells containing a hybrid pBR322 plasmid carrying the xylXZYL genes transformed labeled benzoate to cis-dihydrodiol and catechol (Harayama et al., 1986). The Acinetobacter calcoaceticus benD gene also encodes a cis-diol dehydrogenase (Table 9). In spite of the fact that xylL and benD come from microorganisms of two different genera, the functions of their gene products are similar. DNA sequence comparisons between these two genes have suggested that they are evolutionarily very closely related (Harayama et al., 1991).

Catechol-2,3-dioxygenase (C230, metapyrocatechase) is encoded by the TOL plasmid xylE gene. This enzyme catalyzes the oxygenative cleavage of the catechol ring. Specifically, the catechol-2,3-dioxygenase catalyzes the incorporation of two atoms of oxygen into catechol or substituted catechols) between a hydroxylated carbon atom and an adjacent carbon atom carrying hydrogen (meta cleavage). Nozaki and coworkers purified and crystallized this enzyme in 1963.

Catechol-2,3-dioxygenase from Pseudomonas putida mt-2 is a ferrous iron containing dioxygenase and consists of four identical subunits (each has a molecular weight of 35,000). xylE has been sequenced and its gene product, catechol-2,3-dioxygenase, has been extensively characterized (Nakai et al., 1983).

Although pWW0 has only one C230 gene, some TOL plasmids
Table 9. Comparison among three multicomponent dioxygenases encoding enzyme complexes for benzoate degradation

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Size</th>
<th>Substrate</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toluate-1,2-dioxygenase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(TOL plasmid pWW0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylX</td>
<td>57 kDa</td>
<td>broad</td>
<td>Harayama et al.,</td>
</tr>
<tr>
<td>xylY</td>
<td>20 kDa</td>
<td></td>
<td>1986</td>
</tr>
<tr>
<td>xylZ</td>
<td>38 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzoate-1,2-dioxygenase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Pseudomonas arvilla C-1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benA</td>
<td>53 kDa</td>
<td>narrow</td>
<td>Yamaguchi et al.,</td>
</tr>
<tr>
<td>benB</td>
<td>20 kDa</td>
<td></td>
<td>1982</td>
</tr>
<tr>
<td>benC</td>
<td>38 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzoate-1,2-dioxygenase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Acinetobacter calcoaceticus)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benA</td>
<td>53 kDa</td>
<td>narrow</td>
<td>Neidle et al.,</td>
</tr>
<tr>
<td>benB</td>
<td>19 kDa</td>
<td></td>
<td>1987</td>
</tr>
<tr>
<td>benC</td>
<td>38 kDa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
may have two xylE genes (sometimes as part of two complete meta-cleavage pathway operons) (Chatfield and Williams, 1986). Keil (1985) observed that the TOL plasmid pWW15 (250 kb) has two nonhomologous C230 genes, while Osborne et al. (1988) demonstrated that pWW53 (110 kb) carried two homologous xylE genes. Also, H. Keil (1990) described a third catechol-2,3-dioxygenase in Azotobacter vinelandii which shows little homology to the previously described C230. In 1991, Benjamin and coworkers described the nucleotide sequence of xylE from TOL plasmid pDKl. Comparisons of xylE nucleotide and amino acid sequences among three isofunctional catechol-2,3-dioxygenases from pWW0, pDKl and NAH7 were then reported (Benjamin et al., 1991).

After ring-cleavage there are two principal catabolic routes of the meta-cleavage pathway leading to the Krebs cycle. These are the hydrolytic branch and the dehydrogenative (oxalocrotonate) branch (Harayama et al., 1987). Sala-Trepat and Evans first described the branching of the meta-cleavage pathway in Azotobacter sp. in 1971 (Sala-trepat et al., 1971). In 1987, Harayama et al. described in detail the branching of the meta-cleavage pathway in Pseudomonas (Harayama et al., 1987).

The hydrolytic branch involves only the xylF gene product, 4-hydroxymuconic semialdehyde hydrolase. XylF has been purified and studied by Duggleby and Williams (Duggleby
et al., 1986). XylF is active as a dimer with a monomer molecular weight of 32.5 kDa. The ring cleavage product derived from m-toluate is degraded by the xylF branch, as is the ring-fission product derived from o-toluate.

The dehydrogenative branch requires three enzymes: 4-hydroxymuconic semialdehyde dehydrogenase (XylG), 4-oxalocrotonate tautomerase (XylH) and 4-oxalocrotonate decarboxylase (XylI). Benzoate and p-toluate are degraded by the xylGHI branch. Some of the dehydrogenative branch pathway intermediates exist in two or three isomeric forms (Hamamoto et al., 1989). Differences in enzyme affinity of the dehydrogenase (xylG) toward the initial ring cleavage products plays a decisive role in determining which branch (the hydrolytic or dehydrogenative) each potential substrate enters. This enzyme is NAD$^+$ dependent. The 28 kDa tautomerase (xylH) is formed by the association of extremely small identical protein subunits with a molecular weight of 3,500. The decarboxylase (xylI) and the hydratase (xylJ) have been shown to form a complex in vivo. The formation of this complex may assure efficient transformation of the unstable intermediate, 2-hydroxypent-2,4-dienoate, in vivo (Hamamoto et al., 1989).

The metabolites produced by both the hydrolytic and dehydrogenative branches are directed towards the Krebs cycle by 2-hydroxypent-2,4-dienoate hydratase (XylJ) and
2-oxyhydroxypent-4-enolate aldolase (XylK). The resultant pyruvate and aldehydes are then finally transformed into Krebs cycle intermediates by other cellular enzymes.

Two gene products of the lower xyl operon, XylT and XylQ, have been identified, but their functions are still unknown at this time.

A DNA library of the 40 kb pDK1 TOL region has been constructed. The nucleotide sequence of the 15.5 kb pDK1 meta operon, including the promoter Pm and 13 xyl genes, as well as the xylRS regulatory region, have been determined by Benjamin's group (Hares et al., 1992).

The nucleotide sequences of three consecutive XhoI fragments (8 kb total) encoding the xylXYZLTE region (Benjamin et al., 1991), as well as a 4 kb BamHI fragment encoding the regulatory genes xylR and xylS, were determined initially (Chang and Benjamin, 1991). Later, nucleotide sequences of the xylGFJQKIH region were determined (Luo et al., 1992; Williamson et al., 1992). In addition, some characterization of the proteins encoded by these pDK1 xyl genes have been performed using cell-free extracts from subclones (Table 10) in expression vectors.

**Transposons**

After continuous growth with benzoate some TOL plasmids were "cured". This phenomena was characterized by the loss
Table 10. Hybrid plasmids carrying subcloned DNA fragments derived from the 40 kb TOL region of TOL plasmid pDK1

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Host</th>
<th>Vector</th>
<th>TOL genes</th>
<th>Size</th>
<th>Subclone</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBK990</td>
<td>DH5α</td>
<td>pBR325</td>
<td>xylR and xylS</td>
<td>12.5 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK489</td>
<td>DH5α</td>
<td>pBR322</td>
<td>The meta operon</td>
<td>19.8 kb</td>
<td>HindIII</td>
</tr>
<tr>
<td>PBK288</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylZL</td>
<td>4.3 kb</td>
<td>XhoI</td>
</tr>
<tr>
<td>PBK188</td>
<td>DH5α</td>
<td>pBR325</td>
<td>xylZLTE</td>
<td>11.2 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK189</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylTE</td>
<td>4.9 kb</td>
<td>XhoI</td>
</tr>
<tr>
<td>PBK190</td>
<td>JM101</td>
<td>pUC19</td>
<td>xylXYZ</td>
<td>7.7 kb</td>
<td>HindIII/KpnI</td>
</tr>
<tr>
<td>PBK789</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylXYZ</td>
<td>6.7 kb</td>
<td>HindIII/KpnI</td>
</tr>
<tr>
<td>PBK790</td>
<td>DH5α</td>
<td>pBR325</td>
<td>xylIGF</td>
<td>5.0 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK690</td>
<td>DH5α</td>
<td>pBR325</td>
<td>xylIQKI</td>
<td>9.3 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK289</td>
<td>DH5α</td>
<td>pBR325</td>
<td>xylCMABN</td>
<td>14.1 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK390</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylGFAQ</td>
<td>8.1 kb</td>
<td>XhoI</td>
</tr>
<tr>
<td>PBK490</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylIH</td>
<td>5.6 kb</td>
<td>XhoI</td>
</tr>
<tr>
<td>PBK491</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylIH</td>
<td>5.6 kb</td>
<td>XhoI</td>
</tr>
<tr>
<td>PBK590</td>
<td>DH5α</td>
<td>pUC19</td>
<td>xylIH</td>
<td>4.7 kb</td>
<td>EcoRI</td>
</tr>
<tr>
<td>PBK589</td>
<td>DH5α</td>
<td>pBR325</td>
<td>region</td>
<td>11.8 kb</td>
<td>EcoRI</td>
</tr>
</tbody>
</table>

* EcoRI fragments were inserted into the EcoRI site of pBR325
* XhoI fragments were inserted into the SalI site of pUC18/19
of the meta pathway genes. The spontaneous loss of the approximately 40 kb region encoding the toluene catabolic pathway and its regulatory genes as a result of benzoate curing led Chakrabarty to propose that a transposon may be associated with the TOL plasmid (Chakrabarty, et al., 1972). Tsuda and Iino demonstrated the presence of at least two independent transposable elements on the TOL plasmid pWW0. Tn4651 is a 56 kb transposon and Tn4653 is a 70 kb transposon (Tsudo and Iino, 1988) (Fig. 5). Tn4651 and Tn4653 are class II transposons (Tsudo and Iino, 1989) (Table 11).

Both Tn4651 and Tn4653 of TOL pWW0 are members of the Tn3 family. The genes on these two transposons required for the actual transposition have been identified, including the transposase genes, \textit{tnpA}, and two genes that are necessary for resolution, \textit{tnpS} and \textit{tnpT}. In addition, a specific DNA sequence required in \textit{cis} for resolution, the \textit{res} site, has been identified (Tsudo and Iino, 1988). Transposition of all Tn3 family transposons requires transposase (\textit{tnpA}) and resolvase (\textit{tnpR}). Genetic analysis of Tn4653 has demonstrated that its transposition required these two enzymes. Transposase is encoded in a 3.0 kb fragment at one end of Tn4653. For the resolution of Tn4651, resolvase is necessary (Tsudo and Iino, 1988).

\textbf{Regulation of xyl gene expression}
Figure 5. Transposons Tn4651 and Tn4653.
OP2 XYZTEGFJQKIH

OP1 CMABN

xyIS xyIR

TOL pWW0

tapA

IR

tapS res tapT

IR

merR merT merP merA merD res tnpR

IR

Tn4651

Tn501
Table 11. Transposons

<table>
<thead>
<tr>
<th>Transposon</th>
<th>Size</th>
<th>Specificity</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tn4651</td>
<td>56 kb</td>
<td>Hydrocarbon degradation</td>
<td>xyl genes, res, tnpS, tnpT and tnpA</td>
</tr>
<tr>
<td>Tn4653</td>
<td>70 kb</td>
<td>Hydrocarbon degradation</td>
<td>xyl genes, res, tnpS, tnpT and tnpA</td>
</tr>
<tr>
<td>Tn501</td>
<td>8.5 kb</td>
<td>Carries Mer operon</td>
<td>merR, merT, merP, merA, merD, res, tnpR and tnpA</td>
</tr>
<tr>
<td>Tn3</td>
<td>4.5 kb</td>
<td>Carrying antibiotic markers</td>
<td>Am\textsuperscript{r}, tnpR and tnpA</td>
</tr>
</tbody>
</table>
The catabolic genes of the TOL region are usually organized into two operons. The regulation of the expression of these catabolic genes requires two TOL-encoded positive regulatory proteins, *xylR* and *xylS*, two sigma factors, $\sigma^{54}$ and $\sigma^{70}$, and the RNA polymerase core enzyme.

Worsey, Franklin and Williams proposed a TOL plasmid regulatory cascade model in 1978 (Worsey et al., 1978). Inouye and coworkers (Inouye et al., 1984) demonstrated that the regulation of both the upper pathway and the lower pathway operons occurred at the transcriptional level. Nakazawa and coworkers demonstrated that *xylR* and *xylS* act as positive transcriptional regulators (Nakazawa et al., 1980). Inouye and coworkers have shown evidence that both *xylR* and *xylS* were required for the positive regulation of the meta-cleavage pathway operon. If both XylR and XylS are present, the lower pathway operon is inducible by the upper pathway substrates (Inouye et al., 1983). They also determined the positions of two regulatory genes *xylR* and *xylS* and two operator-promoter regions, OP1 and OP2 (Fig. 6). The regulatory genes are themselves subject to fine regulation in the TOL system (Inouye et al., 1984 and 1988).

The TOL substrates, e.g. toluene, enter the *P. putida* cell and combine with the XylR protein, which is produced constitutively. This binary complex then binds to promoters
P_u (OP1) and P_s (promoter for xylS). This binding promotes the binding of the RNA polymerase $\sigma^{54}$ holoenzyme, giving expression of the upper operon enzymes and xylS. Then the product (benzoate for example) of the upper pathway binds to XylS. This allows XylS to bind to P_m and promoted lower pathway expression.

**TOL regulatory proteins**

A. The role of the regulatory protein XylR

The three domains of the XylR protein each have their own unique functions. The central domain, which harbors a putative ATP-binding site, has been proposed as the region which interacts with the $\sigma^{54}$ RNA-polymerase. The two prokaryotic activator proteins, XylR and NifA (Azotobacter *vinelandii*), show a high degree of sequence similarity within their central domains, each of which contains a putative ATP-binding site. Control of operon expression by the use of an activator protein and ATP to catalyze the formation of transcriptionally productive open complex has been proposed (Abril et al., 1991). Both ATP and the central domain are required for the formation of the open complex between the $\sigma^{54}$-RNA polymerase holoenzyme and the promoter. Second, the C-terminal region of XylR contains the DNA-binding domain.
Figure 6. Regulation of the upper and lower TOL operons. This figure was adapted from Harayama and Timmis, 1989, page 162.
NtrA

Benzoate

Toluene

(Pu) xylCMABN  (Pm) xylXYZLEGFIKIH  xylS (Ps) (Pr) xylR
Finally, the N-terminal region is the effector binding domain interacting with the inducers of the operon.

XylR activates the transcription of the xylCMABN operon in the presence of m-xylene or m-methylbenzyl alcohol (and 23 other compounds). Transcription of xylS also requires the XylR activator. The sigma 54 factor (σ^{54}) is also required for P_s but not for the lower pathway promoter P_m (OP2). In addition, the XylR protein autoregulates its own synthesis from promoter P_R. XylR expression is repressed by its own product (Inouye et al., 1987). Also, in the presence of a substrate-effector for the upper pathway, the XylR protein stimulates expression of the lower operon indirectly, by stimulating a 25-fold increase in transcription of xylS. The XylS protein interacts with benzoate and benzoate analogues (products of the upper operon enzymes) to activate transcription from the lower operon promoter P_m.

It has been observed that two other factors involved in the transcription of σ^{54}-dependent promoters, NtrC and NifA from Klebsiella pneumoniae, can stimulate transcription from P_u and can substitute for XylR (Dixon, 1986).

XylR exhibits a very broad range effector specificity, being able to recognize as effectors a wide variety of unsubstituted, mono- and disubstituted methyl, ethyl-, and
chlorotoluenes, benzyl alcohols, and \( p \)-chlorobenzaldehyde (Abril et al., 1989).

**B. The role of the regulatory protein XylS**

XylS protein is a DNA binding protein with the helix-turn-helix domain located in the C-terminal portion of the polypeptide. It acts as a positive regulatory protein (Inouye, et al., 1986). The XylS protein is a positive activator of the TOL plasmid meta-cleavage pathway operon in \( P. \) putida. After binding to an inducer such as \( m \)-toluate, XylS protein enhances the transcription of the lower pathway genes from \( P_m \). Mermod and coworkers (Mermod et al., 1987) demonstrated that overproduction of the \( xylS \) gene product leads to constitutive expression of the meta-cleavage pathway operon. Transcription of \( xylS \) requires the \( \sigma^{54} \)-RNA polymerase and a special form (possibly a phosphorylated form according to Shiau et al., 1992) of the positive regulatory protein XylR.

**Promoters**

\( xyl \) genes of the TOL plasmid pDK1 are organized in four transcription units and each transcription unit has its own promoter. Among these four promoters, \( P_u \) (OP1) and \( P_s \) (\( xylS \) promoter) belong to the \( \sigma^{54} \)-dependent promoter class and \( P_r \)
and Pm (OP2) belong to the well-known $\sigma^{70}$-dependent promoter class.

A. P_u and P_s belong to a specific promoter class

Two XylR activated promoters, P_u and P_s, share homology to nitrogen metabolism promoters ($ntr$ and $nif$ promoters) (Johnston and Downie, 1984; Reznikoff et al., 1985; Inouye et al., 1984 and 1987). In the case of the best studied $\sigma^{54}$-dependent promoters, the $nif$ promoters in *Klebsiella pneumoniae*, the activator proteins behave as DNA-binding proteins. The XylR-activated promoters on the *Pseudomonas* TOL plasmid (P_u and P_s) are characterized by conserved sequences in the -24 and -12 regions, the expression from which requires the specific sigma factor $\sigma^{54}$ (Abril et al., 1991). Efficient transcriptional initiation from these promoters requires the specific activator protein XylR.

Upstream regulatory protein binding sequences (UBS) are often inverted repeats that represent the binding sites for activator proteins, such as NifA in *Klebsiella*, NtrC in *Escherichia coli*, DctD in *Rhizobium* and XylR in *P. putida*. An inverted repeat ATTTGN_2CAAAT was identified within the xylRS regulatory region (Abril et al., 1991). A requirement
for upstream sequences that could serve as the XylR activator binding sites has been demonstrated. Two XylR binding sites (UBS's) have been located -160 and -130 upstream from the transcription start (+1) (Abril et al., 1991). Interaction between activator proteins and $\sigma^{54}$-RNA polymerase is required for transcriptional initiation (Abril et al., 1989).

B. $P_r$ and $P_m$ belong to a well known promoter class

Host integration factor (IHF protein) was also shown to be required for DNA bending and hence for the initiation of transcription. Thus, DNA looping might be necessary in the initiation of transcription of xyl genes.

Transcription of xylR (from Pr) is constitutive and unaffected by the addition of aromatic inducers (Inouye et al., 1985). xylR is expressed well in E. coli systems and the -10/-35 sequence of the $P_r$ promoter region can be recognized by the $\sigma^{70}$-RNA polymerase complex (Reznikoff et al., 1985).

Initiation of transcription from the well known -10 -35 consensus promoters requires RNA polymerase core enzyme as well as the sigma factor $\sigma^{70}$ in E. coli. The location of the initiation of transcription of the meta-cleavage pathway operon of TOL plasmid pWW0 was determined using the S1 nuclease mapping method (Inouye et al., 1985).
The regulation of lower pathway gene expression

Studies of the TOL plasmid pWW0 originally confirmed that catabolic genes for toluene degradation are organized into two operons on this and related plasmids. The expression of the two operons is regulated by the product of two regulatory genes: xylR and xylS. If P. putida cells carry the TOL plasmid, which specifies a meta-cleavage pathway, the chromosomally-encoded β-ketoacid pathway is not induced. The lower pathway needs both xylR and xylS products for induction. Note that as mentioned previously XylR is required to make XylS. In the presence of lower pathway substrates (e.g., benzoate), the interaction of the xylS gene product with the OP2 operator-promoter region is sufficient for induction of the lower pathway.

Objectives of this study

A. Subcloning of xylRS (the regulatory region) of TOL plasmid pDK1 (Fig. 7)

With the increasing interest in the TOL plasmid catabolic pathways, the need for a detailed genetic analysis by way of extensive DNA sequencing of both the structural genes and the regulatory genes of the TOL region has become urgent. In order to better understand how catabolic plasmids (TOL plasmids for example) regulate the expression of their
Figure 7. Flow chart showing the strategy utilized for the construction of recombinant plasmid pBK990 and the determination of the nucleotide sequence of the 3.9 kb BamHI fragment (the xylRS region) of this plasmid.
Construction of pBK990 (12.5 kb)

BamHI digest

pBK990 BamHI fragment (3.9 kb)

HaeIII, Sau3AI, TaqI, AluI, HincI and BstUI fragments

Subcloning into M13mp18 MCS

Select colorless plaques carrying inserts on E. coli DH5αF' lawns

Recombinant M13 single-stranded templates

Sanger dideoxy sequencing method

Computer analysis of sequence data
degradative genes, detailed genetic fine structure analysis of their regulatory genes is necessary.

Two regulatory genes, xylR and xylS, are located within a 6.5 kb EcoRI fragment of TOL plasmid pDKR1. The pBR325 hybrid plasmid constructed was later identified and designated as pBK990.

B. Determination of the nucleotide sequence of the xylRS region

One of the primary goals of this study was to determine the nucleotide sequence of the pDK1 xylRS region using the Sanger dideoxy chain-termination method.

C. Computer analysis of DNA and predicted amino acid sequences

The third objective of this study was to carry out computer-based comparison of the xylR and xylS sequences with other related sequences obtained from nucleic acid data banks.
CHAPTER II

MATERIALS and METHODS

Isotopes, chemicals and restriction endonucleases

Radioactive materials ([α-\textsuperscript{35}S]dATP and [α-\textsuperscript{32}P]dATP) were obtained from New England Nuclear (Boston, MA) and Amersham Co. (Arlington Heights, IL). Chemicals and bacteriological media were purchased from Difco, Fisher Scientific, Sigma Chemical Co. (St. Louis, MO), and New England BioLabs (Beverly, MA). Sequenase version 2.0 DNA sequencing kits were purchased from United States Biochemical Corporation. T4 DNA ligase was supplied by New England BioLabs. Restriction endonucleases and restriction buffers were purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, MD), New England BioLabs, Inc., Pharmacia LKB Biotechnology, Inc. (Piscataway, NJ), Boehringer Mannheim Biochemicals (Indianapolis, IN) and IBI. Bacterial alkaline phosphatase (BAP) was obtained from Bethesda Research Laboratories and IBI. Isotopes and Sequenase kits were used according to the manufacturer's instructions. Restriction
enzymes, T4 DNA ligase and bacterial alkaline phosphatase were also used as suggested by the suppliers or as noted below.

**Biological media and growth conditions**

Media used in this study were YT broth, LB broth, LA broth and M9 minimal medium. These were supplemented with antibiotics where appropriate. YT broth was composed of 0.8% (w/v) tryptone, 0.5% (w/v) yeast extract and 0.5% (w/v) NaCl. Agar (1.5% w/v) was added in order to make solid media for plates. For the propagation of *E. coli* phages such as M13 and its derivatives (M13mpl8 and M13mpl9), both YT broth and YT agar plates were used. *E. coli* cultures were grown in LB (Luria-Bertani) medium (Miller, 1972) which contained (per liter) 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl. LA medium was composed of 1.0% tryptone (w/v), 0.5% yeast extract (w/v) and 0.5% NaCl (w/v). Both LA broth and LA agar plates were used to maintain *P. putida* strains. *P. putida* strains were also grown on minimal media (Table 12) with either 5 mM *m*-toluate or 5 mM benzoate added as the carbon source. For example, 5 mM *m*-toluate minimal medium plates were made as follows: 50 ml of a 100 mM *m*-toluate stock solution was added to 100 ml of a 10X PN (phosphate and nitrogen) stock solution (91 g KH₂PO₄, 16.8 g NaOH and 12 g
Table 12. Minimal medium for pseudomonads

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>KH$_2$PO$_4$</td>
<td>4.25 g</td>
</tr>
<tr>
<td>NaH$_2$PO$_4$</td>
<td>1.0 g</td>
</tr>
<tr>
<td>NH$_4$Cl</td>
<td>2.0 g</td>
</tr>
<tr>
<td>MgSO$_4$•7H$_2$O</td>
<td>0.2 g</td>
</tr>
<tr>
<td>FeSO$_4$•7H$_2$O</td>
<td>0.012 g</td>
</tr>
<tr>
<td>MnSO$_4$•H$_2$O</td>
<td>0.003 g</td>
</tr>
<tr>
<td>ZnSO$_4$•7H$_2$O</td>
<td>0.003 g</td>
</tr>
<tr>
<td>CoSO$_4$</td>
<td>0.001 g</td>
</tr>
<tr>
<td>nitritotriacetic acid</td>
<td>0.01 g</td>
</tr>
<tr>
<td>water-soluble substrate</td>
<td>0.05% (w/v)</td>
</tr>
<tr>
<td>(m-toluate, benzoate)</td>
<td></td>
</tr>
<tr>
<td>agar</td>
<td>20 g (per liter)</td>
</tr>
</tbody>
</table>

Catechol was added at the concentration of 1% (w/v).
(NH₄)₂SO₄ per liter stock solution). Trace metals were provided by adding 5 ml of a 200X R-salt stock solution (made by mixing 400 ml 10% MgSO₄·7H₂O, 100 ml 1% FeSO₄ and 2 ml conc. HCl). Amino acids such as leucine and tryptophan were added to a final concentration of 50 µg/ml if required. Twenty grams of agar (2.0% w/v) was added and the final volume adjusted to 1 liter. Antibiotics were added as necessary at the following concentrations for E. coli: ampicillin (Ap), 100 µg/ml; tetracycline (Tc), 25 µg/ml; kanamycin (Km), 50 µg/ml; streptomycin (Sm), 100 µg/ml; chloramphenicol (Cm), 25 µg/ml and nalidixic acid (Na), 20 µg/ml. Antibiotics added to media for P. putida were as follows: ampicillin (Ap), 100 µg/ml; tetracycline (Tc), 50 µg/ml; kanamycin (Km), 50 µg/ml; streptomycin (Sm), 200 µg/ml and carbenicillin (Cb), 2,000 µg/ml.

E. coli and P. putida cultures were grown at 37°C and 30°C, respectively.

**Bacterial strains and plasmids**

Bacterial strains and plasmids used are listed in Table 13. *Escherichia coli* DH5α (Fig. 8) was the bacterial host strain utilized for cloning purposes and the plasmid vectors used were pBR322, pBR325 (Fig. 9) and pUC19. E. coli
Table 13. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td><em>lacZΔM15recA</em></td>
<td>Woodcock <em>et al.</em>, 1989&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>DH5αF'</td>
<td>F'</td>
<td>Woodcock <em>et al.</em>, 1989&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>JM101</td>
<td>recA&lt;sup&gt;-&lt;/sup&gt;</td>
<td>Yanish-Perron <em>et al.</em>, 1985&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>HB101</td>
<td>Sm&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Boyer <em>et al.</em>, 1969&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>P. putida</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS1</td>
<td>TOL</td>
<td>Kunz and Chapman, 1981</td>
</tr>
<tr>
<td>PaW630</td>
<td>a derivative of PaW1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RP4::TOL</td>
<td>Kunz and Chapman, 1981&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBR325</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Prentki <em>et al.</em>, 1981&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>pBK990</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pDK1</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Km&lt;sup&gt;R&lt;/sup&gt;, Sm&lt;sup&gt;R&lt;/sup&gt;, Nc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Kunz and Chapman, 1981&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>pDKR1</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Km&lt;sup&gt;R&lt;/sup&gt;, Sm&lt;sup&gt;R&lt;/sup&gt;, Nc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Kunz and Chapman, 1981&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>RP4</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Tc&lt;sup&gt;R&lt;/sup&gt;, Km&lt;sup&gt;R&lt;/sup&gt;, Sm&lt;sup&gt;R&lt;/sup&gt;, Nc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Datta <em>et al.</em>, 1971</td>
</tr>
<tr>
<td>pKT231</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;, Sm&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Timmis <em>et al.</em>, 1989&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>an IncQ plasmid</td>
<td></td>
</tr>
</tbody>
</table>

1. Purchased from Bethesda Research Laboratories
2. A gift from D. A. Kunz
3. R. C. Benjamin
Figure 8. Genetic map of *E. coli* strain K-12 DH5α. This figure was adapted from New England Biolabs 1990-1991 catalog page 102.
**Escherichia coli**

4,720,000 base pairs
Figure 9. Plasmid vector pBR325.
EcoRI 0.00
BamHI 0.38

EcoRI 4.78

pBR325
6.00 Kb

Cm

Tc

Ap

ORI

PstI 3.61
DH5αF' (F' supE44 Δ(lacZYA-argF)U169 (φ80lacZΔM15)

_hsdR17(rK−mK+) recA1 endA1 gyrA96 thi-1 relA1_ (Table 14) was used as the host for the propagation of _E. coli_ phages. M13mp18 (Fig. 10) and its recombinants were propagated in _E. coli_ DH5αF'. Plasmid constructions were made by using either the _E. coli_ plasmid vectors pBR325 (Ap^R^ Tc^R^ Cm^R^) or the _P. putida_ plasmid vector pKT231 (Km^R^ Sm^R^). _E. coli_ plasmid pBR325 (Ap^R^ Tc^R^ Cm^R^) was used to subclone the _xylRS_ region from the recombinant plasmid pDKR1 (RP4::TOL) (Shaw et al., 1988). The broad host range plasmid pKT231 (Km^R^ Sm^R^), which was originally constructed by K. N. Timmis' laboratory in Germany (_E. coli_ MM294/pKT231) was used in the attempt to construct a subclone of the meta-cleavage pathway promoter (P_m) and the regulatory gene _xylS_. _E. coli_ MM294 (supE44 _hsdR endA1 pro thi_) carrying pKT231 was a gift from D. A. Kunz (Department of Biological Sciences, University of North Texas).

A DNA library of the TOL region from pDKR1 was constructed in this laboratory using _E. coli_ plasmid vectors. These recombinant subclones were used for DNA sequencing and protein analysis purposes. Two in particular were utilized in the study reported here.

a. pBK990: a pBR325 clone with a 6.5 kb TOL inserted
Table 14. Genotype of *E. coli* host DH5α (Hanahan, 1983)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>endA1</td>
<td>endonuclease mutation</td>
<td>for plasmid isolation</td>
</tr>
<tr>
<td>gyrA96</td>
<td>DNA gyrase mutation</td>
<td>nalidixic acid resistance</td>
</tr>
<tr>
<td>lacZΔM15</td>
<td>partial deletion of lacZ</td>
<td>X-gal (complementation of lacZ gene)</td>
</tr>
<tr>
<td>recA1</td>
<td>recombination deficient</td>
<td>prevent recombination</td>
</tr>
<tr>
<td>supF</td>
<td>suppressor mutation</td>
<td>UAG mutation</td>
</tr>
<tr>
<td>thi-1</td>
<td>mutation in thiamine metabolism</td>
<td>thiamine required</td>
</tr>
</tbody>
</table>
M13mp18
7,250 base pairs

Polylinker cloning sites
6230-6288

(Met) Thr Met Ile Thr Asn Ser Ser Ser Val Pro Gly Asp Pro Leu Glu Ser Thr Cys Arg His Ala Ser Leu Ala → LacZ'

atgcatatgattacaattCGAGCTCGGTTACCCGGGATCTAGTCGACCGGCGCGATGCAAGCTTGTG

Sac I  Xma I  Xba I  Pst I  Hind III
EcoR I  Kpn I  BamHI  Sal I  Hinc II  Acc I  Sph I
into the vector's *EcoRI* site. This segment encodes the *xylRS* region of pDK1.

b. pBK489: a pBR322 clone with a 15.5 kb TOL segment inserted into the vector's *HindIII* site. This segment encodes the entire *meta*-cleavage pathway operon (*xylXYZLTEGFJQKIH* region).

Both *P. putida* HS1 (pDK1) and *P. putida* PaW630 (RP4::TOL) were kindly provided by D. A. Kunz of the Department of Biological Sciences, University of North Texas.

**Working cultures and Stocks**

All *E. coli* and *Pseudomonas* sp. strain stocks were kept in 40% glycerol at -70°C for long term storage. *E. coli* working stock cultures were maintained on LB agar slants and stored at 4°C. A second set of *E. coli* culture stocks was stored at -20°C in LB broth containing 40% glycerol. A reserve set of *Pseudomonas putida* strains was also maintained at -20°C in 40% glycerol.

**Isolation of *E. coli* plasmid vectors**

Plasmid pBR325 and pUC19 DNA's were isolated by the CsCl/ethidium bromide equilibrium centrifugation method of Tanaka and Weisblum (1975). *E. coli* strain DH5α carrying the desired plasmid (pBR325) was cultured overnight in 50 ml LB
broth. One liter of LB broth in a 2.8 liter Fernbach flask was inoculated with 10-50 ml of the overnight culture and incubated in a New Brunswick shaker incubator at 37°C at 250 rpm. When the absorbance at 550 nm (A$_{550}$) reached 0.8, 2 ml of 85 mg/ml chloramphenicol dissolved in 95% ethanol was added. Incubation was continued for an additional 15 hours in order to allow the amplification of the *E. coli* plasmid. *E. coli* cells were harvested by centrifugation in a Sorvall GSA-3 rotor at 10,000 xg for 8 min. Harvested cells were washed once with 0.15 M NaCl, transferred to Oak Ridge Style tubes and treated with lysozyme (5 mg/ml) in 25% sucrose-50 mM Tris-hydrochloride (pH 8.0) for 5 min at 0°C. An incubation for another 5 min in the presence of 6.25 mM EDTA was followed. Next, 5 ml of a 5 M NaCl solution and 2 ml 10% sodium dodecyl sulfate (SDS) were added in quick succession and mixed thoroughly. The mixture was incubated for 2 h in a 4°C cold room on ice and centrifuged at 4°C with a SA600 rotor at 30,000 xg for 1 h. The supernatant containing the plasmid DNA was precipitated by adding an equal volume of isopropanol and incubating at -20°C overnight. The precipitate was collected by centrifugation and dissolved in TE buffer containing 200 µg/ml of heat-treated RNase with vigorous stirring. The mixture was centrifuged at 12,000 xg for 10 min and the pellet discarded. CsCl was added at a concentration of 1.06 g per ml of supernatant. Samples were
transferred to Sorvall Ultracrimp ultracentrifuges tube and ethidium bromide added to a final concentration of 0.5 mg per milliliter. Ultracentrifuge tubes were carefully balanced (± 0.02 g) and properly sealed. These sealed ultracentrifuge tubes were then centrifuged at 15°C at 100,000 xg for 40 h in a Sorvall Ti1270 rotor. The supercoiled DNA band was collected, and extracted three times with butanol (saturated with TE and 5M NaCl) to remove the ethidium bromide. CsCl was removed by ethanol precipitation as follows. The plasmid containing CsCl solution was diluted with two volumes of ddH₂O. Then, nine volumes of ice-cold absolute ethanol was added to precipitate the DNA. The DNA was redissolved in 0.3 M sodium acetate. The alcohol precipitation was then repeated. This was followed by a 70% EtOH wash, and the pellet was dried in vacuo. The plasmid DNA was redissolved in 500 µl TE buffer and the DNA concentration calculated from its $A_{260/280}$ (1 absorbance unit = 50 µg/ml double-stranded DNA) using a Beckman DU-40 spectrophotometer. Plasmid DNA preparations were properly labeled and stored at -20°C.

**Preparative scale DNA cleavage by restriction endonucleases**

Approximately 500 µg plasmid DNA was digested with 50-200 units of the selected enzyme. The amount of enzyme
and time of digestion varied depending upon the enzyme utilized. A 1% agarose mini-gel electrophoretic analysis was performed on a small quantity of the sample (~1 µg) to verify the completion of the restricting reaction. Agarose mini-gel electrophoresis was performed at 80 volts for 1 h and the gel was stained in a 0.05 µg/ml EtBr solution for 20 min. This was then destained in deionized water for 30 min. DNA bands were visualized using a long wave UV light. The remainder of the sample was then processed for fragment purification as described below.

DNA analysis using agarose gel electrophoresis

Both 1% and 0.7% agarose gels were used for the analysis of DNA samples. Using a 1% agarose gel as an example, 1.0 g of agarose was dissolved in 100 ml of TBE buffer in a 300 ml flask. Before melting the agarose, the flask was weighed. The agarose suspension was brought to boil for 3 min in a microwave oven. In order to replace the water evaporated during this melting process, distilled water was then added back until the flask and its contents were restored to their original weight. An acrylic gel forming tray was prepared by sealing both ends with masking tape. When the temperature of the melted agarose cooled to approximately 55°C, 30-50 ml of the contents were poured into the sealed tray. A well forming comb was placed at one end
of the tray. After the agarose solidified, the masking tape was removed from both ends and the plate was placed in an electrophoresis apparatus containing the 1X TBE electrophoresis buffer (containing per liter 5.0 g boric acid, 4 ml EDTA of 0.5 M, pH 8.0 and 10.8 g Tris-HCl). After loading DNA samples, the power supply was connected to the apparatus with two wire leads. The voltage and the time of electrophoresis were adjusted according to experimental requirements. Generally, agarose gel electrophoresis was performed at 80 volts until the xylene cyanol migrated about two-thirds of the length of the gel.

Molecular Weight Markers

Lambda HindIII fragments were routinely used as DNA molecular weight markers.

Determination of DNA concentrations

A small portion of the CsCl purified plasmid preparation was diluted fifty times (20 µl of plasmid was diluted with 980 µl of TE buffer for example) and the diluted sample was scanned using a Beckman™ DU-40 spectrophotometer. The DNA concentration was then calculated from its $A_{260/280}$ (1 absorbance unit = 50 µg/ml double-stranded DNA) and the dilution factor.
Isolation of plasmids from *Pseudomonas putida*

Both the isolation of *Pseudomonas putida* TOL plasmid pDK1 (125 kb) and that of the cointegrate plasmid pDKR1 (100 kb) were carried out according to Johnston and Gunsalus (1977). *Pseudomonas putida* PaW630 was grown overnight in 50 ml LA broth. One liter of LA broth in a 2.8 liter Fernbach flask was inoculated with 10-50 ml of the overnight culture and incubated at 30°C in a New Brunswick G25 shaker incubator (250 rpm). This large culture was then allowed to grow overnight. *P. putida* cells were harvested by centrifugation at 10,000 xg for 8 min in a Sorvall GSA-3 rotor at 4°C. The resulting pellets were washed once with 0.05 M Na$_2$HPO$_4$-KH$_2$PO$_4$ buffer (pH 7.0). The cells from the 1 liter culture were resuspended in 100 ml of 50 mM Tris-25% sucrose solution. Next, 20 ml of 5 mg/ml lysozyme solution was added and the suspension was incubated on ice for twenty min. At this time 20 ml 0.25 M EDTA (pH 8.0) was added. Sarcosyl-DOC solution (50 g Sarcosyl 97, 10 g desoxycholic acid and ddH$_2$O added to a final volume of 200 ml, pH 8.0) was then slowly added with vigorous stirring. The samples were then incubated on ice in a 4°C cold room for 2 h. After a clearing spin at 27,000 rpm for 30 min in a Du Pont Sorvall ultracentrifuge using either a Ti647.5 rotor or a Ti1270 rotor, 0.25 volumes of 50% polyethylene glycol (PEG-8,000) was added to the supernatant.
and the mixed suspension incubated on ice in a 4°C refrigerator overnight. The mixture was then collected in a Sorvall GSA-3 rotor at 10,500 rpm for 10 min. The pellet was redissolved in 20 ml TES buffer (containing 50 mM NaCl, 5 mM EDTA and 50 mM Tris-Cl, pH 8.0). A CsCl-EtBr gradient purification was then performed. CsCl was added at a rate of 1.06 g per milliliter of TE. Samples were transferred to Sorvall Ultracrimp tubes and ethidium bromide added to a final concentration of 0.5 mg/ml. The ultracentrifuge tubes were carefully balanced (± 0.02 g) and properly sealed. These sealed ultracentrifuge tubes were then centrifuged at 15°C at 110,000 xg for 40 h with a Sorvall Ti1270 rotor. The plasmid bands were collected through the sides of the tubes with a 3 cc syringe and 20 G needle. Ethidium bromide was then removed from samples by three butanol extractions. CsCl was removed by ethanol precipitation as described previously. The isolated plasmid DNA was utilized for restriction endonuclease digestion, plasmid profiling and cloning/transformation experiments (purified through CsCl-EtBr gradients twice in this case).

**Phenol extraction and ethanol precipitation**

One volume of phenol was added to the sample which was then vortexed thoroughly. The microcentrifuge tube was then centrifuged at 10,000 rpm for 1 min. The phenol phase was
removed with a drawn out Pasteur pipette. This was repeated once. One volume of diethyl ether was then added and the sample vortexed thoroughly. It was then centrifuged at 10,000 xg for 1 min. Ether was removed with a drawn out Pasteur pipette and the tube was placed in a chemical fume hood at room temperature for several minutes to allow residual ether to evaporate. The sample was then centrifuged at 10,000 xg for 10 min to remove solids and the supernatant carefully transferred to a fresh microcentrifuge tube using a drawn out Pasteur pipette. One tenth volume of 3 M sodium acetate and 3 volumes of 95% ice-cold EtOH were added. The microcentrifuge tube was vortexed and placed in the -70°C freezer for 10 min. The resulting precipitate was collected by centrifugation at 10,000 xg for 10 min. The supernatant was carefully removed with a drawn out Pasteur pipette. The pellet was washed with 70% EtOH which was removed following centrifugation at 10,000 xg for 5 min. The final pellet was then dried in a Speed Vac™ vacuum concentrator and dissolved in 25 µl TE. DNA samples were stored at -20°C.

**Separation of cloned DNA fragments from the vector by preparative vertical agarose gel electrophoresis**

A 20 by 20 cm by 1.5 mm 1% agarose gel was prepared. Five hundred micrograms of digested recombinant plasmid DNA was loaded onto a 6 cm wide sample well. Gel electrophoresis
was performed at 80 volts for 6 h with constant buffer circulation. After removing the top glass plate from the gel cassette, the agarose gel was illuminated under short wave UV light. The DNA band containing the desired fragment was identified and cut out from the agarose gel with a razor blade.

**Recovery of DNA fragments from agarose gels by electroelution**

Electroelution was performed according to published methods described by Maniatis et al. (1982). The dialysis tubing was cut into pieces about 12 cm long. These pieces were boiled in 1 liter of 2% sodium bicarbonate and 1 mM EDTA (pH 8.0). The tubing was rinsed with distilled water and then boiled in 1 liter of 1 mM EDTA (pH 8.0). These boiled pieces of tubing were again rinsed with distilled water and stored at 4°C in 0.01% sodium azide.

One end of the tubing was sealed with a knot or tubing clip. The gel slice was carefully transferred into the dialysis tubing. Five hundred microliters of 0.25X TBE buffer was added and the remaining end of the tubing sealed with a second dialysis bag clip. The electroelution was performed at 80 volts for 4 h. The current was then briefly reversed for 20 seconds before turning off the power supply. The buffer with the dialysis tube was collected and
transferred to a 1.5 ml microcentrifuge tube. The tubing was washed with 100 µl 0.25X TBE and the sample and the wash were combined. In order to get rid of the unwanted residue from the sample (such as any remaining agarose), a phenol extraction was performed.

One volume of phenol was added and the microcentrifuge tube was centrifuged at 10,000 xg for 1 min. The phenol phase was removed with a drawn out Pasteur pipette. This was repeated once. One volume of diethyl ether was then added. The mixture was vortexed and centrifuged at 10,000 xg for 1 min. Ether was removed with a drawn out Pasteur pipette and the microcentrifuge tube was placed in a chemical fume hood at room temperature for several min to allow residual ether to evaporate. Samples were then centrifuged at 10,000 xg for 10 min and supernatants carefully transferred to fresh microcentrifuge tubes with a pipetor. The transferred supernatants were centrifuged at 10,000 xg for 5 min. The supernatants were then again transferred to fresh microcentrifuge tubes with a pipetor. The volume of each of the preparations was then reduced to approximately 250 µl using a Speed Vac vacuum concentrator. One tenth volume of 3 M sodium acetate (25 µl) and 3 volumes of 95% ice-cold EtOH (750 µl) were then added. The content of the microcentrifuge tubes were vortexed and these were placed in a -70°C freezer for 10 min. The resulting precipitate was collected by
centrifugation at 10,000 xg for 10 min. The supernatant was carefully removed with a drawn out Pasteur pipette and discarded. Pellets were washed with 70% EtOH, which was removed after centrifugation at 10,000 xg for 5 min. The final pellet was dried in a Speed Vac vacuum concentrator. The electroeluted DNA was dissolved in 25 µl TE and stored at -20°C.

**Gel photography using Polaroid positive/negative films**

After electrophoresis, the agarose gel was stained in 0.5 µg/ml ethidium bromide solution for 15 min and then destained in distilled water for 20 min. The destained gel was drained and carefully placed on a Fotodyne System long wave length UV transilluminator (310 nm). The height of the camera was properly adjusted until a clear picture was seen. The red filter was then placed back to its correct position under the lens. A Polaroid type 55 positive/negative film was properly inserted into the camera (inserted the film until a click sound was heard and then pulled the envelope back from the camera). An f/stop of 4.5 and a 1 min exposure was commonly used, depending upon the DNA concentrations of the samples. In cases where DNA was limited, longer exposures were used. Next, the film envelope was placed back into the camera and the switch to the right of the camera was
moved to the process position (P). This releases the film which is then pulled from the camera. After allowing 20 seconds for the film to develop, the outer envelope was opened. The picture (the positive) was carefully removed and coated with a layer of the fixer/hardener. The negative was then farther processed in a 9% sodium sulfite solution for one min and rinsed several times with tap water (about 10 min with running water is best).

**Ligation**

**A. Preparation of the linearized vector DNA**

The *E. coli* plasmid cloning vector was digested with the selected restriction endonuclease and the linearized DNA dephosphorylated by treatment with bacterial alkaline phosphatase. In order to inactivate the bacterial alkaline phosphatase, dephosphorylated plasmid DNA samples were extracted twice with phenol. Phenol was then removed by ether extraction and the DNA concentrated by ethanol precipitation.

**B. Preparation of the target DNA**

The target DNA was cleaved with the same restriction endonuclease used for vector cleavage. In order to verify that the digestion reaction was complete, electrophoretic analysis of the sample was performed using a 1% agarose mini-gel.
C. Ligation of DNA's for transformation

A digested plasmid vector was dephosphorylated by bacterial alkaline phosphatase treatment as described above. This vector DNA (0.6 μg) was mixed with the digested target DNA (0.2 μg target DNA in TE, pH 8.0). Ligation of these fragments was carried out in a total volume of 20 μl. Four μl of 5X ligation buffer (containing, 25% PEG-8,000, 50 mM magnesium chloride, 5 mM ATP, 5 mM DTT and 250 mM Tris-HCl, pH 7.6) and 10 units of T4 DNA ligase (10 units/μl) were included in this volume. The ligation mixture was prepared in a 1.5 ml polypropylene microfuge tube on ice and then 1 μl of T4 DNA ligase was added last to initiate the reaction. The ligation mixture was incubated at room temperature for 5 h and then used to transform competent *E. coli* cells or stored at -20°C until used. Competent *E. coli* DH5α cells were prepared as described below and used as the bacterial recipients for such transformations.

**Preparation of competent *E.coli* cells**

Competent *E. coli* cells were prepared as follows: 500 μl of an overnight *E. coli* DH5α culture was transferred to 50 ml of YT broth. The diluted culture was then incubated at 37°C in a New Brunswick G25 shaker incubator for another 3 to 4 h. After the absorbance at 550 nm (A_{550}) of the *E. coli*
culture reached 0.45, cells were harvested at 4°C in two sterile Corex glass tubes by centrifugation using a Sorvall RT6000 centrifuge (1,800 xg X 10 min) with the brake off. Cell pellets were carefully resuspended in 12.5 ml of 50 mM ice-cold calcium chloride solution. The two Corex tubes were then placed on ice for 30 min. At this time, they were centrifuged at 4°C for 10 min at 2,000 xg. The E. coli cell pellets were then resuspended in 1 ml of 50 mM calcium chloride solution. This suspension of competent cells was then ready for use in transformation experiments.

Competent E. coli DH5αF' cells used for transfection experiment with M13mp18 recombinants were prepared using the same procedure as described above.

Transformation

A. CaCl₂ transformation method

The ligation mixture (about 0.8 μg of DNA total) and 100 μl of CaCl₂-treated competent E. coli cells were added to the bottom of a sterile 12 x 75 mm glass tube. The tube was incubated on ice for 15 min. Next, the tube was placed in a 42°C water bath for 1 min. After a heat shock treatment, the tube was placed back on ice. One ml of YT broth was added and the cell suspension was incubated at 37°C for 1 h to allow the E. coli cells to express plasmid-based phenotypes.
YT plates with the proper antibiotics were used to select for transformants. *E. coli* cells (200 μl each plate) were spread on these YT plates with a hockey stick. Inoculated plates were allowed to dry briefly, inverted and then incubated at 37°C overnight.

**B. Preparation of competent cells and transformation using RbCl method (Kushner, 1978)**

RbCl transformation method was used for transforming the *E. coli* recipient cells with the plasmid pKT231 and its recombinants. A strain of *E. coli* DH5α was cultured overnight in 5 ml LB broth. Ten ml of LB broth was inoculated with 200 μl of the overnight culture. The culture was incubated at 37°C until its A₅₅₀ reached 0.45. *E. coli* cells were harvested in a 15 ml Corex tube for 10 min at 8,000 xg in a SA600 rotor. The pellet was gently resuspended in 1 ml of solution A (containing 1.0 ml 0.1 M RbCl, 8.9 ml distilled water and 0.1 ml MOPS, pH 7.0). The suspension was centrifuged at 8,000 rpm for 10 min in a SA600 rotor. The pellet was resuspended in 1 ml of solution B (containing 1.0 ml 0.5 M CaCl₂, 1.0 ml 0.1 M RbCl, 7.0 ml ddH₂O and 1.0 ml 1 M MOPS, pH 6.5). This cell suspension was then incubated on ice for 30 min. After incubation, the *E. coli* cells were collected by centrifugation at 8,000 xg for 10 min in a SA600 rotor. The supernatant was poured off and the Corex tube was
briefly placed upside down on a layer of Kimwipes™ to allow any residual fluid to drain. The cell pellet was then gently resuspended in 0.2 ml of solution B. Next, 3 μl of dimethylsulfoxide (DMSO) and 0.2 μg of plasmid DNA or ligation mix were added. The E coli cells were then incubated on ice for 30 min. In order to improve the transformation efficiency, the Corex tube was placed in a 43.5°C water bath for one min (a heat shock treatment to increase the uptake of the plasmid DNA by the E. coli) and then placed back on ice again. The volume of the cell suspension was brought up to 5 ml using Z-broth (containing 1.6 g nutrient broth, 1.0 g Bacto peptone and 0.2 g glucose per 100 ml). After incubation at 37°C for 1 h, the E. coli cell suspension (200 μl per plate) was spread on LB antibiotic plates with a hockey stick. LB plates were inverted and incubated at 37°C overnight.

Toothpick assay and the verification of insertional inactivation

To verify the insertional inactivation of selective markers (antibiotic resistance), E. coli transformants that grew on the initial selective plates were patched on an additional set of selective agar plates, one containing the initial selective antibiotic. This was carried out using tooth picks and the plates were incubated at 37°C overnight.
As an example, when pBR325 was chosen as the vector and the insertion was into the EcoRI site, the initial plate was tetracycline and the second was tetracycline plus chloramphenicol. Transformants that grew on agar plates without the second selective antibiotic but not those containing both antibiotics were again transferred to master plates (which also contained initial antibiotic) for further selection purposes. These master plates were incubated at 37°C overnight and bacterial colonies obtained were tentatively assumed to carry recombinant plasmids. These colonies were assigned individual numbers and screened to identify the desired recombinant subclone carrying the DNA fragment (e.g. the xylRS encoding fragment).

**Alkaline lysis plasmid isolatin method**

In order to identify a recombinant E. coli subclone harboring the xylRS region, the alkaline lysis method was used to isolate recombinant plasmids for physical characterization. Alkaline lysis rapid plasmid preparations were prepared using the method described by Maniatis et al. (1982).

Culture tubes containing 5 ml of LB broth were inoculated with single bacterial colonies from the master plates using a sterile inoculation loop. The cultures were incubated at 37°C overnight on a New Brunswick G25 shaker.
incubator (250 rpm). Cells from 3.0 ml of the overnight culture were collected by centrifugation (two centrifugations into the same 1.5 ml microcentrifuge tubes). The pellets were resuspended in 100 µl of solution A (containing 50.0 mM glucose, 10.0 mM EDTA, 25.0 mM Tris-Cl, pH 8.0 and 6.0 mg/ml lysozyme). This suspension was stored on ice for 5 min. Bacterial cells were lysed by treating with 200 µl of solution B (containing 1.0% SDS and 0.2 N NaOH). The contents of the tubes were mixed by inverting each microcentrifuge tube two to three times. A 150 µl volume of ice-cold potassium acetate solution (pH 4.8) was added and the contents mixed by again inverting the microcentrifuge tubes 2 to 3 times. The resulting mixture was centrifuged for 5 min at 10,000 xg in a microfuge centrifuge at 4°C. The supernatant was transferred to a new microcentrifuge tube. A 450 µl volume of phenol:chloroform solution was added and the contents vortexed. The mixture was centrifuged at 10,000 xg for 2 min and the supernatant transferred to a new microcentrifuge tube. Then, 750 µl of 95% ethanol was added to each tube and the samples vortexed. These were then centrifuged at 10,000 xg for 10 min. After removing supernatants with a drawn out Pasteur pipette, 1 ml of 70% ethanol was added to each sample and the contents briefly vortexed. These suspensions were then centrifuged at 10,000 xg for 10 min. Supernatants were removed and pellets dried
in a Speed Vac™ vacuum concentrator. Dried pellets were each redissolved in TE. A 30 µl volume of TE (pH 8.0) containing 25 µg/ml heat-treated RNase was added to each tube. RNase-treated samples were briefly vortexed. Agarose gel electrophoretic analysis of these DNA samples was then performed. Gel loading buffer was added to each of these quick plasmid preparations (5-10 µl) and the samples were loaded into the sample wells of the gel. Undigested E. coli plasmid pBR325 as well as λ HindIII fragments, were used as molecular weight markers for the identification of recombinant plasmids. Samples were electrophoresed at 80 volts for 1 h. Following electrophoresis, the agarose gel was stained with EtBr and a photograph of the gel taken. Unused portions of the plasmid preparations were stored at -20°C.

**Physical characterization of recombinant clones**

Subclones carrying sizable inserted DNA fragments are readily identified by their relative migration as shown by agarose gel electrophoretic analysis.

Unused portions of the samples from analytic scale plasmid preparations were digested with the same restriction endonuclease used for the cloning experiment (EcoRI for example). Loading buffer was then added to these digested
samples and agarose gel electrophoresis (1% agarose gels for digested samples) was performed. Linearized vector DNA, λ HindIII fragments and cleaved target DNA fragments were used as the molecular weight markers for the identification of the desired subclone. The agarose gel was stained with EtBr and then photographed. The desired subclone was identified and set aside for further study.

**Screening for restriction endonuclease cleavage sites**

A small portion of the recovered DNA sample was used for the purpose of screening the fragment for restriction endonuclease cleavage sites. Four microliters of the electroeluted DNA sample was diluted 1 to 4 (4 μl DNA and 16 μl of TE buffer) and 2.5 μl each of this diluted DNA was used for each restriction reaction. This represented approximately 1 μg of DNA for each digestion.

**DNA analysis by acrylamide gel electrophoresis**

A 6% polyacrylamide gel was made as follows: 12.0 ml 40:1 acrylamide: Bis, 8.0 ml 10X TBE buffer, 60.0 ml distilled water and 0.1 g ammonium persulfate. A 5X polyacrylamide loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 0.5% SDS and 25% glycerol) was prepared. For each 10 μl of sample, 2.5 μl of 5X loading buffer was added and the mixture loaded onto the gel. Both HindII and HaeIII
digestions of pBR322 were used as DNA molecular weight markers. DNA samples were loaded and the electrophoresis was performed at 250 volts for about 3.5 h or until the bromophenol blue migrated approximately 80% of the gel length. The polyacrylamide gel was then stained with 0.05 μg/ml ethidium bromide and photographed.

**Cleavage of electroeluted DNA by restriction endonucleases**

The electroeluted DNA (4 μl) was restricted with HaeIII, TaqI and Sau3AI, individually. Restriction reactions were performed as suggested by the suppliers.

**Subcloning DNA fragments into M13mp18**

Digested phage DNA's were treated with the bacterial alkaline phosphatase and extracted with phenol.

For the ligation reaction, a 5X ligation buffer (25% PEG-8,000, 50 mM MgCl₂, 5 mM ATP, 5 mM DTT and 250 mM Tris-HCl, pH 7.6) was prepared and added (20 μl total volume). Ten units (1 μl) of T4 DNA ligase was added to the ligation mix in order to initiate the reaction.

Five hundred microliters of an overnight *E. coli* DH5αF' culture was transferred to 50 ml fresh YT broth and the diluted culture was incubated at 37°C on a New Brunswick G25
shaker/incubator at 250 rpm for 3 to 4 h. After the absorbance at 550 nm ($A_{550}$) of the bacterial culture reached 0.45, *E. coli* cells were harvested by centrifugation. *E. coli* cells were harvested at 4°C in two sterile Corex glass tube by centrifugation using a Sorvall RT6000 centrifuge (1,8000 xg x 10 min) with brake off. Cell pellets were carefully resuspended in 12.5 ml 50 mM ice-cold calcium chloride solution. These two Corex™ tubes were incubated on ice for 30 min and centrifuged at 4°C for another 10 min at 2,000 xg. Cell pellets were then gently resuspended in 1 ml 50 mM calcium chloride solution. This suspension of competent cells was ready for use in transfection experiments.

Transfections were performed in 3 ml glass tubes. One hundred and fifty microliters of competent *E. coli* DH5αF' cells and approximately 50 ng of recombinant M13 RF (10 µl of the ligation mixture) were added to the bottom of the glass tube. Top agar (0.6%) was prepared in advance. Each 5 ml screw capped tube was filled with 3 ml of melted liquid agar and then autoclaved. For each transfection experiment, 200 µl *E. coli* DH5αF' lawn cells, 50 µl 2% X-gal and 10 µl 100 mM IPTG were added. Screw capped tubes containing 3 ml of solidified top agar were melted in a boiling water and transferred to a 42°C water bath. Both X-gal and IPTG were
added to these screw cap tubes before adding the ligation mixture and *E. coli* lawn cells were added last. The screw cap tube was then inverted and the entire contents poured onto a YT agar plate. The plate was rotated so as to get even coverage by the top agar on the whole plate. After the top agar solidified, YT plates were inverted and then incubated at 37°C overnight. In general, recombinant M13 phage gave rise to colorless plaques on *E. coli* lawns after 6 to 8 h.

**Preparation of single-stranded templates for DNA sequencing**

*E. coli* DH5αF' was used as the bacterial host for the propagation of M13mp18 derivatives. A 5 ml YT broth tube was inoculated with 200 μl of a overnight DH5αF' culture and a single colorless plaque (recombinant M13) was transferred from a YT agar plate to this diluted *E. coli* culture with a sterile Pasteur pipette. After these inoculations of recombinant phages, *E. coli* cultures were incubated at 37°C in a New Brunswick G25 shaker incubator (250 rpm) for 5 h. Upon infection of *E. coli* host DH5αF', the plus strand of the RF was packaged and secreted as a single-stranded DNA phage. Secreted recombinant phage were harvested with 5 ml
polypropylene tubes (48 samples at a time for example). For every recombinant phage, 3.6 ml lysate was centrifuged at 10,000 xg for 8 min. Supernatants were transferred to new polypropylene tubes and centrifuged again at 10,000 xg for 8 min. Supernatants were transferred to new polypropylene tubes. Nine hundred microliters of 20% PEG/2.5 M NaCl solution was added to each tube and tubes were vortexed and incubated at room temperature for 15 min. In order to pellet the recombinant phage, these tubes were centrifuged at 10,000 xg for 10 min and supernatants were removed. Pellets were resuspended in 100 µl each of TES buffer (containing 10 mM NaCl, 0.1 mM Na₂EDTA and 20 mM Tris-HCl, pH 7.5). Suspensions were vortexed for 2 to 3 min. Phenol extraction and ethanol precipitation followed. For each sample, 50 µl of phenol was added. The mixture was vortexed and centrifuged for 2 min at 10,000 xg. The lower phase was removed and the aqueous layer extracted with 50 µl of chloroform. Samples were vortexed briefly and centrifuged for 2 min at 10,000 x g. Next, 80 µl of the upper phase was carefully transferred to a new tube. Nine microliters of 3 M sodium acetate was added to each tube and samples were vortexed for 30 seconds. 200 µl of 95% EtOH was added. Suspensions were vortexed again for 30 seconds. Samples were placed in dry ice/ethanol bath for 10 min and centrifuged for 10 min at 10,000 xg. Supernatants were removed. Each pellet was washed with 1 ml 80% ethanol,
frozen at -70°C for 15 min, and centrifuged at 10,000 xg for 5 min. The pellets were then dried using a SpeedVac vacuum concentrator and resuspended in 30 μl TE buffer. Dissolved samples were vortexed for 1 to 2 min and incubated at 65°C for 10 min. Phage DNA preparations were again vortexed for 2 to 3 min and centrifuged briefly before placing at -20°C.

In order to verify the presence and quantity of single-stranded DNA templates, agarose gel electrophoresis was performed. For each recombinant M13 template, 3 μl volume of phage DNA, 2 μl of a 5X loading buffer and 5 μl of ddH₂O were mixed and loaded into a sample well on a 17-well agarose gel (1% in this case). The electrophoresis was carried out at 80 volts for 60 min. The agarose gel was stained in EtBr solution (50 μg/liter) for 15 min and destained in deionized water for 20 min. The gel was photographed and each sample was assigned an individual clone number.

Dideoxyribonucleotide DNA sequencing reactions using single-stranded templates

Nucleotide sequences of these M13 derivatives were determined by the method of Sanger et al. (1977). Both strands of the desired DNA fragment were sequenced by the dideoxyribonucleotide chain-termination method using
Sequenase Version 2.0 DNA sequencing kits (United States Biochemical). Sequencing primers were either the pUC18/19 forward primer (the common primer), or synthetic oligonucleotide primers derived from previously determined pDK1 xyIR sequences. Both single and double-stranded DNA templates were used for the sequencing reactions. M13mp18 and its recombinants were propagated in E. coli DH5αF′.

DNA sequencing reactions (in sets of 12 M13mp18/19 derived samples) were performed as follows: The annealing mixture was prepared using 13.0 μl primer, 26.0 μl sequencing buffer and 26.0 μl sterile water (65 μl total). Five μl of this diluted primer was then distributed among twelve 500 μl microcentrifuge tubes, each containing 5 μl of single-stranded DNA template. These samples containing primer/template annealing mixtures were placed at 65°C in a water bath for 2 min and then allowed to cool slowly to less than 35°C. This was accomplished by transferring them to a glass tray filled with water removed from the 65°C water bath which was allowed to cool at room temperature on the bench. This step routinely took at least 30 min.

While waiting for annealing mix to cool, the termination tubes were labeled and filled. There were four tubes G, A, T and C for each DNA fragment to be sequenced. For running twelve sets of sequencing reactions, forty-eight 500 μl microfuge tubes were needed. One of four
dideoxyribonucleotides (ddGTP, ddATP, ddTTP and ddCTP) was distributed into each of the four sets of twelve 500 µl microcentrifuge tubes. 2.5 µl of dideoxyribonucleotides was used for every reaction.

The labeling solution was prepared by mixing 13.0 µl DTT, 20.8 µl ddH₂O, 5.2 µl labeling mix and approximately 65 µCi (6.5 µl x 10 µCi/µl) ^35^S-dATP (total volume of 45.5 µl). Three and one half microliters of the labeling solution was transferred to each of the twelve reaction tubes containing the primer/template mix.

Sequenase (3.3 µl) was diluted eight-fold with 23.1 µl of the Sequenase dilution buffer. Two microliters (out of a total of 26.4 µl) of the diluted sequenase was added into each reaction tube. Samples were briefly vortexed and incubated for 4 min at room temperature. For each template utilized, the final volume of the reaction mixture was 15.5 µl.

Termination tubes, containing dideoxyribonucleotides (ddGTP, ddATP, ddTTP and ddCTP), were incubated at 37°C. 3.5 µl of the labeling reaction mixture, which started from using 5 µl recombinant template, was then distributed to all four termination tubes (G, A, T and C). These samples were mixed using a pipetor and incubated at 37°C for another 5 min.

Four microliters of stop solution was added to each of the forty-eight reaction samples. All forty-eight samples
were briefly centrifuged for 10 seconds in order to collect these samples down to the bottom of microcentrifuge tubes. In order to avoid confusion when loading samples onto sequencing gels, all twelve sets of reaction samples were placed back to acrylic racks according to individual number assigned to every set of samples following a G, A, T and C order. A sheet of Saran Wrap was covered and samples were kept at -20°C until use.

**DNA sequencing gel electrophoresis**

Sequencing gel electrophoresis was performed at 1,700 V using 6% denaturing DNA sequencing gels. These sequencing gels were pre-electrophoresed for 1 h before loading samples. Wells were then rinsed with buffer to remove unwanted urea which had diffused from the gel. When the pre-electrophoresis was nearly complete, sequencing samples were incubated in a 90°C water bath for 2 min. They were then immediately transferred to racks in a 0°C ice-slurry. Samples were loaded onto the gel using a Hamilton syringe. The Hamilton syringe was rinsed several times with deionized water between each sample loading. The samples of reaction mixture were loaded following the order of G, A, T and C. One lane was skipped before loading the subsequent set of four samples. 3.5 µl of the reaction mix was loaded in each gel lane.
Autoradiography of DNA sequencing gels containing $^{35}$S-labelled nucleotides

Sequencing gels were dried onto Whatman 3 MM paper and exposed to Kodak XAR-5 X-ray film at room temperature. The dried gel was placed in a film cassette and a piece of X-ray film was placed on top of the dried gel in a dark room. The duration of the exposure period varied depending upon the relative amount of the $^{35}$S-dATP label incorporated into the DNA samples. Typically, a two-day exposure was enough to obtain a good autoradiogram. Films were developed in a dark room using both the developer and the fixer solutions indicated by the manufacturer (two min each). Developed films were then rinsed with tap water and air dried for 30 min. Autoradiograms were assigned individual reference numbers and all sample lanes appropriately labelled. Nucleotide sequences shown on these autoradiograms were read and any M13 sequences identified (using the DNA Strider software program) and deleted.

DNA sequencing reactions using double-stranded plasmid DNA templates

Several procedures have been developed in the past several years in order to sequence double-stranded plasmid DNA templates directly. For this experiment, the chain
termination sequencing method for double-stranded DNA templates as described by Chen (Chen et al., 1985) was followed. Alkaline-denatured plasmid DNA templates and $^{35}$S-dATP were utilized for sequencing reactions. Sequenase Version 2.0 DNA sequencing kits were purchased from United States Biochemical and used for these sequencing reactions. The quality of autoradiograms produced using this method was approximately similar to that obtained using single-stranded templates. In case of need, phenol extraction and/or PEG precipitation of template samples was repeated to clean up the supercoiled plasmid DNA. The oligonucleotide primers were prepared by dissolving the commercially synthesized sample in 100 μl ddH$_2$O. In order to obtain a concentration of around 10 pmoles/μl, 10 μl of the dissolved oligonucleotide sample was diluted fifty-fold (490 μl of ddH$_2$O was added). The actual DNA concentrations of these primer samples were then determined from their $A_{260}$ using a UV spectrophotometer. Fifty microliters of the diluted sample was utilized (concentration = $A_{260}$/0.01 x N, N is the number of nucleotides in the primer) for this purpose. From 0.5 to 1 pmole of oligonucleotide primer was used per sequencing reaction.

Safety precautions
Radioactive waste was discarded properly according to radiation safety rules. Several precautions were taken for safety reasons when hazardous materials or high voltage power were used. For example, a correct sequence to turn off a power supply after completing the electrophoresis of a DNA sequencing gel was as follows: the power supply was turned off first and then the leads disconnected at the power supply. Finally the leads connected to the gel stand were removed. Many laboratory related accidents can be avoided if the necessary precautions are taken and safety rules are strictly enforced.
CHAPTER III

RESULTS

For this report I have divided the results into three major areas. The first is the subcloning of the desired xylRS DNA fragment from the P. putida TOL plasmid pDKR1. The second section will cover the determination of its nucleotide sequence using the dideoxynucleotid method of Sanger (Sanger et al., 1977). In the final section, computer analyses of both the DNA and the predicted amino acid sequences encoded within the xylRS region will be described.

Subcloning of the xylRS DNA fragment from the TOL plasmid pDKR1

The xylRS regulatory region from the 40 kb TOL region of the cointegrate plasmid pDKR1 (RP1::TOL) was subcloned into the EcoRI site of the E. coli plasmid cloning vector pBR325 (Ap\(^R\), Tc\(^R\), Cm\(^R\)).

A. Construction of the recombinant plasmid pBK990

The procedure for the construction of the recombinant plasmid pBK990 is shown on Fig. 11.
Figure 11. Construction of the recombinant plasmid pBK990
pDK1 117 kb

RP4 56 kb

pDKR1 95 kb

pBR325 6 kb

PBR990 12.5 kb
EcoRI digested plasmid vector pBR325 (0.1 mg) was prepared as described in Materials and Methods and treated with bacterial alkaline phosphatase. Digested samples were extracted with phenol and ethanol precipitated. The final DNA pellet was dissolved in 25 µl of TE buffer and stored at -20°C.

The target DNA (pDKR1) was cleaved with the restriction endonuclease EcoRI. In order to verify that the digestion reaction was complete, 1% agarose mini-gel electrophoresis was performed. A 6.5 kb EcoRI fragment (fragment D, the 4th largest fragment of pDKR1, see Shaw and Williams, 1988) covers the entire xylRS region and was the objective of this subcloning effort.

Linearized pBR325 DNA (0.6 µg) was mixed with pDKR1 EcoRI fragments (0.2 µg target DNA in TE, pH 8.0). Ligation reactions were carried out in a total volume of 20 µl each.

Competent E. coli DH5α cells were prepared and used as the bacterial recipients for transformation. The procedures for the preparation of competent E. coli cells were described in chapter II.

The ligation mixture (about 0.8 µg of DNA total) was used to transform CaCl2-treated competent E. coli DH5α cells.

Insertional inactivation of the CAT (chloramphenicol
acetyl transferase) gene was verified by patching the original ampicillin<sup>R</sup>/tetracycline<sup>R</sup> colonies on plates containing ampicillin/tetracycline plus chloramphenicol. Transformants that grew on agar plates without chloramphenicol but not those containing all 3 antibiotics were again transferred to master plates. Approximately 175 recombinant subclones were obtained. These colonies were assigned individual clone numbers and further screened to identify the desired recombinant subclone carrying specific TOL DNA fragments (e.g. the xylRS encoding fragment).

**B. Physical characterization of recombinant clones**

In order to identify the recombinant *E. coli* subclone carrying the desired xylRS recombinant plasmid (subsequently designed pBK990), the alkaline lysis method was utilized to isolate the DNA necessary for physical analysis. These alkaline lysis rapid plasmid preparations were prepared using the method described by Maniatis (Maniatis et al., 1982).

Recombinant pBR325 subclones were initially characterized by their relative migrations during agarose gel electrophoresis as shown by the photograph in Fig 12. In order to identify specific recombinant plasmids carried the xylRS region of pDK1, samples from these alkaline lysis preparations were digested with EcoRI and XhoI. Agarose gel electrophoresis was again performed. EcoRI and XhoI double
Figure 12. Electrophoretic analysis of restriction endonuclease products from the recombinant plasmid pBK990.

1. λ HindIII fragments. 2. Double digestion of pBK990 using EcoRI and BamHI. The sizes of these fragments show that EcoRI cut the 6.5 kb BamHI fragment into two smaller ones (4.4 kb and 2.1 kb). 3. Digestion of pBK990 with BamHI. Three BamHI fragments (6.5 kb, 3.9 kb and 2 kb) were shown on this lane.
digestion fragments of pDKRl, EcoRI linearized pBR325 DNA and λ HindIII fragments were used as the molecular weight standards. The agarose gel was stained with EtBr and an example is shown in Fig. 13. A pBR325 recombinant subclone having a 6.5 kb insert into the vector's EcoRI site was identified and designated pBK990.

The orientation of the 6.5 kb EcoRI fragment within the pBR325 was determined from further restriction endonuclease digestion analysis of the subclone. Using this data and the known restriction endonuclease cleavage maps of pBR325 (Prentki et al., 1981) and the pDK1 xylRS (Shaw et al., 1988) region, it was possible to determine the insertional orientation of the fragment. As an example, the sizes of three BamHI fragments obtained were consistent with only one of the two possible orientations. The correct orientation of the xylRS fragment within the subclone pBK990 is shown in Fig. 14.

The E. coli DH5α strain carrying the recombinant plasmid pBK990 was maintained on selective medium (ampicillin and tetracycline at concentrations of 100 and 25 μg/ml, respectively).

C. Preparative scale isolation of the recombinant plasmid pBK990
Figure 13. Identification of the recombinant subclone pBK990. EcoRI and XhoI double digestion of eight recombinant subclones (clones 300, 63, 115, 116, 146, 202, 203 and 206).

Lanes 5 and 10 are λ HindIII fragments. Plasmid isolated from clone 203 shown on lane 8 was identified as carrying the xylRS region and designated pBK990.
Figure 14. Restriction endonuclease cleavage map showing the orientation of the *xylR* and *xylS* genes within the 12.5 kb recombinant plasmid pBK990.
After the identification of the recombinant pBK990 subclone, a large scale plasmid preparation was carried out using the Tanaka and Weisblum method as described in Materials and Methods. The purified plasmid DNA was redissolved in 0.5 ml TE buffer and the DNA concentration of the plasmid preparation (from 1 liter _E. coli_ culture) was calculated from its $A_{260/280}$. Normal plasmid yields were 0.75-1.5 mg/liter of bacterial culture.

D. Purification of the cloned _xylRS_ fragment

The 3.9 kb _BamHI_ fragment which contained both the _xylR_ and the _xylS_ genes was part of a larger, 6.5 kb _EcoRI_ fragment (see restriction endonuclease cleavage map, Fig 2). Approximately 500 μg pBK990 plasmid DNA was cleaved with 50-200 units of _BamHI_. A 1% agarose mini-gel electrophoretic analysis was performed to verify the completion of the restriction reaction. Three DNA fragments of 6.5 kb, 3.9 kb and 2.0 kb appear in the photograph of the gel (Fig. 12). The 3.9 kb _BamHI_ fragment encoded the _xylRS_ region of _pDK1_. In order to separate this fragment from the other two DNA fragments, vertical preparative scale agarose gel electrophoresis was performed. A 20 x 20 x 1.5 mm 1% vertical agarose gel was prepared. 500 μg of _BamHI_ digested recombinant plasmid (pBK990) DNA was loaded onto a 6 cm wide sample well. Agarose gel electrophoresis was performed at 80
volts for 6 h with constant buffer circulation. The power supply was then disconnected. The glass cassette was removed and the agarose gel was illuminated under short wave length UV light in a dark room. The 3.9 kb fragment was identified and carefully cut out from the agarose gel using a razor blade.

The desired 3.9 kb BamHI DNA fragment was recovered from the agarose by electroelution as described in Methods. Phenol extractions and ethanol precipitations were carried out as described in chapter II. The final DNA pellet was dissolved in 25 µl TE. The amount of DNA recovered was estimated to be approximately 80 µg. Purity was verified by gel electrophoretic analysis (Fig. 15).

E. Screening for restriction endonuclease cleavage sites

A small portion of the recovered 3.9 kb DNA sample was used for the purpose of screening the fragment for restriction endonuclease cleavage sites (Fig. 16). Nine restriction endonucleases (HaeIII, AluI, Rsal, BstUI, Sau3AI, TaqI, HinfI, HincII and MspI) were chosen. 5µl of the electroeluted DNA sample, the cloned 3.9 kb BamHI fragment, were diluted 1 to 4 (5 µl DNA and 20 µl of TE buffer) and 2.5 µl each of this diluted DNA was used for each restriction reaction. This represented approximately 1 µg of DNA for each digestion.
Figure 15. Agarose gel analysis of the 3.9 kb BamHI fragment encoding xylRS. 1. The electroeluted DNA sample showing the 3.9 kb BamHI fragment. 2. λ HindIII fragments. 3. The electroeluted DNA sample showing the 6.5 kb BamHI fragment.
Figure 16. Screening of the 3.9 kb BamHI fragment for restriction endonuclease cleavage sites (HaeIII, AluI, Rsal, BstUI, Sau3AI, TaqI, HinfI, HincII and MspI). Lane 3 is a HinfI digestion of pBR322 and lane 8 is a HaeIII digestion of pBR322.
A 6% polyacrylamide gel was made as described in Materials and Methods. The restriction endonuclease cleaved samples were then loaded onto the gel. Both pBR322 *Hinfl* and *HaeIII* fragments were used as DNA molecular weight standards. Following electrophoresis, the polyacrylamide gel was then stained with 0.05 µg/ml ethidium bromide and photographed. DNA sequencing strategies using *HaeIII*, *Sau3AII* and *TaqI* for the 3.9 kb *xylRS* *BamHI* fragment were then determined based upon the sizes and number of fragments generated by each enzyme as shown in Fig. 17.

F. Subcloning TOL fragments into M13mp18

The electroeluted DNA (4 µl) was restricted with *HaeIII*, *Sau3AII* and *TaqI*, individually. These digested DNA samples were phenol extracted and redissolved as described in Materials and Methods and utilized for subcloning into M13 DNA sequencing vectors.

Zabarovsky and coworkers described the *XhoI* half-site arms cloning strategy in 1986. A similar cloning strategy was taken to subclone portions of the 3.9 kbp electroeluted *BamHI* fragment into M13mp18. The *Sau3AII* fragments were cloned into the *SalI* site of M13mp18. *HaeIII* fragments were inserted into the *HincII* site of the phage. An attempt was made to clone *TaqI* fragments into the *AccI* site, but no recombinant phage were obtained. Three samples of 0.2 µg
Figure 17. The DNA sequencing strategy for the *xylR* and *xylS* genes. a. DNA sequences obtained from recombinant phage carrying *HaeIII* fragments. b. DNA sequences obtained from recombinant phage carrying *Sau3AI* fragments. c. DNA sequence obtained using dITF reactions and the 15-mer oligonucleotide primer (filled arrow to the right).
each of the M13mp18 RF DNA's were restricted with SalI, HincII and TaqI, respectively. Digested phage DNA samples were treated with bacterial alkaline phosphatase, phenol extracted and redissolved in TE buffer as described in Materials and Methods. For the ligation reaction, 1 µl of T4 DNA ligase was added to each ligation mix (20 µl total volume) to initiate the reaction. Competent E. coli DH5αF' cells were utilized for transfection. In general, recombinant M13 phage gave rise to colorless plaques on the E. coli lawn after 6 to 8 h. Approximately 15 to 75 colorless plaques were obtained from each plate and these plaques were then utilized for making single-stranded DNA templates.

**Determination of xylRS nucleotide sequence using the dideoxyribonucleotide method**

**A. Preparation of recombinant single-stranded templates for DNA sequencing reactions**

M13mp18 and its recombinants were propagated in E. coli DH5αF'. A total number of 168 recombinant phage samples (colorless plaques) were prepared. In order to verify the existence of single-stranded templates after the preparation procedures, agarose gel electrophoresis was performed. For
every single-stranded template, 3 μl of recombinant DNA sample, 2 μl of a 5X loading buffer and 5 μl of ddH₂O were mixed and loaded into a well of a 17-well 1% agarose gel. The electrophoresis was carried out at 80 volts for 60 min. The agarose gel was stained in EtBr solution (50 μg/liter) for 15 min and destained in deionized water for 20 min. The gel was photographed and the single-stranded DNA samples were assigned individual clone numbers (Fig. 18).

B. DNA sequence analysis of the cloned _xylRS_ region

Nucleotide sequences of these M13 derivatives were determined by the method of Sanger (Sanger et al., 1977). Both strands of the 3.9 kb BamHI fragment (_xylRS_ region) were sequenced by the dideoxyribonucleotide chain-termination method using Sequenase version 2.0 DNA sequencing kits (Fig. 19). Sequencing primers were either the T7 promoter primer (Fig. 20), or synthetic oligonucleotides derived from previously determined pDK1 _xylRS_ sequences. Both single- and double-stranded DNA's were used as templates for sequencing.

Three of the regions within _xylRS_ were found to exhibit substantial compression artifacts after sequences from these autoradiograms were read. Six recombinant templates were therefore chosen for performing dITP reactions. Results of these were shown in Figure 21. For every template chosen, both dGTP reactions and dITP reactions were loaded on a
Figure 18. Single-stranded templates prepared from recombinant M13 plaques.
Figure 19. Example of an autoradiogram from dNTP DNA sequencing reactions
Figure 20. Examples of M13mp18/19 DNA sequencing primers. This figure was adapted from the New England Biolabs 1992 Catalog, page 147.
Figure 21. Autoradiogram of a DNA sequencing gel showing the radioactive products resulting from two sets of dITP sequencing reactions.
sequencing gel adjacent to each other. These two sets of sequences were then compared. The nucleotide sequences from all three compression areas were successfully read and corrected according to the results derived from these dITP reactions.

Dideoxy sequencing was carried out using synthetic oligonucleotide primers complementary to xylRS sequences in order to obtain additional information on regions flanking xylRS sequences.

For this experiment, a 15-mer oligonucleotide was purchased from Biosynthesis Inc. (Lewisville, Texas). The synthetic oligonucleotide primer was designed based upon the previously obtained sequence of the pDK1 xylS C-terminal region. As a result of using this synthetic oligonucleotide primer, the xylS C-terminal sequence was extended downstream into another unknown ORF (Fig. 22).

Some DNA sequence analysis was also performed using double-stranded templates. A preparative scale isolation of the plasmid pBK990 was carried out and its DNA concentration calculated from its A_{260/280}. The procedures for preparative scale plasmid isolation from E. coli have been described in Chapter II. A small portion (0.1 mg) of this plasmid pBK990 DNA sample was extracted with phenol. Next, Chen's method (Chen et al., 1985) for sequencing double-stranded templates
Figure 22. The restriction endonuclease cleavage map of xylS (2209 to 3923).
383 Bgl I
371 Sst I
366 EcoNI
293 Bgl II
534 Bsa I
534 Bal I
350 Rsa I
1054 Xho I
1496 Bam I
1443 Nhe I

xyLR S # 2209-3923  xyLS

xyLR S #2470  xyLR S #3465

Unique Sites
was followed. United States Biochemical Sequenase Version 2.0 sequencing kits were also used for these sequencing reactions. Typical results from these reactions are shown in Fig. 23.

Several phage subclones carrying large segments of TOL DNA were selected for running "stretched" gels. Sequencing reactions were performed as described above. The gel was run at 1,700 volts for up to 10 h. The xylene cyanol marker dye was loaded two to three times until it had migrated up to 100 cm. Under these conditions sequences up to 500 bp from the primer were successfully determined.

Two approaches were taken in order to read sequences close to the primer. One can run the xylene cyanol marker no further than 15 cm. The primary difficulty encountered when this was carried out using standard reactions was that bands near the primer were very faint or completely absent (due to insufficient numbers of terminations within this region). The use of Mn buffer greatly increases termination in the early region and was therefore used in these situations.

C. DNA sequence of the \textit{xylRS} region

The DNA sequence of the 3.9 kb \textit{BamHI} fragment was completely determined (Fig. 24) for both strands with all segment junctions overlapped by additional sequence. Open reading frame analysis of this sequence identified two major
Figure 23. Autoradiogram showing nucleotide sequence analysis using double-stranded templates and oligonucleotide primers.
Figure 24. The xylRS region nucleotide sequence of TOL plasmid pDK1. xylR is 1,701 bp in size (2024 to 324) and xylS is 996 bp in size (2470 to 3465). These two regulatory genes are divergent and two promoters (2025 to 2469) are on top of each other.
GATCCAGCAG GTGCCTGTCG CCGACGTGGC TAGACCGGTT CGTGAATCAA CCGACAGGGC 60
TCTAGACAGC GCCACACCGG GGCACCTCGG TGGCCTGGGG TTTGACCCCG ATCCAGCAGG 120
GTATTTTCAG CTACGGGGCG AGAGSCGACA AGCTGTGAGC CGTGGCAAGA AAGCCAGAGA 180
TCAATTAGAT GCACGCGGAC AGCAAGTGGT GGGGTGGGAT TGTGAGCCCG ACTGCCAATT 240
TACCGACTCC CCGCGCTTGC AAATTGTGCT TTCCCCAGAA TCCGGAAGGG CCCTTGAGAAA 300
GGCCCAAAACC TGGCGCAAGG CAGCTATGGG CCGGTGCTTT TCAGAGATGAG GCTGGGGGTC 360
AGCTTCTTTAA GTCACTAGGC CATGCGCGGG CAGGTCAATTC CCAGCAACCG CGCCAGCTGT 420
GAGATAATTCT GCCACAAACG ATCCATGCGG GTGCACGTGTA AACCCGCTCG GAGATCTCTG 480
AGACTGACGC CCGGGTGCGAT AATTGCTCTA AACAACTCCT GACGGGACTT CTCACTCAAAC 540
CGGGCCCTGC TTGATAGCGT GTCGCGCTTC GAGGTATTCC CAGCTTACAG AAGACGCAGAC 600
TCGACATTGA TGCTTTCCGG TCCGCTGCGG AGAATCAGCC CGCGCTCAGG GGGTTCTCC 660
AGCTCGGCGA TATTGCCTGG CCAATGGTAG TGAGGCGAGG CCCTCACTGG TGCACTAGAC 720
AGGCAACAGG TCTCCTGACT ATTCCCTGTA TGGTCGCTCT CAGAAAGATG TCGAACAGAC 780
AGGCGGGATGT CTCCAGCGGC TCGCCGAACC GCGGGCATCT GCAGGGGAAG AAGACCTGTCG 840
CGAAAAAACG GGCTGCGGCG AAGCGCCCGCC ATCTGTAGCG CCTCTCCAGT GCTCTCTGT 900
GGTGCCTGCGT TAAACCTCAC GTCGACCTTT CGGCTGCGGT CCGCTCCGAC CSGCTCTGACG 960
TCTCTTCCTT CCAATCGCGG TAGCAGGGGT GCCTAGCGCC TCGGGCTCAA TCGATCACC 1020
TCATCAAGAA AGATGGTGCC GCCGGTGGCC CGTTCAAAC GTCCAGCGCG TGCATTGACC 1080
GCGCCCGTAT AGGGCGCCCTT ATCGACACCA AACAGTTCCG ACTCGATCAG ATCCGGCGGA 1140
ATTGCCGCAC AGTTCACCSC GACGAAACGGT TGCTCTGCCA GCTCACTGGG CAAATGCACG 1200
CTGCGCGGGA TTACCTCCTT GCCCACCACA GTCTACACCA GTAGCAGGAC CGAAACCTTG 1260
CCACGTGCAG CCTTGTGCAG GTGTCGCCAG ATGGCGTTGTC ATGGCTGGCGA ATGGCGGAATG 1320
CCGAATAGCT GCCCCATGTA CTGCTTCAGG CGGTTGGCGCA GGTTGGCGAC CTGGGTCTGC 1380
AGCTCGTAGC GCTCAGCTAC GATCCGGTGC CTTTTGAAGGT AGGCTTCGGAA ACTGCTGACA 1440
TCGCCCCACT CTCTCCGCCG CTTGGCGAGG ATAAAACGGC TACCGCTACC GCACCCCAGG 1500
CAGCTATGTTT CTTGAAAAT GATCTCAGGG CCCATGAAAC CGAAGCTAGCA 1560
TAGGTAGGCA GCTCCCCGCA GACGCGGTCA TCTCCAGGGC CGAGCTCAGAT ACGGAGAGAA 1620
TCACCTTCAA AGGAATCAAT CCACGCCGCC TCACCGTTGA AAGTGCCGTC CCGGATGGCG 1680
ATATCCATTG TCAGCAGAGG TACTTGTACC ATCCCATGGA GCSCATACAA TTGAGCGCCA 1740
GCCAGGAACA CCTCTCTCCT GCAGTGCCGC GCCGGGTGGCT TCGCCGGTCC CTCGGCATCC 1800
ATCAGGGCGG ACTGAGRGCC CAAGCCAGGG AAGAARACCT TGCCCCGCTC GATGGCGGTC 1860
AAGCTGATAA TTTGCCGGCC GAAAGCTGGGC AGCGTGTGACATTTGCGATAG GATCGCATGC 1920
TGCTCTCCCG GCCAGATCTT GCCTTCCGGC GCAACGAAAC GGATCGTGGT GATPAGGGTT 1980
TGCGATATCA CTATGCTGCA CTGGGGATTG TATATAAGGG ACATTGGTTT CCAATTTGGG 2040
TTATCAGGAT GATCCACCC CGACGAGAATG ATGGGCTCAC TGGGTCATG ACTACCTTGC 2100
AAGCTCAAGG CCTGTCCTTG CCAGGCTCAGC ATCCCATCGG CCACCTTACC ATTTGCCTAG 2160
GTGGTCCCTGA AAAATTTACC AATTGATTTA CTGAAATCCA CCTCAACGCA GATCCCTTT
ATCCGCCAAAT TGCTGCCGGC ATGGCCCGAC AGGACGAACT TCTGCTGGTTT CCGCTTCTTA
AAAAGAACGT CTTGGCTTCTG CTGGCCCTTA TTTTTGCTTTG GAAAGATGTT CACTAATTGC
AAAAAGGATG CCSCAAGCTG GCATGCGGAG TAACCCAGTA CCGCATCAAG TCGAGATGCA
TTTTCATCGA CCGGGCGCCT TTGGACATTA G CTGGCTGCAG CCAGCAATCT TCGGATGCTT
CGAACGAGCTT GCCTATGTTT TGGATATTTG CAGAAAAGTA AGATCTTTGC CCACGGCGAG
CTGGCAATGAA TACTCAAGGG CACGCAACGG TGGCCTTGGAT ACGCATCAGT CACGATCTG
GATTCGGGAG AATTTATCGT TAAATTGCCG TCAGTGTCTC TTGATCGGGC ATGCGATGAC
AAACATTGGC ACAAGCCGAG GAGATTTCAG GCTGATGGCT TTATCAATCT ACTCGGGTTA
ATGGAAGGTA GTTTTGAGAT AAGCGGAAGT CCGCTCGCGG TGCACTGTTA GGAGCTGGCG
CTGATGAGCC CACGCGGATC AATTTATCGT CAAAGCCGAG AGGGTTCTT CCAGGCGTGA
GTGCTGCTGGT TTGGGATCGG GTTTGAGAT AAGCGGAAGT CCGCTCGCGG TGCACTGTTA
ATGGAAGGTA GTTTTGAGAT AAGCGGAAGT CCGCTCGCGG TGCACTGTTA GGAGCTGGCG
CTGATGAGCC CACGCGGATC AATTTATCGT CAAAGCCGAG AGGGTTCTT CCAGGCGTGA
AACTACATCC GCAACCACCA GCTGAGTGTC ATACGGGCCC GCTTGAGCGA TCCCAATGCA 3300
ARATGCCGTA GTGTCACCAG GATGGGCGTA GACTACGGCT TCTTTTCATA GGGACGTTTC 3360
GCCGAGAAGC ATAGGAGGAC TTTGCGGAG CTGCCCTCCG ACACCCGGCG TCAGCGCAAA 3420
ATGAAGTGCC TTAGTCCOGG GGAAGCGCTG CCCCGACCCG GCTGACTGAA CGTTACTACA 3480
GCGGAGGTTT TGAAGGCCAT GCTAAGGGCA ACAACTGGTG CTGAAAATGA TGATGGAATT 3540
GCGTCAGACG CTCCGGACCG GCAAGATTCA TCCGCGGGGT GTGGTTGTTG ATTTTAGACC 3600
ACAATCATGC CTTGCGATGG CGACCCCCGC TCCACTCTTT TTGCCGCGTT GCTAGCGGAC 3660
AGAAATTCAC TCCCGGTGGG GTGTTTGCGA AGAGCCCTGT TCGGAATGCC AATTATGAAG 3720
AATAGATCTT AGGGTTCTCG TAGGGTTTGC GACCCGGGCG TACCCAGCAC GATCGCCTGG 3780
ATCGTTTACC GCTCGGCCTC AAGGCCAGAA GGGTAACTGG ATATTGGATG CGGATATCAC 3840
GTGTTTCTTG GATCTCGCAT AACAATCGCT TTGGTTATAG AGGTGTTGTC GAAGAGGCTA 3900
GGCTGATCGT CGACCGCGCTT GAT
open reading frames of 1,701 bp and 996 bp, respectively, separated by an approximately 400 bp intergenic region (Fig. 25). These were tentatively identified as encoding xylR (Fig. 26) and xylS. As expected for Pseudomonas genes, both of these open reading frames exhibit a high frequency usage of codons ending in C or G (Fig. 27 and 29). Comparisons of putative xylR and xylS sequences to those of the pWW0 TOL plasmid confirmed that these two open reading frames were in fact the xylR and xylS genes of the TOL plasmid pDK1. xylR is 1,701 bp in size (Fig. 28) and encodes a 63.7 kDa protein. xylS has a 996 bp coding region (Fig. 30) giving rise to a 37.5 kDa protein. The xylR and xylS genes are transcriptionally divergent, running in opposite directions from the central intergenic region. My results show that the pDK1 P_R and P_S promoters are separated and found in the 400 bp intergenic region similar to that observed for pWW0. The xylR promoter P_R and the xylS promoter P_S appear to overlap. Homology analysis of the xylRS regions of pDK1 and pWW0 will be discussed below.

Computer analysis of DNA and encoded amino acid sequences

For DNA sequence analysis, a Macintosh SE/30 personal
Figure 25. The restriction endonuclease cleavage map of the xylRS region (1 to 3923). This figure shows the location of xylR (2024 to 324) and xylS (2470 to 3465) genes. The relevant restriction endonuclease sites are indicated.
Figure 26. The restriction endonuclease cleavage map of xylR (2308 to 1).
527 Pst I
904 Pvu I
513 Bgl I
15 Xmn I
370 Bgl II
750 Nhe I
15 Xmn I
1383 Sal I
1861 Nco I
2243 Xma I
1719 Xma I
1719 Xba I
1944 Hind III
2243 Xma I

xyIRS #2308 to 1

xyIRS #2024

xyIRS #324
Figure 27. Codon usage table for the xylR gene of pDK1.
<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTC</td>
<td>Phe F</td>
<td>1</td>
</tr>
<tr>
<td>TCC</td>
<td>Ser S</td>
<td>4</td>
</tr>
<tr>
<td>TAA</td>
<td>Tyr Y</td>
<td>2</td>
</tr>
<tr>
<td>TAT</td>
<td>Tyr Y</td>
<td>3</td>
</tr>
<tr>
<td>TTT</td>
<td>Phe F</td>
<td>9</td>
</tr>
<tr>
<td>TCT</td>
<td>Ser S</td>
<td>2</td>
</tr>
<tr>
<td>TAG</td>
<td>Tyr Y</td>
<td>7</td>
</tr>
<tr>
<td>TGC</td>
<td>Cys C</td>
<td>6</td>
</tr>
<tr>
<td>ATT</td>
<td>Ile I</td>
<td>7</td>
</tr>
<tr>
<td>ACT</td>
<td>Thr T</td>
<td>4</td>
</tr>
<tr>
<td>AAT</td>
<td>Asn N</td>
<td>8</td>
</tr>
<tr>
<td>ATG</td>
<td>Met M</td>
<td>17</td>
</tr>
<tr>
<td>GTT</td>
<td>Val V</td>
<td>3</td>
</tr>
<tr>
<td>GCT</td>
<td>Ala A</td>
<td>6</td>
</tr>
<tr>
<td>CTT</td>
<td>Val V</td>
<td>13</td>
</tr>
<tr>
<td>TTA</td>
<td>Leu L</td>
<td>2</td>
</tr>
<tr>
<td>TCA</td>
<td>Ser S</td>
<td>5</td>
</tr>
<tr>
<td>TAG</td>
<td>Amb Z</td>
<td>1</td>
</tr>
<tr>
<td>TGG</td>
<td>Trp W</td>
<td>6</td>
</tr>
<tr>
<td>CTP</td>
<td>Leu L</td>
<td>8</td>
</tr>
<tr>
<td>CCT</td>
<td>Pro P</td>
<td>1</td>
</tr>
<tr>
<td>CAT</td>
<td>His H</td>
<td>5</td>
</tr>
<tr>
<td>CTC</td>
<td>Leu L</td>
<td>15</td>
</tr>
<tr>
<td>CCT</td>
<td>Pro P</td>
<td>1</td>
</tr>
<tr>
<td>CAT</td>
<td>His H</td>
<td>5</td>
</tr>
<tr>
<td>CTA</td>
<td>Leu L</td>
<td>8</td>
</tr>
<tr>
<td>CCA</td>
<td>Pro P</td>
<td>2</td>
</tr>
<tr>
<td>CAA</td>
<td>Gin Q</td>
<td>5</td>
</tr>
<tr>
<td>CTG</td>
<td>Leu L</td>
<td>28</td>
</tr>
<tr>
<td>CCG</td>
<td>Pro P</td>
<td>10</td>
</tr>
<tr>
<td>CAG</td>
<td>Gin Q</td>
<td>16</td>
</tr>
<tr>
<td>ATT</td>
<td>Ile I</td>
<td>1</td>
</tr>
<tr>
<td>ACC</td>
<td>Thr T</td>
<td>9</td>
</tr>
<tr>
<td>ACA</td>
<td>Thr T</td>
<td>2</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr T</td>
<td>7</td>
</tr>
<tr>
<td>CTA</td>
<td>Ala A</td>
<td>10</td>
</tr>
<tr>
<td>GCA</td>
<td>Ala A</td>
<td>14</td>
</tr>
<tr>
<td>CCA</td>
<td>Ala A</td>
<td>12</td>
</tr>
<tr>
<td>GCA</td>
<td>Ala A</td>
<td>10</td>
</tr>
<tr>
<td>CCA</td>
<td>Ala A</td>
<td>12</td>
</tr>
</tbody>
</table>

The table above represents the codons and their corresponding amino acids, along with their positions in the sequence.
Figure 28. The nucleotide and predicted amino acid sequences of xyIR from the TOL plasmid pDK1.
xyIR

AGGCGCTGAC GTGCAAGGT AGTCAGAGGC GCAGTGAGCC TCTGATGTTT CGCCGGTGG ATCATCCCGA TAAAAACAG AGGAAAAACAA
1
ATG TCC CTT ATA TAC AAT CCC AAG ATG CAG CAT GAG GAT ATG CAA GAC CTT ATC AGC CAG ATC GTG TTC GCT GCC
met ser leu ile tyr asn pro lys met gln his glu asp met gln asp leu ile ser gln ile are phe val ala
10
GCC GAA GGC AAG ATC TGG CTG GGA GAG CAG CGC ATG CTG CTA ATG CAG CTA TCT AGC CTG GCC AGC TTC CGC CGC
ala glu gly lys ile trp leu gly glu gln arg met leu leu met gln leu ser thr leu ala ser phe arg arg
20
GAA ATT ATC AGC TTG ATC GCC ATC GAG CGG GCC AAG GGT TTC TTC CTG CCG TTG GCC TAT CAG TCC GCC CTG ATG
glu ile ile ser leu ile gly ile gly glu arg ala lys gly phe phe leu arg leu gly tyr gln ser gly leu met
30
GAT GCC GAG CTG GCA CGC AAG CTG CGG CCG GCC ATG CGC GAG GAG GAT TGG TTC CTG CGT GGG CCT CAA TTG TAT
asp ala glu leu ala arg leu pro ala met arg glu glu glu val phe leu ala gly pro gln leu tyr
40
GCC CTC AAG GGG ATG GTC AAA GTA GCC TTG CTG ACA ATG GAT ATC GCC ATC CGG GAC GGA CTG GCC AAG TTC AAC GTG GAG
ala leu gly lys met val lys val arg leu leu thr met asp ile ala ile arg asp gly arg phe asn val glu
50
GCC GAG TGG ATT GAT TCC TTT GAA GTG GAT ATC TGC CGT ACT GAG CTG GCC CTG ATG AAT GAG CCC GTC TGG
ala glu trp ile asp phe phe glu val asp ile asp ser phe glu val asp ile cys arg thr glu leu gly leu met asn glu pro val cys trp
60
AGC GTC TAA GCC TAT GCT AGC GGC TAT AGT TGG GCA TTC ATG GGC CGC AGA ATC ATT TTC CAG GAA ACT AGC TGT
thr val leu gly tyr ala ser gly tyr ser ser ala phe met gly arg arg ile ile phe glu thr ser cys
70
GCC GGG TGC GGC CAT AAA TGC CTT ATC GTC GGC AAG ACC GCA GAA GAG TGG GCC GAT GTC AGC AGT TTC GAA
arg gly cys gly asp asp lys cys leu ile val gly lys thr ala glu glu trp gly asp val ser ser phe glu
80
GCC TAC TTC AAA AGC GAC CCG ATC GTA GAT GAG CCG TAC GAG CTG CAG ACC CAG GTG GCC AAC CTG CGC AAC CGC
ala tyr phe lys ser asp pro ile val asp glu arg tyr glu leu gin thr gin val ala asn leu arg asn arg

leu lys gin tyr asp gly gin tyr tyr gly ile gly his ser pro ala tyr lys arg ile cys glu thr ile asp

AAG GCT GCA CGT GGC AGG GTT TCG TGC CTA CTG GGT GAG ACT GSG GTG GGC AAC GAG GTA ATC GCG CGC AGC
lys ala ala arg gly arg val ser val leu leu leu gly glu thr gly val gly lys glu val ile ala arg ser

val his leu arg ser glu arg ala glu gin pro phe val ala val asn cys ala ala ile pro pro asp leu ile

GAG TCG GAA CTG TTT GTT GTC GAT AAG GCC CCC TAT ACC GGC GCG TAT AAG GCC GCC TTC TCG GCC GTG AAT GCC GTG
GAA ATC CGT ATC GCC AAC ACC ACG TTT TGT GCC AAC GCC ACC CTG CTA CGG GTA

ala asn gly gly thr ile phe leu asp glu val ile glu leu thr pro arg ala gin ala thr leu leu arg val

TTG CAG GAA CGA GAG CTA GAG CGG TGC GCC GGC CGC CGC ACG CGA AAG CTC GAC GTG AGG TTA ATC ACC CGA ACA
leu gin glu gly leu glu arg val gly gly asp arg thr lys val asp val arg leu ile thr ala thr

410

AAC GAG AAC CTG GAA GAG CGC GTC AAG ATG GGG CGC TTT CGC GCC GAC CTG TTC TTT CGG CGT AAT GTT TTT CCC
asn glu asn leu glu ala val lys met gly arg phe arg ala asp leu phe phe arg leu asn val phe pro

420

GTG CAG ATC CCG CCG TTG CGC GAG CGC GTG GAA GAC ATC CCG CTG GTC CAG CAT TTT CTC AGA AGG CAC CAT
val gin ile pro pro leu arg glu arg val glu asp ile pro leu leu val glu his phe leu arg arg his his
AAG GAA TAC GCC AAG AAG ACT CTT GCC CTG TCT GAT CGA CGG ATG GAG GCC TGC CTC CAC TAC CAA TGG CCA GCC lys glu tyr gly lus lys thr leu gly leu ser asp arg ala met glu ala cys leu his tyr gln trp pro gly

AAT ATC CGC GAG CTG GAG AAG GCC CTT GAG CGC GGG GTG ATT CTC ACC GAG AGC AAC GAA AGC ATC AAT GTC GAG

asn ile arg glu leu glu asn ala leu glu arg gly val ile leu thr glu ser asn glu ser ile asn val glu

tcg ctt ccc ggg tgg cag atg cgc acc gaa ggc gac agg cta tca agc gat ggg gcc cgg tgg gag gag ggc tcc

ser leu phe pro gly leu ala met ala thr glu gly asp arg leu ser ser gly gly arg leu glu glu glu ser

ggt gac agt tgg ttt agg cca att atc gac cag ggc gtc agt ctc gac gat ctc gag ggc ggt tta atg cgc agg

gly asp ser trp phe arg glu ile ile asp gli gly val ser leu glu asp leu glu ala gly leu met arg thr

gcc atg gac cgt tgt ggg cag aat atc tca cag ggc ggc cgg tgg ctc gga tgt tgg ACC CCG GCA ATG GCC TAT

ala met asp arg cys gly glu asn ile ser glu ala leu leu glu thr arg pro ala met ala tyr

cga ctt aag aag ctt gac ccc agc cta tct gtg aaa gca aag ggc gca tag ctgctttgcc ccaggtttgg gcccttctaa

arg leu lys lys leu asp pro ser leu ser val lys ala thr gly arg ter
Figure 29. Codon usage table for the xylS gene of pDK1.
<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTC</td>
<td>phe</td>
<td>1</td>
</tr>
<tr>
<td>TTC</td>
<td>phe</td>
<td>3</td>
</tr>
<tr>
<td>TTA</td>
<td>leu</td>
<td>1</td>
</tr>
<tr>
<td>TTA</td>
<td>leu</td>
<td>2</td>
</tr>
<tr>
<td>TTA</td>
<td>leu</td>
<td>3</td>
</tr>
<tr>
<td>TTC</td>
<td>phe</td>
<td>5</td>
</tr>
<tr>
<td>TTC</td>
<td>phe</td>
<td>7</td>
</tr>
<tr>
<td>TTA</td>
<td>leu</td>
<td>9</td>
</tr>
<tr>
<td>TTA</td>
<td>leu</td>
<td>10</td>
</tr>
<tr>
<td>TTC</td>
<td>phe</td>
<td>12</td>
</tr>
<tr>
<td>TCA</td>
<td>ser</td>
<td>2</td>
</tr>
<tr>
<td>TCA</td>
<td>ser</td>
<td>3</td>
</tr>
<tr>
<td>TCA</td>
<td>ser</td>
<td>4</td>
</tr>
<tr>
<td>TCA</td>
<td>ser</td>
<td>5</td>
</tr>
<tr>
<td>TCA</td>
<td>ser</td>
<td>6</td>
</tr>
<tr>
<td>CCA</td>
<td>pro</td>
<td>2</td>
</tr>
<tr>
<td>CCA</td>
<td>pro</td>
<td>3</td>
</tr>
<tr>
<td>CCA</td>
<td>pro</td>
<td>4</td>
</tr>
<tr>
<td>CCA</td>
<td>pro</td>
<td>5</td>
</tr>
<tr>
<td>CCA</td>
<td>pro</td>
<td>6</td>
</tr>
<tr>
<td>ACT</td>
<td>thr</td>
<td>2</td>
</tr>
<tr>
<td>ACT</td>
<td>thr</td>
<td>3</td>
</tr>
<tr>
<td>ACT</td>
<td>thr</td>
<td>4</td>
</tr>
<tr>
<td>ACT</td>
<td>thr</td>
<td>5</td>
</tr>
<tr>
<td>ACT</td>
<td>thr</td>
<td>6</td>
</tr>
<tr>
<td>ACA</td>
<td>thr</td>
<td>7</td>
</tr>
<tr>
<td>ACA</td>
<td>thr</td>
<td>8</td>
</tr>
<tr>
<td>ACA</td>
<td>thr</td>
<td>9</td>
</tr>
<tr>
<td>ACA</td>
<td>thr</td>
<td>10</td>
</tr>
<tr>
<td>AGC</td>
<td>thr</td>
<td>1</td>
</tr>
<tr>
<td>AGC</td>
<td>thr</td>
<td>2</td>
</tr>
<tr>
<td>AGC</td>
<td>thr</td>
<td>3</td>
</tr>
<tr>
<td>AGC</td>
<td>thr</td>
<td>4</td>
</tr>
<tr>
<td>AGC</td>
<td>thr</td>
<td>5</td>
</tr>
<tr>
<td>GAT</td>
<td>asp</td>
<td>8</td>
</tr>
<tr>
<td>GAT</td>
<td>asp</td>
<td>9</td>
</tr>
<tr>
<td>GAT</td>
<td>asp</td>
<td>10</td>
</tr>
<tr>
<td>GAT</td>
<td>asp</td>
<td>11</td>
</tr>
<tr>
<td>GAT</td>
<td>asp</td>
<td>12</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>6</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>7</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>8</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>9</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>10</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>11</td>
</tr>
<tr>
<td>GAC</td>
<td>glu</td>
<td>12</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>13</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>14</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>15</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>16</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>17</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>18</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>19</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>20</td>
</tr>
<tr>
<td>GGC</td>
<td>gly</td>
<td>21</td>
</tr>
</tbody>
</table>
Figure 30. The nucleotide and predicted amino acid sequences of \textit{xylS} from the TOL plasmid pDK1.
ATG GAT TTT TGC TTA TTG AAC GAG AAA AGT CAG ATC TTC GTC CAC GCC GAG
met asp phe cys leu leu asn glu lys ser gln ile phe val his ala glu

CCC TAT GCA GTC TCC GAT TAT GTT AAC CAG TAT GTC GGT ACG CAC TCT ATT
pro tyr ala val ser asp tyr val asn gln tyr val gly thr his ser ile

CGC CTG CCC AAG GGC GGS CGC CCG GCA GGC AGG CTG CAC CAC AGA ATC TTC arg leu pro lys gly gly arg pro ala gly arg leu his his arg ile phe

GGA TGC CTC GAC CTG TGT CGA ATC AGC TAC AGC GCC GAC AAG CTG GGA ATC gly cys leu asp leu cys arg ile ser tyr gly gly ser val arg val ile

TCG CCT GCA TTA GAG ACC TGT TAT CAT CTG CAA ATA ATA CTC AAA GCC CAT ser pro gly leu thr cys tyr his leu gln ile ile leu lys gly his

TGC CTG TGG CGT GCC TAT GGC CAG GAC CAC TAT TTT TCG CGG GCC GAA CTA cys leu trp arg gly tyr gly gly gln tyr phe ser pro gly gly leu

TGG CTG CTC AAT CCG GAT GAC CAA GCC GAC CTG ACC TAT CTA GAA GAT TGC leu leu leu asn pro asp asp gln ala asp leu thr tyr ser glu asp cys

GAG AAA TTT ATC GTT AAA TTG CCC TCA GTG GTC CTT GAT CGG GCA TGC AGT glu lys phe ile val lys leu pro ser val val leu asp arg ala cys ser

GAC AAT TGG CAC AAG CCG AGG GAG GTT ATC GGT TTC GCC GCC CGA CAC asp asn asp leu asp asp asp gln ala asp leu thr tyr ser glu asp cys

AAT CTC CAG CAA CTC GAT GCC TTT ATC AAT CTA CTC GGG TTA GTT TGT GAC asn leu gln gln leu asp gly phe ile asn leu gly leu val lys asp

GAA GCC GAA CAT ACA AAG TCG ATG CCT CGG CCA GAG CAC TAT CGG GGG glu ala glu his thr lys ser met pro arg val gln glu his tyr ala gly

ATC ATC GCT TCC AAG CTG CTC GAA ATG CTG GCC AGC AAT GTC AGC CTG GAA ile ile asl ser leu leu glu met pro arg glu ser asn val ser arg glu

ATT TTC AGC AAA GGT AAC CCG TCT TCC GAG CGA GTC GTT CAA TTC ATT GAG ile phe ser lys gly asn pro ser phe glu arg val gln phe ile glu

GAG AAT CTC AAA CGG AAT ATC AGC CTT GAG CGG TTA GCC GAG CAC CTG GCC CAC TGC glu asp leu lys arg asn ile ser leu glu arg leu ala glu

ATG AGC CCA CGC TCG CTC TAC ACT TTG TTC GAG AAG CAT GCT GCC ACC ACG met ser pro arg ser leu tyr thr leu phe glu lys his ala gly thr thr

CCG AAG AAC TAC ATC CGC AAC CGC AAG CTC GAG TGC AAG AAG ACG TGG pro lys asn tyr ile arg asn arg lys leu glu cys ile arg ala arg leu
AGC GAT CCC AAT GCA AAT GTG CGT AGT GTC ACC GAG ATG GCC CTA GAC TAC
er asp pro asn ala asn val arg ser val thr glu met ala leu asp tyr

GGC TTC TTI CAT ACG GGA CGT TTC GCC GAG AAC TAT AGG AGC ACT TTC GCC
gly phe phe his thr gly arg phe ala glu asn tyr arg ser thr phe gly

GAG CTG CCT TCC GAC ACC CTG CGT CGG CGC AAA ATG AAG TGG CTT GAT CCC
glu leu pro ser asp thr leu arg arg lys met lys trp leu asp pro

GAG GAG AGC CTG CCC CCA CTC CCT TGA
 glu glu ser leu pro pro leu pro OPA
computer was used. The Mac DNASIS Pro™ version 1.0 DNA and protein sequence analysis software for Apple Macintosh computers was obtained from Hitachi Software Engineering America, Ltd. (Brisbane, California). Both GenBank® and the EMBL DNA bank were utilized for DNA and amino acid sequence homology searches using CD-ROM (Compact Disc-Read Only Memory). The DNA Strider program was also used for DNA analysis. Both nucleotide and amino acid sequences were compared. Potential RNA secondary structures were evaluated as well. Hydrophobicity profiles (Kyte and Doolittle, 1982) for xylR and xylS proteins are shown in Figures 31 and 32.

A. Computer analysis of DNA sequences

DNA sequences of the pDK1 xylRS were compared to those derived by others for the archetypal TOL plasmid pWW0. Sequences from other TOL plasmids such as pWW53 and pWW15 were also used for comparisons. Both xylR and xylS nucleotide sequences and their derived amino acid sequences were used for these analysis.

DNA and amino acid sequence comparisons of two genes encoding the TOL transcriptional activators (xylR and xylS) of pDK1 and pWW0 show strong homology between both pairs of genes.

A comparison of the xylR nucleotide sequences from the two TOL plasmids (Fig. 33) reveals remarkable 98% level of
Figure 31. Kyte and Doolittle hydrophobicity profile of pDK1 xylR gene product. Positive numbers are assigned to amino acids with hydrophobic (apolar) side chains and negative numbers to those with hydrophilic (polar) side chains.
Figure 32. Hydrophobicity profile of pDK1 xylS gene product.
Figure 33. Comparison of xylR nucleotide sequences from pWW0 and pDK1.
overall sequence homology (1669/1701 nucleotides identical). The 32 nucleotide differences affect 32 codons, with 5 changes in first position, 4 in position two and 23 in position three. Of the 32 altered codons, 23 exhibit no change in the amino acid encoded due to the degeneracy of the genetic code. Further, several amino acid substitutides are quite neutral, being ile/val/leu switched. The encoded proteins are thus an equally high 98.4 % identical (557/566 amino acid identical) (Fig. 34). The high level of nucleotide sequence homology (98%) continue for about 190 bp beyond the end of the xylR ORF and then abruptly drops to an insignificant level (ran down).

A similar comparison of the xylS genes from the pDK1 and pWW0 TOL plasmids again reveals a high level of sequence homology (Fig. 35). Only 32 nucleotides vary between the pDK1 sequence and the coding region of the pWW0 gene. However one of these (at position #955) is a single base insertion/deletion. This cause the pDK1 xylS ORF to be 30 nucleotides longer than that for pWW0. Up to the point of this frameshift the two sequences exhibit an overall level of homology at 96.8% (935/966 nucleotides). The 31 changes affected 29 codons. Of these, 10 gave no change in the amino acid sequence. Thus, the overall amino acid sequence homology was 94.1% (302/321 identical before frameshift).
Figure 34. Comparison of the amino acid sequences of XylR proteins from the TOL plasmids pDK1 and pWW0.
MSLTYKPKMQ HEDMQDLSSQ IRFVAEGKI WLGEQRMLVM QLSTLASFRR pWW0
MSLIYNPKMQ HEDMQDLISQ IRFVAEGKI WLGEQRMLLM QLSTLASFRR pDK1
EIISLIGVER AKGFFRLRLGY QSGLMDAEIKA RKLPRPAMRE EVFLAGPQLY 100
EIISLIGIER AKGFFRLRLGY QSGLMDAEIKA RKLPRPAMRE EVFLAGPQLY
ALKGMVKVRL LTMDIAIRDG RFNVEAEWID SFEVDCRTLE LGLMNEPVCW
ALKGMVKVRL LTMDIAIRDG RFNVEAEWID SFEVDCRTLE LGLMNEPVCW
TVLGYASGYS SAFMGRRIIF QETSCRGGCD DKCLIVGTKTA EEWGDVSSFE 200
TVLGYASGYS SAFMGRRIIF QETSCRGGCD DKCLIVGTKTA EEWGDVSSFE
AYFKSDPIVD ERYELQTQVA NLRNRLKQYD GQYGIGHSP AYKRICEITID
AYFKSDPIVD ERYELQTQVA NLRNRLKQYD GQYGIGHSP AYKRICEITID
KAARGRVSVL LLGETGVGKE VIARSVHRLS ERAEQPFVAV NCAAIIPDLI 300
KAARGRVSVL LLGETGVGKE VIARSVHRLS ERAEQPFVAV NCAAIIPDLI
ESELFAGVDKG AYTGAVNARA GRFERANGGT IFLDEVIELT PRAQATLLRV
ESELFAGVDKG AYTGAVNARA GRFERANGGT IFLDEVIELT PRAQATLLRV
LQEGELERVG GDTRKVDVR LITATNENLE EAVKMGFRFA DLFFRLNVFP 400
LQEGELERVG GDTRKVDVR LITATNENLE EAVKMGFRFA DLFFRLNVFP
VHIPPLRERV EDIPLLVEHF LRRHKEYGYK KTLGKSDRAM EACLHYQWPG
VQIPPLRERV EDIPLLVEHF LRRHKEYGYK KTLGKSDRAM EACLHYQWPG
NIRELENALE RGVILTESNE SINVESLFPG LATATEGDRL SSEGRLBBS 500
NIRELENALE RGVILTESNE SINVESLFPG LAMATEGDRL SSEGRLBBS
GDSWFRQIID QGVSELEDLEA GLMRTAMDRC GQNISQAARL LGLTRPAMAY
GDSWFRQIID QGVSELEDLEA GLMRTAMDRC GQNISQAARL LGLTRPAMAY
RLKKLDPSLS VKAMGRZ
RLKKLDPSLS VKATGRZ
Figure 35. Comparison of xylS nucleotide sequences from pWW0 and pDK1.
<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'</td>
<td>ATGGATTTTTGGCTTATGAAGGAAAATCAGATCTCTGCACTGCGACACCGAGCCGCTGCTGCACTCCTCTCCGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>71</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
</tr>
<tr>
<td>141</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
</tr>
<tr>
<td>211</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
<td>270</td>
<td>280</td>
</tr>
<tr>
<td>281</td>
<td>290</td>
<td>300</td>
<td>310</td>
<td>320</td>
<td>330</td>
<td>340</td>
<td>350</td>
</tr>
<tr>
<td>351</td>
<td>360</td>
<td>370</td>
<td>380</td>
<td>390</td>
<td>400</td>
<td>410</td>
<td>420</td>
</tr>
<tr>
<td>421</td>
<td>430</td>
<td>440</td>
<td>450</td>
<td>460</td>
<td>470</td>
<td>480</td>
<td>490</td>
</tr>
<tr>
<td>491</td>
<td>500</td>
<td>510</td>
<td>520</td>
<td>530</td>
<td>540</td>
<td>550</td>
<td>560</td>
</tr>
</tbody>
</table>
GATCATCGCTTCCAAAGCTGCAGTTGCTGAGCCAGCAATGTCAGCCGTGAAATTTTCAGCAAAGGTAC
GATCATCGCTTCCAAAGCTGCAGTTGCTGAGCCAGCAATGTCAGCCGTGAAATTTTCAGCAAAGGTAC

CGGTCTTTTCAGGGTCGTTTCAATTCATTGAGGAGAATCTCAAACGGAATATCAGCCTTGAGCGGTTAG
CGGTCTTTTCAGGGTCGTTTCAATTCATTGAGGAGAATCTCAAACGGAATATCAGCCTTGAGCGGTTAG

CGGAGCTGGCGGTAGTGAACGGATGCGACGGCTTCTTTCATACGGGACGTTTCGCCGAGAACTATAGGAGCA
CGGAGCTGGCGGTAGTGAACGGATGCGACGGCTTCTTTCATACGGGACGTTTCGCCGAGAACTATAGGAGCA

GAACTACATCCGCAACCGCAAGCTCGAGTGCATACGCGCCCGCTTGAGCGATCCCAATGCAAATGTGCGT
GAACTACATCCGCAACCGCAAGCTCGAGTGCATACGCGCCCGCTTGAGCGATCCCAATGCAAATGTGCGT

AGTGTCACCGAGATGGCCCTAGACTACGGCTTCTTTCATACGGGACGTTTCGCCGAGAACTATAGGAGCA
AGTGTCACCGAGATGGCCCTAGACTACGGCTTCTTTCATACGGGACGTTTCGCCGAGAACTATAGGAGCA

CTTTCCGGAGSGCTGGCTCCGACACCTCGCGCTGCGCCCGCAATGAAATGGCCTGATCGCCGGAGACGGAGCC
CTTTCCGGAGSGCTGGCTCCGACACCTCGCGCTGCGCCCGCAATGAAATGGCCTGATCGCCGGAGACGGAGCC

CCCCACCTCCCCCTGGACAGTGGCCCTTACAGGCCG 3' pDK1
CCCCACCTCCCCCTGGACAGTGGCCCTTACAGGCCG 3' pDK1

pWWO
This homology is not evenly distributed across the gene, however. The N-terminal (or 5' end) 711 nucleotides are 99.7% homologous (709/711 identical). The C-terminal 249 nucleotides (712-961) are therefore only 87.5% homologous. 10 of the 31 nucleotide changes and 4 of 20 amino acid substitutions are found in the putative helix-turn-helix DNA binding domain of the C-terminal end. The homology drops to random values at the termination signal for the pWW0 xylS gene.

The intergenic region (Fig. 36), including the promoters (Fig. 37 and Table 15) and other regulatory sequences (Fig. 38), is 444 nucleotides in length for pDK1 (445 for pWW0). Here we again find a very high level of homology (Fig. 39) with only 10/445 nucleotides being altered (9 substitutions, 1 deletion/addition). This gives a 97.7% level of overall homology.

A further comparison of the amino acid sequence of the XylS protein from pDK1 to that from the TOL plasmid pWW53 also reveals an interesting regional variation in the level of homology (Fig. 40). In this case the lowest level of homology is observed for the N-terminal two-thirds of the protein and the highest (including an identical insertion/deletion) in the C-terminal region. The "break point" for this differential homology is again in the
Figure 36. The GC content of the xyRS intergenic region (2,000 to 2,500).
Figure 37. $\sigma^{54}$- and $\sigma^{70}$-dependent promoters of the TOL plasmids pWW0 and pDK1.
TOL plasmid pWW0

1. $\sigma^{54}$-dependent promoters

$P_s$ TTTCTACATC ACACCAAGCA GCCCACATTA AAATAAGAGA ACCGTGAACCT ATG
$OP1$ GCGGTGTGCTA GCTATACGAG ACTTTAAATA AAAATAGTGG TGACCCCTCA ATG

2. $\sigma^{70}$-dependent promoters

$P_r$ TCTGATGTTC CGCCGGGTGG ATCATCCCGA TAAAAACAAG AGGAAAACAA ATG
$OP2$ CTACCCCTTA GGCTTTATGC AACAGAAACA ATAATAATGG AATTTATG AGTCATGACC ATG

TOL plasmid pDK1

1. $\sigma^{54}$-dependent promoters

$P_s$ TTTCTACATC ACACCAAGCA GCCCACATTA AAATAAGAGA ACCGTGAGCT ATG
$OP1$ ATG

2. $\sigma^{70}$-dependent promoters

$P_r$ TCTGATGTTC CGCCGGGTGG ATCATCCCGA TAAAAACAAG AGGAAAACAA ATG
$OP2$ CTAGCTCTTA GGCTTTATGC AACTGAAACA ATAATAATGG AATTTATG AGTCACGAAA ATG
Figure 38. Comparison of putative ribosome binding sites for xyl genes of the TOL plasmids pDK1 and pWW0.
Putative ribosome binding site (RBS) underlined

<table>
<thead>
<tr>
<th>pWWO</th>
<th>xylR</th>
<th>UAAAACAAAGGAAAAACAA AUG</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAAAACAAAGGAAAAACAA AUG</td>
<td>xylR</td>
<td>UAAAACAAAGGAAAAACAA AUG</td>
</tr>
<tr>
<td>UAAAACAAAGGACAACCU AUG</td>
<td>xylS</td>
<td>UAAAACAAAGGACAACCU AUG</td>
</tr>
<tr>
<td>GAACACAGGACGGACGUC AUG</td>
<td>xylE</td>
<td>GAACACAGGACGGACGUC AUG</td>
</tr>
<tr>
<td>AAUAA AUAGGCAUGGACGACC AUG</td>
<td>xylX</td>
<td>AAUAA AUAGGCAUGGACGACC AUG</td>
</tr>
<tr>
<td>GAUCCACGGGAGGCGUGUAA AUG</td>
<td>xylY</td>
<td>GAUCCACGGGAGGCGUGUAA AUG</td>
</tr>
<tr>
<td>AUACCACGUCUGGUGCCGCC AUG</td>
<td>xylZ</td>
<td>AUACCACGUCUGGUGCCGCC AUG</td>
</tr>
<tr>
<td>GGCAACACGACGACGGGUUC AUG</td>
<td>xylL</td>
<td>GGCAACACGACGACGGGUUC AUG</td>
</tr>
<tr>
<td>ACUUAUUGCAAGGAUGGCAUG AUG</td>
<td>xylG</td>
<td>ACUUAUUGCAAGGAUGGCAUG AUG</td>
</tr>
<tr>
<td>GAAACUUGGAGGGACGUGGUGC AUG</td>
<td>xylF</td>
<td>GAAACUUGGAGGGACGUGGUGC AUG</td>
</tr>
<tr>
<td>TTUUCUGAGGAGACCCGCAU AUG</td>
<td>xylJ</td>
<td>TTUUCUGAGGAGACCCGCAU AUG</td>
</tr>
<tr>
<td>AUAUAA AUAGGCAUGGCAUG AUG</td>
<td>xylM</td>
<td>AUAUAA AUAGGCAUGGCAUG AUG</td>
</tr>
<tr>
<td>CGUTCACUUGGAGGCAUGGCUU AUG</td>
<td>xylA</td>
<td>CGUTCACUUGGAGGCAUGGCUU AUG</td>
</tr>
<tr>
<td>xylO</td>
<td>UCAUGUGACGCAUGGCAUGGCUU AUG</td>
<td></td>
</tr>
<tr>
<td>xylK</td>
<td>GAGGACGAGCC AUG</td>
<td></td>
</tr>
<tr>
<td>xylI</td>
<td>CAACUCUCAAGGACGCAUGC AUG</td>
<td></td>
</tr>
<tr>
<td>xylH</td>
<td>UUCAGGAGGCAACAC AUG</td>
<td></td>
</tr>
</tbody>
</table>
Figure 39. Comparisons of xyIRS operator-promoter region from pWW0 and pDK1.
Comparison of pDK1 and pWW0 xyIRs Promoter Regions

\[\text{xyIR} \quad 2030\]
GATTGTATAT AAGGGACATT -GTTTCCCTCT TGTTTTTATC GGGATGAATCC ACCCAGCAGC ACATCAGTGG CTCACTGCCG
GT TTGTATgT AAGggACATT GTTTTTCCCTCT TGTTTTTATC GGGATGAATCC ACCCggCCgGa ACATCAGgGgGg CTCACTGCCG

\[\text{xyIS} \quad 2100\]
TCATGACTAC CTGCAACGTT CAAGGCCCTGT GCTTGCCGGGC TCAGCATTTCC ATCTGCCAAT TTAGGATTGCT TTAGGGTGCT
TCATGACTAC CTGCAACGTT CAAGGCCCTGT GCTTGCCGGGC TCAGCATTTCC ATCTGCCAAT TTAGGATTGCT TTAGGGTGCT

\[\text{xyIS} \quad 2200\]
CCTGAAAAAT TAACCAATTTG ATTAACAGAA ATCCACTCA CACGAGATCC CCCCCATCTGG CCGCGCAGCC
CCTGAAAAAT TAACCAATTTG ATTAACAGAA ATCCACTCA CACGAGATCC CCCCCATCTGG CCGCGCAGCC

\[\text{xyIS} \quad 2300\]
CCTGACAGCAC GAACTTCGG TGTTTCGCCT TGTTAAAAAG AAGGGCTTCG TTCTGCGTAGG CTTTAATTTT GCCGGAAAAA
CCTGACAGCAC GAACTTCGG TGTTTCGCCT TGTTAAAAAG AAGGGCTTCG TTCTGCGTAGG CTTTAATTTT GCCGGAAAAA

\[\text{xyIS} \quad 2400\]
GTTGGTCTACA ATGCAAAAAT GATGCCCGCA AGCTGGGCAAT GGGGTTAACC CGTATCCGCA TCAGCTCGAG ATGCCATTTTC
GTTGGTCTACA ATGCAAAAAT GATGCCCGCA AGCTGGGCAAT GGGGTTAACC CGTATCCGCA TCAGCTCGAG ATGCCATTTTC

\[\text{xyIS} \quad 2500\]
AATCGACCTGG CGCCCTTCTTA CATCACACCA AAGCGACCAC AITAAAAATT AAAACCTCGTTTAT
AATCGACCTGG CGCCCTTCTTA CATCACACCA AAGCGACCAC AITAAAAATT AAAACCTCGTTTAT
Table 15. Comparison of different classes of TOL promoters.

<table>
<thead>
<tr>
<th>Promoters</th>
<th>Consensus sequence</th>
<th>Sigma factors</th>
<th>Activator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_u$ (OP1)</td>
<td>-12/-24</td>
<td>Sigma 54</td>
<td>XylR</td>
</tr>
<tr>
<td>$P_m$ (OP2)</td>
<td>-10/-35</td>
<td>Sigma 70</td>
<td>XylS required</td>
</tr>
<tr>
<td>$P_R$</td>
<td>-10/-35</td>
<td>Sigma 70</td>
<td>Not required</td>
</tr>
<tr>
<td>$P_S$</td>
<td>-12/-24</td>
<td>Sigma 54</td>
<td>XylR</td>
</tr>
</tbody>
</table>
Figure 40. Comparison of XylS amino acid sequences from pWW0, pDK1 and pWW53 (S3). The N-terminal proteins of both pDK1 and pWW0 are nearly identical (amino acid 1 to 90) while the C-terminal ends of pDK1 and pWW53 (S3) are more similar (227 to 331).
<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDFCLLNEKS QIFVHAEPYA VSDYVNYQVG THSIRLPKGG RPAGRLHHR1 FGCLDLCRIS pWW0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDcrlLLNEKS rIFVnAdPYf VSDYVQhVG sHcIRLPKsG cPeasLnHst FGsLDLCRIS pWW53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YGGSVRVISP GLETCHLYQI ILKGHCLWRG hGQEHYFspG ELLLNPDDQ ADLTYESDECE S3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XFIvKLPSVv LDLACSDNNW HKPREGIRFA ARHNLQQLDLG EELLGLVLCD EAEHTKSMPR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VQEHYAGIA SKLLEMLGSN VSREIFSKGN PSFERVQFQI EENLKRNISSL ERLAEALMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VQSTFGELPS DTLRqckkev aZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YRSTFGELPS DTLRRRKMkw LDPEESLVP PL PZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(S3)
vicinity of nucleotide 700 of the xylS ORF.
Microbes play an important role in the maintenance of the carbon cycle of the ecosystem. The study of microbial catabolic genes is therefore ecologically important. In addition, the study of catabolic pathways can be significant in terms of industrial applications. Molecular biologists are interested in modifying soil microorganisms for the purpose of producing useful chemicals at reduced prices. For example, cis-diols are compounds of biotechnological interest because they can be used as building blocks for polyphenylene synthesis or serve as reactants for organic chemical synthesis (Wubbolts and Timmis, 1990). Thus, the oxidation of substituted benzoates to cis-diols by the TOL enzyme toluate-1,2-dioxygenase (xylXYZ) provides a valuable mechanism for the production of a wide variety of fine chemicals. At least two of the TOL enzymes are also potentially useful as reporter functions in cloning vehicles.

The study of TOL plasmids is academically important as well. For example, TOL plasmids can serve as useful model systems for the study of gene structure, gene regulation and
DNA-protein interactions.

TOL plasmids are often conjugative and the TOL functions are frequently associated with transposable elements. Conjugative plasmids and transposons are important agents in the process of microbial evolution. DNA sequence analysis will undoubtedly contribute evidence indicative of how these conjugative plasmids and transposons have evolved. This will simultaneously provide additional insight into the evolution of aromatic catabolic pathways in bacteria.

The pBK990 plasmid constructed during my research efforts carries the complete \textit{xylR} and \textit{xylS} transcriptional units, as well as the OP1 promoter region. By encoding \textit{xylR} as well as the two promoters regulated by this protein, pBK990 will be very useful for studying gene regulation and transcriptional control. This recombinant plasmid can also be used for studying \textit{σ}^{54}-dependent promoters by researchers using microorganisms such as \textit{Klebsiella}, \textit{Rhizobium}, \textit{Azotobacter}, \textit{Pseudomonas} and the enterics.

The nucleotide sequence of a 4 kilobase segment of the \textit{P. putida} TOL plasmid pDK1 which I have determined, encodes the two positive regulatory proteins required for the expression of the upper (\textit{xylCMABN}) and lower (\textit{xylXYZLTEGRJQKIH}) TOL operons. Comparisons of this region
to the equivalent regions of the related *P. putida* TOL plasmids (pWW0 and pWW53) have revealed several interesting observations. An overall DNA sequence homology for a 3.3 kb segment, encoding primarily the *xylR* and *xylS* genes, was found to be greater than 97% for the pWW0 and pDK1 TOL plasmids. Comparisons of the encoded amino acid sequences of the *xylR* and *xylS* genes between TOL plasmids also reveals high levels of sequence identity between homologous proteins. The *XylR* proteins of pWW0 and pDK1 are 98.4% identical, while the *XylS* proteins are 96.8% identical. The *XylS* proteins of pWW53 and pDK1 were found to have similar high levels of homology, being 87.6% identical (290/331 amino acid identical).

Beyond the *xylRS* genes the homology falls off rapidly. In the case of *xylR* it becomes random about 190 nucleotides beyond the translational stop signal when comparing pWW0 and pDK1 sequences. For *xylS* the homology ends with the stop signal itself. This loss of homology beyond the central *xylRS* region is not surprising and presumably reflects the break junction points of the rearrangements involving the lower operon, upper operon and regulatory gene regions which have occurred since the evolutionary lines leading to these plasmids diverged. It should be remembered that the arrangement of these three regions within the TOL DNA varies.
from one plasmid to another. It is, however, somewhat surprising that the level of nucleotide sequence homology between pWW0 and pDK1 is so high within this \textit{xylRS} region (>97%) when one considers the previous findings for the \textit{xylE} region of the lower operon. The two \textit{xylE} genes from pDK1 and pWW0 were found to share only approximately 80% homology with each other and in fact a very similar level with the \textit{nahH}, the isofunctional C230 gene of the NAH7 naphthalene degradation plasmid. While some of the nearly 98% homology shared by the two \textit{XylR} proteins may be explained by the multiple binding sites with the protein (promoter region, inducer, RNA polymerase) and restrictions related to the functions of these proteins, it does not explain the limited number of silent third position substitutions observed in the two \textit{xylR} genes which were previously found to occur rather commonly in the \textit{xylE} genes. Although several possible explanations for this difference exist, one would be that the \textit{xylRS} regions of the two plasmids share a more recent common ancestral sequence than the \textit{xylE} regions. This would suggest that TOL plasmids have continued to "share" nucleotide sequence information via interplasmid recombination events long after individual plasmid lines have diverged. Thus the rearrangements leading to organizational differences with the TOL regions may reflect a combination of intraplasmid and
interplasmid recombinational events.

A possible "hot spot" for such recombinational events may exist in \textit{xylS}. The \textit{xylS} genes of pWW0 and pDK1 are 97% identical. However, most of the differences occur after position 700, giving a 99.7% identity up to that point and only 88.9% for the remainder of the C-terminal region. A comparison of XylS amino acid sequences of pDK1 and pWW53 reveals a similarly unbalanced distribution of homology, with low homology up to chart position 210 (position 3100 in nucleotides) and near identity for the remainder (Fig. 41). In addition, growth on benzoate cause the loss of TOL\textsuperscript{+} phenotype by strains carrying pDK1. This loss is due to a deletion from the upper operon to the \textit{xylS} gene (D. A. Kunz, personal communication). All of these observations are consistent with the \textit{xylS} region being a common point for both the proposed interplasmid and intraplasmid recombination events. Further nucleotide sequence data from additional related TOL plasmids may provide additional insight into whether or not this is a viable hypothesis.

It will also be of interest to determine if any functional genes lie between the \textit{xylRS} region and the upper and lower operons. There are the beginning of potential ORF's downstream of \textit{xylS} and \textit{xylR}. Two oligonucleotide primers and the pBK990 clone will be used for the purpose of
Figure 41. Comparison of the XylRXylS region of related TOL plasmids.
Comparisons of the XyIRXyIS Region of Related TOL Plasmids
attempting to verify and identify these unknown ORF's. Hopefully, the sequence can be extended until it overlaps with the region of the OP1 sequence. A similar effort will be made to further characterize the open reading frame (ORF) downstream of xylS which also remains to be identified.
BIBLIOGRAPHY


**Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983.


Hydroquinone as the ring-fission substrate in the catabolism of 4-ethylphenol and 4-hydroxyacetophenone by *Pseudomonas putida* JD1. J. Gen. Microbiol. **133**:2137-2146.


Ensley, B. D., B. J. Ratzkin, T. D. Osslund, M. J.


Franklin, F. C. H., P. R. Lehrbach, R. Lurz, B.


Construction of a partial diploid for the degradative pathway encoded by the TOL plasmid (pWW0) from Pseudomonas putida mt-2: evidence for the positive nature


Hansen, J., and R. H. Olsen. 1978. Isolation of large
bacterial plasmids and characterization of the P2 incompatibility group plasmids pMG1 and pMG5. J. Bacteriol. **135**:227-238.


Harayama, S., M. Rekie, A. Wasserfallen, and A. Bairoch. 1987. Evolutionary relationships between catabolic pathways for aromatics: conservation of gene order and nucleotide sequences of catechol oxidation

Harayama, S., M. Rekik, M. G. Wubbolts, K. Rose, R. A.
Leppik, and K. N. Timmis. 1989. Characterization of
five genes in the upper-pathway operon of TOL plasmid
pWW0 from Pseudomonas putida and identification of the

Harayama, S., and K. N. Timmis. 1989. Catabolism of
aromatic hydrocarbons by Pseudomonas, pp. 151-174. In D.
A. Hopwood, and K. F. Chater (eds.), Genetics of

Nucleotide sequence of the xylXYZ region of the
Pseudomonas putida TOL plasmid pDK1 and expression of the
encoded toluate-1,2-dioxygenase in Escherichia coli.
Soc. Microbiol., Washington, D.C.

Hartnett, C. S., E. L. Neidle, K.-L. Ngai, and L. N.
Ornston. 1990. DNA sequences of genes encoding
Acinetobacter calcoaceticus protocatechuate 3,


Inouye, S., A. Nakazawa, and T. Nakazawa. 1986. Nucleotide sequence of the regulatory gene xylS on the Pseudomonas putida TOL plasmid and identification of the


Company, New York.


Lowry O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the folin


Mermod, N., P. R. Lehrbach, W. Reineke, and K. N.


Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.


Nakai, C., H. Kagamiyama, M. Nozaki, T. Nakazawa, S.


Nakazawa, T., E. Hayashi, T. Yokota, Y. Ebina, and A.


Ribbons, D. W., and R. W. Eaton. 1982. Chemical transformation of aromatic hydrocarbons that support the
growth of microorganisms, pp. 59-84. In A. M.
Chakrabarty (ed.), Biodegradation and Detoxification of
Environmental Pollutants. CRC Press, Boca Raton, Fla.

1977. Labelling of deoxynucleotide acid to high specific
activity in vitro by nick translation with DNA

Knackmuss. 1987. Assemblage of ortho-cleavage route
for simultaneous degradation of chloro- and

Rommermann, D., J. Warrelmann, R. A. Bender, and B.
eutrophus and Pseudomonas facilis controls expression of
diverse metabolic pathways, including hydrogen oxidaton.

Ronson, C. W., B. T. Nixon, L. M. Albrighet, and F. M.
Ausubel. 1987. Rhizobium meliloti ntrA (rpoN) gene is
169:2424-2431.


Shaw, L. E., and P. A. Williams. 1988. Physical and


Spooner, R. A., K. Lindsay, and F. C. H. Franklin. 1986. Genetic, functional and sequence analysis of the


Tabor, S., and C. C. Richardson. 1990. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase:


Willetts, N., and R. Skurray. 1987. Structure and


Williams, P. A., S. D. Taylor, and L. E. Gibb. 1988. Loss of the toluene-xylene catabolic genes of TOL plasmid pWW0 during growth of *pseudomonas putida* on benzoate is due to a selective growth advantage of 'cured'


**Worsey, M. J., F. C. H. Franklin, and P. A. Williams.**


Biotransformation of substituted benzoates to the corresponding cis-diols by an engineered strain of Pseudomonas oleovorans producing the TOL plasmid-specified enzyme toluate-1,2-dioxygenase. Appl. Environ. Microbiol. **56**:569-571.


Zhou, L., K. N. Timmis, and J. L. Ramos. 1990. Mutations leading to constitutive expression from the TOL plasmid meta-cleavage pathway operon are located at the
C-terminal end of the positive regulator protein XylS.