PHYSICAL AND FUNCTIONAL CHARACTERIZATION OF THE xy1XYZ REGION FROM TOL PLASMID PDK1 AND ITS ASSOCIATED DOWNSTREAM REGULATORY ELEMENTS

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

Douglas R. Hares, B.S.
Denton, Texas
August, 1998

The nucleotide sequence for the pDK1 TOL plasmid region encoding toluate-1,2-dioxygenase (XylXYZ, TO) was determined. TO is the first enzyme in the meta-cleavage operon, responsible for the conversion of toluates and benzoates to their carboxy-substituted diols. DNA sequence analysis revealed the presence of three open reading frames (ORF). The three ORFs correspond to xylX (1353 bp), xylY (486 bp) and xylZ (1008 bp), encoding predicted protein products of 51370 Da, 19368 Da and 36256 Da, respectively. The average G+C content of the xylXYZ region was 58.9%, substantially lower than the 64.5% average for the remaining meta operon genes (xylLTEGFPJQKIH). High nucleotide and deduced amino acid sequence homologies were found between xylXYZ pDK1 and three other sequences, xylXYZ (toluate-1,2-dioxygenase) from archetype TOL plasmid pWW0, benABC (benzoate dioxygenase) from Acinetobacter calcoaceticus, and cbdABC (2-halobenzoate-1,2-dioxygenase) from Burkholderia cepacia. All four enzymes belong to Class IB hydroxylating bacterial dioxygenases. Further DNA sequence analysis of the pDK1 meta-cleavage operon uncovered the existence of a putative rho-independent
transcriptional terminator structure immediately downstream of xylZ and another possible stem and loop structure in the intergenic region between xylL and xylT. Site-directed mutagenesis was used to destabilize the terminator-like structure, lowering the predicted ΔG value by approximately one-third. The effects of mutations were analyzed by comparing DHCDH (1,2-dihydrocyclohexa-3,5-diene carboxylate dehydrogenase, encoded by xylL) and C230 specific activities between E. coli strains carrying mutant and wild-type constructs. Carefully controlled studies provided evidence that this structure has little or no effect on the transcription of downstream genes in E. coli and may simply be an evolutionary relic. Discussions on the evolutionary significance of the two intergenic gaps between xylZ/xylL and xylL/xylT are presented.
PHYSICAL AND FUNCTIONAL CHARACTERIZATION OF THE xy1XYZ REGION FROM TOL PLASMID PDK1 AND ITS ASSOCIATED DOWNSTREAM REGULATORY ELEMENTS

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

Douglas R. Hares, B.S.
Denton, Texas
August, 1998
ACKNOWLEDGEMENTS

I would like to express my appreciation to my friend, mentor, and major professor, Dr. Robert C. Benjamin for his support and guidance throughout my studies at the University of North Texas. Many thanks to Dr. Gerard O’Donovan for his scientific input and his belief in me. My thanks also go to Dr. Arthur Eisenberg, Dr. Earl Zimmerman and Dr. Sarah McIntire for their effort, time and patience serving as committee members. I would also like to thank Dr. Daniel Kunz for his advice and use of his laboratory for the many enzyme assays completed during the course of this study. To my fellow graduate students in Dr. Benjamin’s lab, thanks for all the support and advice throughout the years.

Most of all I would like to thank my entire family for their support through the years. A special thank you to my wife Heather and my son Timothy, whose patience and support for me never failed.
TABLE OF CONTENTS

LIST OF TABLES................................................. vi
LIST OF ILLUSTRATIONS........................................ vii
LIST OF ABBREVIATIONS....................................... xi

CHAPTERS

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

   Organization of TOL pathway genes
   Ortho-cleavage pathway
   Regulation of the TOL pathway
   The xylXYZ genes and toluate-1,2-dioxygenase
   Evolution of the TOL meta-cleavage operon
   Genetic Engineering and biodegradative pathways

II. MATERIALS AND METHODS................................. 36

   Bacterial strains and plasmids
   Biological media and growth conditions
   Long-term storage of cultures
   Large scale plasmid isolation from E. coli strains
   Quantitative and qualitative identification of DNA
   Complete and partial digestion of DNA with
   restriction endonucleases
   Bacterial alkaline phosphatase treatment of vector
   DNA
   Phenol extraction of DNA solutions
   Ligation of DNA for cloning into plasmids and RF
   phage
   Preparation of competent E.coli cells for
   transformation
   Transformation of E. coli strains DH5α and DH5αF'
   with recombinant CoIE1 plasmids and M13 phage
   DNA
   Screening of recombinant plasmids and phage
   Rapid plasmid isolation using alkaline lysis
   technique (Birnboim and Doly 1979)
   Preparation of M13 templates (Schreier et al.
   1979)
   Horizontal agarose gel electrophoresis
Sanger dideoxyribonucleotide DNA sequence analysis of the single- and double-stranded templates (Sanger et al. 1977)
Denaturing polyacrylamide sequencing gels (Sanger and Coulson 1978)
Autoradiography of sequencing gels
Preparation of dihydrodiols (Kunz and Chapman 1981b)
Preparation of cell-free extracts (Kunz and Chapman 1981b)
Modified Lowry procedure for determining protein concentration (Lowry et al. 1951)
Enzyme assays (Kunz and Chapman 1981b)
Site-directed mutagenesis

III. RESULTS ........................................ 86

Nucleotide sequence determination
DNA sequence analysis
Subcloning of pBK190 and expression of toluate-1,2-dioxygenase
Evolutionary possibilities of the TOL meta-cleavage operon
Site-directed mutagenesis of possible regulatory elements
Enzyme assays
Site-directed mutagenesis of terminator-like structure
Enzyme assays
Subcloning of pLTE(wt), pLTEmut6 and pLTEmut9 into cloning expression vector pSP73
Enzyme assays

IV. DISCUSSION...................................... 179

BIBLIOGRAPHY...................................... 201
LIST OF TABLES

1. Examples of degradative plasmids found within soil bacteria ........................................ Page 2
2. Examples of multicomponent hydroxylating dioxygenases ............................................. 23
3. Comparison of intergenic regions of the TOL meta-cleavage operon .......................... 120
4. Comparison of G+C content between xylXYZ and the downstream genes of the pDK1 TOL meta-cleavage operon ................................................................. 142
5. Nucleotide sequence of oligonucleotide primers used for site-directed mutagenesis ....... 147
6. Comparison of C230 activities expressed by E. coli strains carrying constructs pTE(wt), pTEmut3, pTEmut5 and pTEmut7 .............................................................. 160
7. Comparison of DHCDH assays expressed by E. coli strains carrying constructs pLTE(wt), pLTEmut6 and pLTEmut9 ................................................................. 170
8. Comparison of C230 assays expressed by E. coli strains carrying constructs pLTE(wt), pLTEmut6 and pLTEmut9 ................................................................. 171
9. Comparison of DHCDH activities expressed by E. coli strains carrying constructs pSLTE(wt), pSLTEmut6 and pSLTEmut9 ................................................. 177
10. Comparison of C230 activities expressed by E. coli strains carrying constructs pSLTE(wt), pSLTEmut6 and pSLTEmut9 ......................................................... 178
LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Illustration</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparison of physical maps for TOL plasmids pDK1, pWW0 and pWW53</td>
<td>5</td>
</tr>
<tr>
<td>2. Pathway for plasmid- and chromosomally-encoded degradation of toluene and related compounds</td>
<td>7</td>
</tr>
<tr>
<td>3. Current scheme governing the regulation of the TOL operon expression</td>
<td>14</td>
</tr>
<tr>
<td>4. Nucleotide sequence of the Pm promoter region from TOL plasmid pWW0</td>
<td>17</td>
</tr>
<tr>
<td>5. Subunit composition of the various classes of hydroxylating bacterial dioxygenases</td>
<td>24</td>
</tr>
<tr>
<td>6. Mechanism of the toluate-1,2-dioxygenase complex</td>
<td>26</td>
</tr>
<tr>
<td>7. Catabolic pathways for the degradation of toluene, phenol and naphthalene and the prospective genes encoding the required catabolic enzymes</td>
<td>29</td>
</tr>
<tr>
<td>8. Possible evolution of the naphthalene, toluene and phenol pathways</td>
<td>32</td>
</tr>
<tr>
<td>9. CsCl buoyancy gradient used to separate (chromosomal) DNA from closed circular supercoiled (plasmid) DNA</td>
<td>41</td>
</tr>
<tr>
<td>10. Screening grid for master plating of recombinant strains</td>
<td>53</td>
</tr>
<tr>
<td>11. Summary of the DNA sequencing reaction protocol</td>
<td>64</td>
</tr>
</tbody>
</table>
12. Sample standard curve used for the Lowry method of determining protein concentration

13. Enzymatic reactions catalyzed by DHCDH and C230

14. Transformer™ site-directed mutagenesis strategy

15. Genetic organization of the xyl genes found on the cointegrate pDK1 TOL plasmid pDKR1

16. Cloning strategy for the construction of subclones pBK190 and pBK789

17. DNA sequencing strategy and restriction endonuclease cleavage map of the xylXYZ region

18. Autoradiograph of a representative 6% denaturing polyacrylamide DNA sequencing gel

19. Nucleotide and predicted amino acid sequence of the xylXYZ region on TOL plasmid pDK1

20. Comparison of the pDK1 xylX nucleotide sequence with xylX (pWW0), benA and cbdA

21. Comparison of the pDK1 xylY nucleotide sequence with xylY (pWW0), benB and cbdB

22. Comparison of the pDK1 xylZ nucleotide sequence with xylZ (pWW0), benC and cBdC

23. Comparison of the pDK1 XylX amino acid sequence with XylX (pWW0), BenA and CbdA

24. Comparison of the pDK1 XylY amino acid sequence with XylY (pWW0), BenB and CbdB

25. Comparison of the pDK1 XylZ amino acid sequence with XylZ (pWW0), BenC and CbdC

26. Possible nucleotide deletions within xylX of pWW0
27. Putative transcriptional terminator found immediately downstream of xylZ

28. Nucleotide sequence comparison of the meta-cleavage Pm promoter between TOL plasmids pDK1 and pWWO

29. Comparison of the nucleotide sequence upstream of the Pm promoter from pDK1 to a Pseudomonas putida nucleotide sequence of unknown function

30. Diagrammatic representation of the strategy used to construct pBK192, pBK292 and pBK392

31. Relative levels of toluate-1,2-dioxygenase activity from pBK190, pBK192, pBK292 and pBK392 constructs

32. Putative stem and loop structures found in the xylZxylL and xylLxylT intergenic regions

33. Diagrammatic representation of TOL constructs pBK692, pBK189 and pBK591, and their expressed levels of C230 activity

34. Nucleotide and predicted amino acid sequence of xylT

35. Agarose gel showing sensitivity of mutant clones to BglII restriction endonuclease digestion

36. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence changes for pTEmut3 and pTEmut5

37. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence change for pTEmut7

38. Putative xylLxylT intergenic region stem and loop structure and the calculated free-energy predictions for wild-type and mutant constructs
39. Representative C230 assay

40. Putative xylZxylL intergenic region stem and loop structure and the calculated free-energy predictions for wild-type and mutant subclones

41. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence change for pLTEmut6 and pLTEmut9

42. Representative DHCDH assay

43. Diagrammatic representation of the cloning strategy used to construct pSLTE(wt), pSLTEmut6 and pSLTEmut9

44. Phylogenetic analysis of oxygenase components (α subunits) (A) and reductase components (B)
## 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>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>Amp&lt;sup&gt;r&lt;/sup&gt;</td>
<td>ampicillin (50 milligrams per milliliter)</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BAP</td>
<td>bacterial alkaline phosphatase</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Centigrade</td>
</tr>
<tr>
<td>Cb&lt;sup&gt;r&lt;/sup&gt;</td>
<td>carbenicillin resistance</td>
</tr>
<tr>
<td>cc</td>
<td>cubic centimeter</td>
</tr>
<tr>
<td>cm</td>
<td>centimeter</td>
</tr>
<tr>
<td>C230</td>
<td>catechol-2,3-dioxygenase</td>
</tr>
<tr>
<td>Da</td>
<td>dalton</td>
</tr>
<tr>
<td>DHCDH</td>
<td>1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase</td>
</tr>
<tr>
<td>ETOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>IR</td>
<td>intergenic region</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kbp</td>
<td>kilobase pair</td>
</tr>
<tr>
<td>kcal</td>
<td>kilocalories</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>Kn&lt;sup&gt;r&lt;/sup&gt;</td>
<td>kanamycin resistance</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>l</td>
<td>liter</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>mol</td>
<td>mole</td>
</tr>
<tr>
<td>m-tol</td>
<td>meta-toluate</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>μl</td>
<td>microliter</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>Tc&lt;sup&gt;r&lt;/sup&gt;</td>
<td>tetracycline resistance</td>
</tr>
<tr>
<td>TO</td>
<td>toluate-1,2-dioxygenase</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>U</td>
<td>uracil</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
<tr>
<td>xylZ*</td>
<td>partial segment of xylZ gene</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The genus *Pseudomonas* is comprised of many different organisms and representatives of the genus can generally be found in nearly all natural ecosystems (Haas 1983). Pseudomonads are considered to be metabolically diverse organisms, having the ability to utilize a large variety of compounds as carbon and energy sources (Chakrabarty 1976). The catabolic enzymes responsible for this phenomenon may be encoded either chromosomally or extrachromosomally (e.g. on transmissible plasmids) within the organism (Williams and Worsey 1976). The first biodegradative plasmid discovered in *Pseudomonas* was the SAL plasmid, which encodes the genes necessary for the degradation of salicylate (Chakrabarty 1972). Since then many other catabolic plasmids have been discovered and studied to greater or lesser extents. Several examples of plasmids encoding catabolic pathways are listed in Table 1.

A number of toluene degrading (TOL) plasmids have been isolated independently, and three of them have been extensively studied with regard to the locations and organization of their catabolic genes and regulatory elements. These plasmids are the archetype pWWO plasmid (Worsey and Williams 1975; Nakazawa et al. 1978), the pWW53
Table 1. Examples of degradative plasmids found within soil bacteria (adapted from Haas, 1983).

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Growth Substrate</th>
<th>Principle metabolic product(s)</th>
<th>Original Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOL</td>
<td>Toluene, m-xylene, p-xylene</td>
<td>Pyruvate + acetaldehyde, Pyruvate + prionaldehyde (meta-cleavage)</td>
<td>P. putida</td>
</tr>
<tr>
<td>(pWW0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAH/C</td>
<td>Naphthalene</td>
<td>Salicylate -&gt; pyruvate + acetaldehyde</td>
<td>P. putida</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(meta-cleavage)</td>
<td></td>
</tr>
<tr>
<td>pWW60-1C</td>
<td>Naphthalene</td>
<td>Salicylate -&gt; pyruvate + acetyl-CoA (ortho-cleavage)</td>
<td>P. putida</td>
</tr>
<tr>
<td>SAL; pMWD1</td>
<td>Salicylate</td>
<td>Pyruvate + acetaldehyde (meta-cleavage)</td>
<td>P. putida</td>
</tr>
<tr>
<td>pND50</td>
<td>p-Cresol</td>
<td>p-Hydroxybenzoate -&gt; succinate + acetyl-CoA (ortho-cleavage)</td>
<td>P. putida</td>
</tr>
<tr>
<td>CAM</td>
<td>Camphor</td>
<td>Isobutyrate</td>
<td>P. putida</td>
</tr>
<tr>
<td>OCT</td>
<td>n-Octane (C5-C10 alkane)</td>
<td>Alkane aldehyde</td>
<td>P. putida</td>
</tr>
<tr>
<td>pWW17</td>
<td>Phenylacetate</td>
<td>Not Known</td>
<td>P. putida</td>
</tr>
<tr>
<td>'pB13'</td>
<td>3-Chlorobenzoate</td>
<td>Succinate + acetyl-CoA (ortho-cleavage)</td>
<td>P. sp. B13</td>
</tr>
<tr>
<td>pAC25</td>
<td></td>
<td></td>
<td>P. putida</td>
</tr>
<tr>
<td>pRF1</td>
<td>4-Chlorobiphenyl</td>
<td>4-Chlorobenzoate</td>
<td>Acinetobacter sp.P6</td>
</tr>
<tr>
<td>pCSI</td>
<td>Parathion</td>
<td>p-Nitrophenol + diethylthiophosphate</td>
<td>P. diminuta</td>
</tr>
</tbody>
</table>

a The most extensively studied TOL plasmid
b Cleavage for catechol intermediate
c NAH7 and pWW60-1 are examples of naphthalene degradative plasmids.
plasmid (Keil et al. 1987) and the pDK1 plasmid (derivatives of this one used for this study, Shaw and Williams 1988). The plasmid of interest for this study was discovered in Pseudomonas putida HS1 (pPCl) and encodes the genes required for the degradation of toluene, meta- and para-xylene, 1,3,4-trimethyl benzene and 3-ethyltoluene to pyruvate and the corresponding aldehydes (Kunz and Chapman 1981b, Shaw and Williams 1988). This plasmid is a 125-kbp (82.5 MDa), non-transmissible plasmid and was given the name pDK1. The plasmid pDKR1 is a cointegrate derived from pDK1 and the broad host range resistance plasmid RP1. RP1 confers transmissibility and antibiotic resistance to carbenicillin, tetracycline and kanamycin to the plasmid. For this study, recombinant constructs derived from the TOL region of the plasmid pDKR1 were used.

**Organization of TOL pathway genes**

The genes of the TOL pathway encoded by pDK1 are organized into two metabolic operons. The 'upper pathway' operon contains the genes $\text{xylUWCMABN}$ (Harayama et al. 1989, Williams et al. 1997), that are responsible for the conversion of a methyl substituent of the aromatic substrate to a carboxylate group. The meta-cleavage operon is comprised of the genes $\text{xylXYZLTEGFJQKIH}$ (Harayama and Rekik 1990), that are responsible for the conversion of the carboxylic acids to the common metabolites pyruvate and an aldehyde. These two operons are separated by a segment of
DNA encoding the two regulatory genes \textit{xylR} and \textit{xylS}. The products of these genes are both positive acting transcriptional regulators. The XylS protein regulates the expression of the meta-cleavage operon, while XylR regulates the 'upper pathway' operon and \textit{xylS} expression. Maps showing the organization and location of each of these pDK1 genes and operons, along with similar data for the related TOL plasmids pWW0 and pWW53, are presented in Figure 1. Figure 2 shows the TOL encoded catabolic pathway, as well as the related chromosomally-encoded pathway for degradation of aromatic compounds.

The initial enzyme in the 'upper' pathway is a multicomponent oxygenase composed of two proteins encoded by \textit{xylMA}, xylene monooxygenase (XO). XylM functions as the electron-transfer component where XylA functions as a terminal hydroxylase (Suzuki et al. 1991). Xylene monooxygenase is responsible for the conversion of toluene to benzyl alcohol. Benzyl alcohol is then attacked by benzylalcohol dehydrogenase (BADH, encoded by \textit{xylB}), transforming benzyl alcohol to benzaldehyde. The final step in the 'upper' pathway is the conversion of benzaldehyde to benzoate by benzaldehyde dehydrogenase (BZDH, encoded by \textit{xylC}). Currently, three genes in the 'upper' pathway have unknown functions and are not required for the conversion of toluene to benzoate: \textit{xylU}, \textit{xylW} and \textit{xylN} (Harayama et al. 1989 and Williams et al. 1997). The XylU protein joins the many protein sequences in the Genbank database with unknown
**Fig. 1.** Comparison of physical maps for TOL plasmids pDK1, pWW0 and pWW53. The arrows designate the direction of transcription. Upper refers to the upper pathway genes (xylUWCMABN), whereas meta, meta1 and meta2 refer to the meta-pathway genes (xylXYZLTFIJEKQI). pWW53 contains two identical copies of the meta-cleavage operon (meta1 and meta2). Letters represent genes from the TOL pathways. xylS2pDK1 and xylS2pWW53 are both nonfunctional truncated pseudogenes. pWW53 contains two functional xylS genes, xylS1 and xylS3. This figure was adapted from Assinder et al. (1993).
Fig. 2. Pathway for plasmid- and chromosomally-encoded degradation of toluene and related compounds. This figure was adapted from Harayama et al. (1987a) and Ramos et al. (1997).
functions (Williams et al. 1997). The XylW protein has strong homology to other long-chain Zn-containing alcohol dehydrogenases, but no NAD$^+$-dependent dehydrogenase activity was observed against benzyl alcohol, mandelate, or benzylamine (Williams et al. 1997). Finally, the XylN protein has been postulated as a probable toluene-specific outer membrane porin (Genbank accession #1841317).

Benzoate is a substrate for the meta-cleavage operon. Benzoate is converted to 1,2-dihydrocyclohexa-3,5-diene carboxylate by toluate-1,2-dioxygenase, a multicomponent dioxygenase encoded by xylXYZ. XylX and XylY together serve as the oxygenase component, whereas XylZ serves as a reductase (Harayama et al. 1991a). The 1,2-dihydrocyclohexa-3,5-diene carboxylate is converted to catechol by 1,2-dihydrocyclohexa-3,5-diene carboxylate dehydrogenase (DHCDH, encoded by xylL). Catechol is next cleaved by catechol-2,3-dioxygenase (C230, encoded by xylE), forming 2-hydroxymuconic semialdehyde. This compound is bright yellow and can be measured spectrophotometrically. Depending upon the substitution of the initial compound, the ring fission products are further metabolized by the hydrolytic branch or the oxalocrotonate branch of the pathway. Para-substituted and unsubstituted derivatives like benzoate, toluene, and para-toluene are metabolized by the oxalocrotonate branch. The oxalocrotonate branch is actually the more energetically favorable of the two due to the formation of NADH (Harayama et al. 1987a). In this branch,
2-hydroxymuconic semialdehyde dehydrogenase (HMSD, encoded by xylG) takes 2-hydroxymuconic semialdehyde and converts it to its enol form, 4-oxalocrotonate. Equilibrium conversion between the keto and enol form is completed by 4-oxalocrotonate tautomerase (OT, encoded by xylH). The keto form is then attacked by 4-oxalocrotonate decarboxylase (OD, encoded by xylI), leaving 4-oxapenta-4-enoate. Next, 4-oxapenta-4-enoate hydratase (DEH, encoded by xylJ) converts 4-oxapenta-4-enoate to 4-hydroxy-2-oxovalerate. The xylK gene product, 4-hydroxy-2-oxovalerate aldolase (HOA) converts 4-hydroxy-2-oxovalerate to pyruvate and acetaldehyde. Finally, acetaldehyde dehydrogenase (an acylating enzyme, encoded by xylQ) adds CoA to acetaldehyde, forming acetyl-CoA. These two products, pyruvate and acetyl-CoA, can now be shuttled into the Kreb's cycle as intermediates.

Meta-substituted and ortho-substituted derivatives are metabolized by the hydrolytic branch. Here 2-hydroxymuconic semialdehyde is converted to 4-oxapenta-4-enoate by 2-hydroxymuconic semialdehyde hydrolase (HMSH, encoded by xylF). Acetate is released as a byproduct as the result of this action. Next, 4-oxapenta-4-enoate hydratase (DEH, encoded by xylJ) converts 4-oxapenta-4-enoate to 4-hydroxy-2-oxovalerate. The xylK gene product, 4-hydroxy-2-oxovalerate aldolase (HOA) converts 4-hydroxy-2-oxovalerate to pyruvate and an aldehyde. The use of these two branches (hydrolytic and oxalocrotonate) allows the meta-cleavage pathway to accommodate a wider variety of closely related compounds.
One of the 13 genes comprising the meta-cleavage operon, xylT, was initially thought to be unnecessary for the conversion of toluene and related xylenes to Kreb's cycle intermediates. The xylT gene product is closely related to chloroplast-type ferredoxins (Harayama et al. 1991b). XylT may actually be required for the regeneration of catechol-2,3-dioxygenase (C23O, encoded by xylE) when grown in the presence of p-toluate (Polissi and Harayama 1993). It was found that C23O activity in xylT minus cells was 1% of the activity as compared to wild-type cells when the meta-cleavage operon was induced by p-toluate, whereas all other meta-cleavage enzymes were at normal levels (Polissi and Harayama 1993). XylT is therefore another example of how substrate specificity is expanded for these pathways.

Ortho-cleavage pathway

It is possible for a strain such as in Pseudomonas putida mt-2 to contain both the chromosomally-encoded ortho-pathway and the plasmid-encoded meta-pathway (Nakazawa and Yokota 1973). Benzoate is first converted to 2-hydro-1,2-dihydroxybenzoate by benzoate-1,2-dioxygenase (encoded by benABC). This diol is then converted to catechol by 2-hydro-1,2-dihydroxybenzoate dehydrogenase (encoded by benD). Benzoate is metabolized by the meta-cleavage pathway unless a functional meta-cleavage pathway is not available (Nakazawa and Yokota 1973). The ortho-cleavage pathway (also called the β-keto adipate pathway) is responsible for the conversion
of catechol to the Kreb’s cycle intermediates acetyl-CoA and succinate (Fig. 2). Catechol is converted to \textit{cis,cis}-muconate by catechol-1,2-dioxygenase (encoded by \textit{catA}). This product is further metabolized by \textit{cis,cis}-muconate lactonizing enzyme (MLE, encoded by \textit{catB}) to muconolactone, which is subsequently converted to \(\beta\)-ketoadipate by muconolactone isomerase (\textit{catC}) and \(\beta\)-ketoadipate enol-lactone hydrolase (\textit{catD}). The \(\beta\)-ketoadipate is then metabolized by \(\beta\)-ketoadipate succinyl-coenzyme A transferase (\textit{catE}) to \(\beta\)-ketoadipyl CoA and then to Kreb’s cycle intermediates acetyl CoA and succinate.

**Regulation of the TOL pathway**

There are many aspects to the regulation of the TOL operons. As described earlier, the genes encoding the TOL pathway of pDK1 and other TOL plasmids are organized into two operons, the ‘upper operon’ and the ‘meta-cleavage operon’. In addition to these two operons, there are two associated regulatory genes, \textit{xylR} and \textit{xylS}, whose products are involved in the transcriptional control of the pathway. Each of the regulatory genes is transcribed from its own promoter. The promoters for \textit{xylR}, \textit{xylS}, the upper operon and the \textit{meta}-cleavage operon are referred to as \textit{Pr}, \textit{Ps}, \textit{Pu} and \textit{Pm}, respectively. The control of the two operons accomplished by two regulatory loops, the \textit{meta} loop (growth on toluates) and the cascade loop (growth on xylones, ensures that both upper and \textit{meta} operons are expressed) (Ramos et al. 1997).
current model for the transcriptional control of expression is shown in Figure 3.

Growth on toluates (alkylbenzoates) requires only the expression of the meta-cleavage operon (the meta loop). The XylS protein, which stimulates the transcription of the meta-cleavage pathway, is activated by the binding of a benzoate effector molecule (Inouye et al. 1981). The nucleotide sequence of xylS (pWWO) was determined and it was found to be 963 bp in length encoding a 321 aa protein product of 36502 Da (Inouye et al. 1986). It belongs to the XylS/AraC family of transcriptional regulators which are involved in transcriptional stimulation of such processes as carbon metabolism and virulence factors (Ramos et al. 1990, Gallegos et al. 1993). Low constitutive levels of inactive XylS are produced under the control of a σ70-dependent promoter called Ps2 (Gallegos et al. 1996a). The addition of benzoate (meta-cleavage substrate) causes an immediate expression of the meta-cleavage operon from the Pm promoter, thereby suggesting that preexisting XylS regulator protein becomes active with effector binding (Marques et al. 1994). Approximately 80 bp upstream from the transcriptional initiation point are needed for the activation of Pm (Kessler et al. 1993). Two protein binding regions have been identified within this area: a XylS binding region (extending from -40 to -80 bp) and an RNA polymerase binding region, which displays divergent -35 and -10 DNA sequences (Mermod et al. 1984, Inouye et al. 1984). Figure 4 shows the nucleotide sequence of the Pm promoter.
**Fig. 3.** Current scheme governing the regulation of the TOL operon expression. Filled arrows and (+) signs demonstrate stimulation of transcription. Open arrows and (-) signs demonstrate inhibition of transcription. The dashed arrow represents stimulation of transcription due to overproduction of XylS in the absence of effector binding. The DNA-bending proteins integration host factor and protein HU are represented by IHF and HU respectively. This figure was adapted from Kessler et al. (1994) and Ramos et al. (1997).
region of pWW0 with the proposed binding regions.
Interestingly, Marques et al. (1995) offered an explanation for the divergent -35 and -10 RNA polymerase binding region. They discovered that high levels of transcription from Pm in a strain lacking the sigma-S factor (\(\sigma^S\)) only occurred during the early logarithmic phase of growth, while a \(\sigma^S\) capable strain showed high levels of transcription throughout the normal growth curve. Thus it seems that \(\sigma^{70}\) maintains high level expression from Pm during early exponential phase while \(\sigma^S\) maintains the expression during the late exponential and stationary phases of growth. The divergence from normal -35 and -10 sequences may be due to the recognition requirement for both \(\sigma^{70}\) and \(\sigma^S\) within the Pm promoter region.

Regulation of the cascade loop is more complex. Transcriptional regulation of the cascade loop revolves around the xylR gene product. The need here is to get both the upper and meta-cleavage operons expressed. XylR (pWW0) is constitutively expressed from two tandem overlapping \(\sigma^{70}\)-dependent promoters, where XylR controls its own synthesis (Inouye et al. 1985, Gomada et al. 1992). The 1698 bp nucleotide sequence of xylR was determined to encode a 566 aa protein product of 63741 Da (Inouye et al. 1988). It belongs to the NtrC/NifA family of transcriptional regulators which are best known for their involvement in nitrogen metabolism (Inouye et al. 1988). XylR becomes activated upon binding of
5' AGGGATAAAGITGCAAGAAGCGGATACAAGTCGAAAAATGGCTA1CTCTTAGAAAGCCCTAC

3' CCTTTACGCTTTATGCAAACAGAAACATATAATAATGGACATGACCATTG

pww0

XbaI

RBS

xylX-->
Fig. 4. Nucleotide sequence of the Pm promoter region from TOL plasmid pWW0. The boxed areas are the TACAN$_4$TGCA motifs upstream of the -35 region found to be the minimum DNA segment required for transcriptional stimulation by XylS (Gallegos et al. 1996b). The arrows designate the direct repeats in the operator sequence where XylS binds along one side of the DNA covering four helical turns specifically demonstrated by hydroxyl radical footprinting and methylation interference assays (Kaldalu et al. 1996). As seen from the sequence, the -10 box is TAGGCT instead of the consensus TATAAT. The -35 sequence CTATCT also differs substantially from the consensus TTGACA sequence. This figure was adapted from Gallegos et al. (1996b) and Kessler et al. (1993).
an effector molecule; for example, toluene, xylenes, benzyl alcohol, alkylbenzyl alcohols, p-chlorobenzaldehyde, m-amino tolune, o-nitrotoluene, or p-nitrotoluene (Abril et al. 1989, Delgado and Ramos 1994). Once activated, XylR stimulates transcription of the xylUWCMABN genes (upper operon) from promoter Pu and simultaneously activates the xylS gene from promoter Psl. Stimulation of Pu also requires the DNA bending protein IHF and a $\sigma^{54}$-dependent RNA polymerase (Kohler et al. 1989, Holtel et al. 1990, de Lorenzo et al. 1991). XylR mediated transcription of xylS occurs at a different promoter than the $\sigma^{70}$-dependent Ps2 described earlier in the meta loop. Here XylR stimulates transcription from the Psl promoter, requiring a $\sigma^{54}$-dependent RNA polymerase and the DNA-bending protein HU (Perez-Martin and de Lorenzo 1995). Therefore, when cells are grown on xylenes, xylS is expressed from two promoters: the XylR-independent $\sigma^{70}$ constitutive Ps2 promoter and the $\sigma^{54}$-XylR-dependent Psl promoter. Expression from both promoters causes the XylS protein to be hyperproduced, which in turn elicits high expression from the Pm promoter even in the absence of meta pathway inducers (Inouye et al. 1987, Gallegos et al. 1996a). Apparently a portion of XylS becomes activated under these conditions and allows the expression of the meta-cleavage operon, even though no lower operon effector is yet present. This allows for the harmonious
induction of both the upper and meta pathways without requiring the accumulation of metabolic intermediate products to act as meta operon inducers (Ramos et al. 1987).

The TOL plasmids provide Pseudomonas with the ability to use a wider variety of carbon sources, but at a rather large energy expense. The synthesis of 22 peptides is elicited in response to an aromatic inducer such as toluene. Thus, the expression of this system can actually be a hindrance to the organism if other more easily utilized carbon sources are available. Several studies have produced evidence for catabolite repression of the TOL pathway operons (Hugouvieux-Cotte-Pattat et al. 1990, Holtel et al. 1994, Duetz et al. 1994, Marques et al. 1994). Marques et al. (1994) showed that the induction of the TOL pathway enzymes was slow (hours) following effector introduction when cells were grown in rich medium. This compared to immediate induction when cells were grown on minimal medium containing the aromatic compound as the only carbon source. Pseudomonads do not abide by the classical enteric system of catabolite repression involving cAMP (Holtel et al. 1994). The role of σ^{54} proves to be an interesting alternative. Cases et al. (1996) demonstrated that an overproduction of σ^{54} slightly attenuated the lack of expression from Pu in cells growing in rich medium. Much work remains to be done in order to further the understanding of catabolite repression and its effects on TOL operon expression.
The *xylXYZ* genes and toluate-1,2-dioxygenase

The first three genes of the *meta*-cleavage operon, *xylXYZ*, encode the three subunit enzyme toluate-1,2-dioxygenase (TO). When toluene is acted upon by upper pathway enzymes, the methyl group is converted to a carboxylic acid. This product is a substrate for TO, which converts this carboxylic acid to a cis-carboxylic diol. A dioxygenase is an enzyme which inserts two atoms of oxygen into its substrate. An important function of dioxygenases in biological processes is the cleavage of aromatic rings. The first dioxygenase discovered was responsible for the conversion of anthranilic acid to catechol (Kobayshi et al. 1964). Since this time, a number of other dioxygenases have been discovered and studied. There are two basic types of dioxygenases, those involved in ring fission and those involved in ring hydroxylation. Toluate-1,2-dioxygenase (encoded by *xylXYZ*) is involved in ring hydroxylation.

Hydroxylating dioxygenases are multicomponent enzymes that utilize NADH and O$_2$ for the degradation of aromatic compounds. With the advent of DNA sequence analysis and other recombinant DNA techniques, the fine structure of numerous dioxygenase enzymes has been studied. Recently, Batie et al. (1991) classified hydroxylating dioxygenases by the nature of their electron transfer systems. Class I dioxygenases are two-component enzymes composed of a reductase, which contains both a flavin and a [2Fe-2S] redox center, and an oxygenase. Class II dioxygenases are three-
component enzymes where components of the reductase occur on separate proteins, a flavoprotein (flavin) and a ferredoxin ([2Fe-2S]). Class III dioxygenases are also three-component enzymes, but they also require a second [2Fe-2S] center located on the ferredoxin. Subdivision within the classes is based upon the number of terminal oxygenase subunits, placement of the [2Fe-2S] in the ferredoxin, and the whether the flavin utilizes FAD or FMN in the reductase. Examples of the different dioxygenase classes are described in Table 2 and Figure 5.

Toluate 1,2-dioxygenase in *Pseudomonas putida* mt-2 (encoded by the TOL plasmid pWWO) is composed of three different polypeptides, 57, 20 and 39 kDa, encoded by *xylXYZ*, respectively (Harayama and Rekik 1990). The *xylXY* products are responsible for the oxygenase reaction (hydroxylase component) and each is present at three copies per enzyme. *XylZ* is an NADH-cytochrome C reductase (electron transfer component) and is present at one copy per molecule. A diagrammatic example of the toluate-1,2-dioxygenase reaction can be seen in Figure 6. TO has been studied extensively and is quite similar in structure and function to the chromosomally-encoded benzoate dioxygenase system (Harayama et al. 1991a, Harayama and Rekik 1990, Harayama et al. 1986, Neidle et al. 1987, Neidle et al. 1991). The nucleotide sequences of the genes encoding the pWWO TOL plasmid TO and the benzoate dioxygenase system exhibit greater than 50% homology (Harayama et al. 1991a). Although the two enzymes
Table 2. Examples of multicomponent hydroxylating dioxygenases separated by Classes (adapted from Mason and Cammack, 1992).

<table>
<thead>
<tr>
<th>Enzyme System</th>
<th>Components</th>
<th>Prosthetic group(s)</th>
<th>Gene Designation Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phthalate dioxygenase (P. cepacia)</td>
<td>Reductase</td>
<td>34,000</td>
<td>P450, [2FeS-2S]</td>
</tr>
<tr>
<td></td>
<td>Oxygenase</td>
<td>α(44,000)</td>
<td>4[2FeS-2S]R, 4Fe</td>
</tr>
<tr>
<td>4-chlorophenylacetate 3,4-dioxygenase (P. sp. CBS)</td>
<td>Reductase</td>
<td>34,000</td>
<td>3[2Fe-2S]</td>
</tr>
<tr>
<td></td>
<td>Oxygenase</td>
<td>α(46,000)</td>
<td></td>
</tr>
<tr>
<td>Benzoate dioxygenase (P. arvilla C-1, A. calcoaceticus)</td>
<td>Reductase</td>
<td>37,500</td>
<td>FAD, 2Fe-2S benC (xy12)</td>
</tr>
<tr>
<td></td>
<td>Oxygenase</td>
<td>α(50,000)</td>
<td>3[2Fe-2S]R, 3Fe benAB (xy1XY)</td>
</tr>
<tr>
<td>4-methoxybenzoate O-demethylase (P. putida DSM No. 1868)</td>
<td>Reductase</td>
<td>42,000</td>
<td>P450, [2Fe-2S]</td>
</tr>
<tr>
<td></td>
<td>Paradiamonoxin</td>
<td>α(41,000)</td>
<td>[2Fe-2S], Fe</td>
</tr>
<tr>
<td>Benzene 1,2-dioxygenase (P. putida NL2 NCIB 12130)</td>
<td>Reductase&lt;sub&gt;RED&lt;/sub&gt;</td>
<td>α(42,000)</td>
<td>FAD</td>
</tr>
<tr>
<td></td>
<td>Ferredoxin&lt;sub&gt;RED&lt;/sub&gt;</td>
<td>11,850</td>
<td>[2Fe-2S]R</td>
</tr>
<tr>
<td></td>
<td>ISP&lt;sub&gt;RED&lt;/sub&gt;</td>
<td>α(54,500)</td>
<td>2[2Fe-2S]R, Fe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β(23,500)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyrazon dioxygenase (Pseudomonas)</td>
<td>Reductase</td>
<td>67,000</td>
<td>FAD</td>
</tr>
<tr>
<td></td>
<td>Ferredoxin</td>
<td>12,000</td>
<td>[2Fe-2S]</td>
</tr>
<tr>
<td></td>
<td>Oxygenase</td>
<td>180,000</td>
<td>[2Fe-2S]</td>
</tr>
<tr>
<td>Toluene dioxygenase (P. putida F1)</td>
<td>Reductase&lt;sub&gt;TOL&lt;/sub&gt;</td>
<td>46,900</td>
<td>FAD</td>
</tr>
<tr>
<td></td>
<td>Ferredoxin&lt;sub&gt;TOL&lt;/sub&gt;</td>
<td>11,900</td>
<td>[2Fe-2S]R</td>
</tr>
<tr>
<td></td>
<td>ISP&lt;sub&gt;TOL&lt;/sub&gt;</td>
<td>α(52,500)</td>
<td>2[2Fe-2S]R, Fe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β(30,800)</td>
<td></td>
</tr>
<tr>
<td>Naphthalene dioxygenase (P. putida NCIB 9816)</td>
<td>Reductase&lt;sub&gt;NAP&lt;/sub&gt;</td>
<td>36,300</td>
<td>FAD, [2Fe-2S]</td>
</tr>
<tr>
<td></td>
<td>Ferredoxin&lt;sub&gt;NAP&lt;/sub&gt;</td>
<td>15,300</td>
<td>[2Fe-2S]</td>
</tr>
<tr>
<td></td>
<td>ISP&lt;sub&gt;NAP&lt;/sub&gt;</td>
<td>α(55,000)</td>
<td>2[2Fe-2S], 2Fe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β(20,000)</td>
<td></td>
</tr>
</tbody>
</table>
Electron transport chain

<table>
<thead>
<tr>
<th>Class</th>
<th>Oxygenase</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td></td>
<td>4-sulphobenzoate 3,4-dioxygenase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4-chlorophenylacetate 3,4-dioxygenase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>phthalate 4,5-dioxygenase</td>
</tr>
<tr>
<td>IB</td>
<td>2-oxo-1,2-hydroquinoline 8-monoxygenase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>benzoate 1,2-dioxygenase toluate 1,2-dioxygenase 2-halobenzoate 1,2-dioxygenase</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>dibenzofuran 4,4a-dioxygenase</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>benzene 1,2-dioxygenase</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>naphthalene 1,2-dioxygenase</td>
<td></td>
</tr>
</tbody>
</table>

- FMN, [2Fe-2S] Reductase
- FAD, [2Fe-2S] Reductase
- FAD Reductase
- [2Fe-2S] Ferredoxin
- [2Fe-2S] Fe(II) Oxygenase (ISPα)
- [2Fe-2S] R, Fe(II) Oxygenase (ISPβ)
Fig. 5. Subunit composition of the various classes of hydroxylating bacterial dioxygenases. This figure was adapted from Butler and Mason (1997).
Toluene-1,2-dioxygenase

\[ \text{Benzoate} \]

\[ \text{hydroxylase component } \text{xy}lXY \]

\[ \text{electron transfer component } \text{xy}lZ \]

\[ \text{1,2-dihydrocyclohexa-3,5-diene carboxylate} \]
Fig. 6. Mechanism of the toluate-1,2-dioxygenase complex. Toluate-1,2-dioxygenase (TO) initiates benzoate degradation by converting it to 1,2-dihydroxycyclohexa-3,5-diene carboxylate. TO is composed of a two-subunit hydroxylase (encoded by xylXY) and an electron transfer oxidoreductase (encoded by xylZ). This figure was adapted from Neidle et al. (1991).
seem to have a number of similarities, benzoate dioxygenase has a limited substrate range as compared to the broad substrate specificity of TO. More recently, TO was also found to be similar (~50% amino acid sequence homology) to the plasmid-encoded 2-halobenzoate 1,2-dioxygenase system from *Burkholderia cepacia* (formerly *Pseudomonas*) (Haak et al. 1995). All three of these dioxygenases are considered to be members of Class IB dioxygenases as set forth by Batie et al. (1991).

**Evolution of the TOL meta-cleavage operon**

Comparisons by Harayama and Rekik (1990) of the TOL pathway operons to other catabolic pathways allowed them to suggest that the meta-cleavage operon is the product of a fusion between two DNA modules. These transcription units would have been *xylXYZL* and *xylTEGFJQKIH*, where each unit originally possessed its own promoter/operator regions. The idea of creating a new pathway through the fusion of two separate operons can also be envisioned for other related pathways, for example, the NAH and Phenol degradation pathways. Catechol is an important intermediary metabolite in the TOL, phenol and the naphthalene (NAH) pathways (Platt et al. 1995, Patel and Barnsley 1980, Patel and Gibson 1974, Ensley and Gibson 1983, Yen and Gunsalus 1982, Shingler et al. 1992). As seen in Figure 7, the three pathways merge at catechol. Because of this structural relationship between the three pathways, it is possible to postulate that the *xylXYZL* genes may have become part of the "modern" TOL meta-
Fig. 7. Catabolic pathways for the degradation of toluene, phenol and naphthalene and the prospective genes encoding the required catabolic enzymes. Catechol (shown in bold) is the common intermediate in several aromatic degradative pathways. The complete pathways for degradation of naphthalene, phenol and toluene are shown. Each pathway contains its own unique set of enzymes allowing for the conversion naphthalene, phenol, or toluene to catechol. The remaining enzymes responsible for the degradation of catechol to Kreb's cycle intermediates are homologous between the three pathways. This figure was adapted from Harayama and Rekik (1990), Platt et al. (1995) and Shingler et al. (1992).
cleavage operon through recombination with an early
xylTEGFJQKIHX meta-cleavage unit. The NAH meta-cleavage
operon contains the genes nahTHINLOMKJ and the phenol meta-
cleavage operon contains the genes dmpQBCDEPCHI; both which
correspond to the xylTEGFJQKIHX genes of the TOL meta-cleavage
operon (Yen and Gunsalus 1982, Platt et al. 1995, Shingler et
al. 1992). The complete NAH meta-cleavage operon could then
have been constructed from the same early xylTEGFJQKIHX operon
and the nahG gene needed for the merging of the two pathways
at catechol. Likewise, the entire phenol operon could have
been constructed from a similar xylTEGFJQKIHX operon by
merging with dmpKLMNOP (phenol hydroxylase). A diagrammatic
representation of the evolutionary relationships is shown in
Figure 8.

Genetic engineering and biodegradative pathways

Ever since the discovery that some microorganisms have
the unique ability to detoxify aromatic compounds, much work
has been done to further understand the processes involved.
The biodegradative capability of a particular strain is
commonly restricted to a single class of chemical compounds.
Since most oil spills and chemical waste sites are usually
not a composed of just one toxic compound, it might be useful
to expand substrate specificity of existing biodegradative
pathways or to genetically engineer new pathways to handle a
wider variety of substrates.

Reineke and Knackmuss (1980) demonstrated that 3-
Ancestral Transcriptional Units

Modern Transcriptional Units
**Fig. 8.** Possible evolution of the naphthalene, toluene and phenol pathways.
chlorobenzoate degradation in *Pseudomonas* sp. B13 was expanded by the addition of the entire TOL plasmid from *Pseudomonas putida* mt-2. As a direct result, the newly engineered strain was able to utilize 4-chlorobenzoate and 3,5-dichlorobenzoate. It was later discovered that the entire TOL pathway was not necessary for the substrate expansion. In fact, the same result seen by the addition of the entire TOL plasmid was accomplished by the addition of a cloned segment containing only the *xylXYZL* genes (Lehrbach et al. 1984). The degradative capability was further increased to include salicylate and 3-, 4-, and 5-chlorosalicylate by the addition of *nahG* (Lehrbach et al. 1984). Another example of combining pathways to expand substrate specificity was described by Lee et al. (1995). They combined the chromosomally encoded *tol* (not TOL) and *tod* pathways into one organism that successfully degraded the typical petroleum components benzene, toluene and p-xylene. The *tol* pathway cannot recognize benzene as a substrate and the *tod* pathway cannot completely metabolize p-xylene. A hybrid strain was created by cloning the *todC1C2BA* genes encoding toluene dioxygenase and inserting them into *Pseudomonas putida* mt-2. This strain was found to mineralize a benzene, toluene and p-xylene mixture without the accumulation of any metabolic intermediates (Lee et al. 1995).

In order to enhance the capabilities for bioremediation, a further understanding of biodegradative pathways is needed at the genetic and biochemical level. Toluate-1,2-
dioxygenase (encoded by xylXYZ) is the first enzyme of the TOL meta-cleavage pathway and is partially responsible for the broad substrate specificity demonstrated by the TOL degradative pathway. The effort of this research is to further characterize the xylXYZ genes found on the pDK1 TOL plasmid and its associated downstream regulatory elements.
CHAPTER II

MATERIALS and METHODS

Bacterial strains and plasmids

*Escherichia coli* DH5α (supE44, hsdR17, recA1, endA1, gyrA96, thi-1 and relA1) (Sambrook et al. 1989) was the bacterial host strain used for purposes of cloning into CoIE1 origin of replication vectors pUC 18/19.

*E. coli* DH5αF' (F', supE44, hsdR17, recA1, endA1, gyrA96, thi-1 and relA1) (Sambrook et al. 1989) was used as the recipient strain for recombinant M13mpl8/19 phage constructs. Such constructs were used for the production of single-stranded templates used in DNA sequence analysis.

The pDKR1 plasmid was the original donor molecule for subclones created by Azadpour (1991) and Baker (1992), from which all subclones were derived for this study. The plasmid pDKR1 is an in vivo construct produced through conjugation of strains carrying RP4 with the pDX1 TOL plasmid. This construct contains the RP4 replicon along with a 40 kb segment of pDX1, encoding all necessary TOL functions (Shaw and Williams 1988). The *P. putida* strain PaW630 (tryptophan auxotroph and streptomycin resistant (1 mg/ml)) contained the pDKR1 plasmid and was obtained from Dr. D. A. Kunz, Department of Biological Sciences, University of North Texas. Genes encoding the antibiotic resistance to kanamycin (100
µg/ml), tetracycline (50 µg/ml), and ampicillin (1 mg/ml) are conferred through the presence of the RP4 replicon.

**Biological media and growth conditions**

Media used for this study included YT broth, LB broth and *E. coli* minimal medium. These were supplemented with antibiotics when appropriate. Solid media were produced by adding 1.5% w/v agar to any of these liquid media.

YT broth was used for the propagation of the M13 *E. coli* phage hosts and was composed of 0.8% (w/v) Bacto Tryptone™, 0.5% (w/v) yeast extract and 0.5% (w/v) NaCl. *E. coli* cultures were grown in LB (Luria-Bertani) medium which contained 1.0% (w/v) Bacto Tryptone™, 0.5% (w/v) yeast extract and 1.0% (w/v) NaCl. *E. coli* strains were also grown on minimal medium. *E. coli* minimal medium was made by dissolving 10.5 g K₂HPO₄, 4.5 g KH₂PO₄, 1.0 g (NH₄)₂SO₄ and 0.5 g Na₃Citrate in 990 ml of ddH₂O and autoclaved. After the solution cooled, 1.0 ml of 1 M MgSO₄, 1.0 ml of 10 mg/ml thiamine and 7.3 ml of 50% glycerol (carbon source, 50 mM final concentration) were added.

**Long-term storage of cultures**

In order to store bacterial strains for extended periods of time, ultra cold temperatures of -70°C were utilized. A single colony from a particular bacterial strain was chosen from a selective plate and inoculated into the appropriate 5 ml liquid medium containing a selective agent.
(usually 50 \( \mu \text{g/ml} \) ampicillin). The culture was incubated overnight at 37°C at 250 rpm. Five-hundred \( \mu \text{l} \) of overnight culture was added to a sterile tube containing 500 \( \mu \text{l} \) of sterile 80% glycerol. After vortexing the mixture, the tube was placed at -70°C for long term storage.

For routine use, bacterial cultures were maintained on a rich medium plate containing the applicable antibiotic (if appropriate) and were stored at 4°C.

**Large scale plasmid isolation from *E. coli* strains**

Plasmids were isolated using CsCl/ethidium bromide density gradient centrifugation according to the procedure of Tanaka and Weisblum (1975). A 50 ml flask of rich media containing 50 \( \mu \text{g/ml} \) ampicillin was inoculated from a single colony of the desired strain of *E. coli*. After overnight incubation at 37°C and 250 rpm, 10-50 ml of the overnight was inoculated into a Fernbach flask containing 1000 ml of LB medium and ampicillin (50 \( \mu \text{g/ml} \)). This flask was incubated at 37°C and 250 rpm for 18-24 hours. The cells were collected by centrifugation using a precooled Sorvall GS3 rotor at 5213 x g for 6 minutes. After discarding the supernatant, the cells were resuspended in ice-cold 0.15 M NaCl (20 ml per liter of original culture). The resulting cell suspension equivalent to 750-1000 ml of culture was transferred to a 45 ml Oak Ridge style centrifuge tube and centrifuged in a Sorvall SA600 rotor at 5213 x g for 5 minutes at 4°C. From this point forward, all work was
performed in the cold room at 4°C. The supernatant was
discarded and the remaining pellet was thoroughly resuspended
in 10 ml of 50 mM Tris-HCl, 25% sucrose, pH 8.0 until no
cellular clumps remained. Two ml of 5 mg/ml lysozyme was
added, followed by thorough mixing accomplished by several
tube inversions. After a five minute incubation on ice, 4 ml
of 0.25 M Na₂EDTA (pH 8.0) was added. The tube was mixed by
inversion and placed on ice for 5 minutes. Next, 5 ml of 5 M
NaCl was added, followed by mixing by inversion. Immediately
after mixing, 2 ml of 10% SDS (SDS must be kept at room
temperature to prevent the SDS from precipitating out of
solution) was added. The solution was mixed gently yet
completely and the resulting mixture was placed on ice for
two hours in the cold room. The solution was centrifuged at
38712 x g for 60 minutes in an SA600 rotor to pellet out the
resulting cellular debris. The plasmid containing
supernatant was poured into a graduated cylinder and an equal
volume of isopropanol was added. The mixture was then
transferred to a clean 250 ml GSA centrifuge bottle and the
bottle was placed into a dry ice-ethanol bath until the
contents were frozen solid. The tube was briefly transferred
to a room temperature water bath until the contents were no
longer frozen and then immediately centrifuged in a Sorvall
GSA rotor for 20 minutes at 9268 x g. The supernatant was
discarded and the pellet was resuspended with gentle stirring
by a magnetic stirrer in 16 ml of ice-cold TE buffer (10 mM
Tris-HCl, 1 mM EDTA, pH 8.0) at 4°C. In order to degrade any
RNA present in the sample, DNase-free RNase was added to a final concentration of 20 μg/ml. The mixture was left to stir for 60 minutes, then the solution was transferred to a Oak Ridge style tube and centrifuged in an SA600 rotor for 10 minutes at 14481 x g. The plasmid containing supernatant was poured into a clean graduated cylinder. Solid molecular biology grade cesium chloride was added at a ratio of 1.06 grams per ml of original solution. The solution was divided equally between two T1270 ultra centrifuge tubes to which 400 μl of 10 mg/ml ethidium bromide was added. The two tubes were balanced within 0.02 g of each other using a balancing solution containing 1.06 g CsCl per ml of TE. The tubes were centrifuged in a Dupont OTD65 ultracentrifuge for 40 hours at 43467 x g.

The centrifugation established the formation of a buoyancy gradient where supercoiled plasmid DNA can be isolated free from linear, nicked or chromosomal DNA. The bands of DNA can be visualized with the help of the ethidium bromide in the tube and a long wave ultraviolet lamp (~310 nm). All work from this point forward was done in reduced light. A diagrammatic representation of the various bands formed after centrifugation can be observed in Fig. 9. By inserting a 20G needle just below the lower plasmid-containing band, the supercoiled plasmid DNA was removed. To insure the purity of the plasmid, this solution was added to a new T1270 tube, balanced, and centrifuged again. The ethidium bromide was removed by several extractions with an
Protein

Chromosomal DNA (linear or nicked)

Plasmid DNA (supercoiled)

RNA
Fig. 9. CsCl buoyancy gradient used to separate (chromosomal) DNA from closed circular supercoiled (plasmid) DNA. A diagramatic representation of a T1270 tube after completion of the 40 hour centrifugation at 43467 x g.
equal volume of water-saturated isobutanol. The extractions were repeated until the upper isobutanol layer was no longer pink in color. Once the extractions were complete, it was no longer necessary to work under subdued light. Two volumes of water and nine volumes of cold ethanol were added to the DNA solution. The resulting suspension was transferred to two 30 ml siliconized Corex™ tubes, mixed and placed in a dry-ice ethanol bath for 10 minutes and then centrifuged at 9240 x g for 15 minutes. The DNA pellet was resuspended in 300 µl of 0.3 M sodium acetate, transferred to a microcentrifuge and 1 ml of cold 100% ethanol was added. After the mixture was vortexed, it was centrifuged in a microcentrifuge at 14481 x g for 10 minutes at 4°C. The resulting DNA pellet was washed with 1 ml of cold 70% ethanol and centrifuged again for five minutes. The supernatant was discarded and the pellet was dried for 5 minutes in a Savant Speed Vac™ vacuum concentrator. The dried DNA pellet was then resuspended in one ml of TE buffer. Qualitative and quantitative evaluations were performed by gel electrophoresis and spectrophotometric analysis as described below.

Quantitative and qualitative identification of DNA

DNA can be visualized with the use of ethidium bromide (EtBr). EtBr intercalates between the base pairs of DNA and when exposed to UV radiation, emits an orange fluorescence. DNA samples were electrophoresed adjacent to marker DNA of known size(s) and concentration. The gel was stained in a 50
μg/ml EtBr solution. By comparing the experimental samples to the marker DNAs, estimations of size and concentration were made. EtBr also allows for the visualization of any RNA and degraded DNA that may have been present in the sample, therefore the quality of DNA was observed.

If sufficient quantities of DNA were available, another method was also used, spectroscopy. DNA absorbs UV radiation optimally in the 256 nm - 260 nm range. An absorbance reading of 1.0 in this range is equivalent to 50 μg/ml of duplex DNA. Diluted DNA samples were subjected to a wavelength scan from 220 nm to 320 nm. The presence of contaminants such as phenol or protein can cause a shift of the absorbance maximum to the right (270 nm - 280 nm). A Beckman DU-40 spectrophotometer was used for the analysis.

**Complete and partial digestion of DNA with restriction endonucleases**

Restriction endonuclease digestion conditions were set by the manufacturer. Generally 2 units of enzyme per μg of duplex DNA was used. A typical reaction mixture would be as follows:

- 2.0 μl 10x Reaction buffer (provided by the manufacturer)
- 12.0 μl sterile distilled deionized water
- 5.0 μl DNA (5 μg)
- 1.0 μl restriction enzyme (10 units)
- 20.0 μl total volume
The volume of restriction enzyme solution may be increased, but under no circumstances should it ever exceed 10% of the total reaction volume. If this volume is increased above 10%, the enzyme may not cleave as expected (altered specificity) due to the glycerol found within the enzyme storage buffer. Different enzymes may require different digestion conditions. Each enzyme is supplied with a specific 10x reaction buffer for that enzyme. Reaction mixtures were generally incubated at 37°C for 2 hours, unless otherwise specified by the manufacturer.

Partial restriction endonuclease digestions were performed by adding 10 to 100 times less enzyme to the reaction mixture. The reaction was carried out at the appropriate temperature and aliquots of the mixture were removed at various time points. The addition of phenol to the aliquot denatured the enzyme, rendering it inactive. Samples were then electrophoresed on an agarose gel and analyzed to determine the completeness or degree of restriction.

After the restriction endonuclease digestion was complete, the resulting DNA solution was subjected to phenol extraction and ethanol precipitation (described later) to remove any salts and/or enzymes present. The clean DNA now can be used in a variety of ways which will be further discussed below.
**Bacterial alkaline phosphatase treatment of vector DNA**

In order to help prevent a cloning vector from being reconstructed during ligation, bacterial alkaline phosphatase (BAP) was used to remove the 5' phosphate moiety. This procedure was completed after restriction endonuclease digestion. A typical reaction mixture was as follows:

- 5.0 μl linear DNA (no more than 1 μg)
- 44.0 μl TE buffer
- 1.0 μl bacterial alkaline phosphatase (100-200 units)
- 50.0 μl total volume

The reaction mixture was placed at 65°C for 3 hours. Flash spins in a microcentrifuge were done at 30 minute intervals to prevent changes in salt concentration in the reaction due to evaporation.

**Phenol extraction of DNA solutions**

After restriction endonuclease digestion and/or BAP treatment of DNA, it was necessary to remove the salts and enzymes added to these mixtures and precipitate the DNA. The removal of enzymes was accomplished by adding the organic solvent phenol. Phenol denatures the proteins, causing them to precipitate from the aqueous solution without harming the DNA. A distilled phenol stock was saturated with TE buffer and 0.1% 8-hydroxyquinoline was added. The 8-
hydroxyquinoline plays two roles. First, it is an anti-
oxidant. Second, it changes the phenol layer to a yellow
color, making it more visible than its normal clear color and
easily distinguished from the aqueous layer.

TE buffer was added to the tube containing the protein
DNA solution to a total volume of 100 μl. The contents of
the tube were mixed and 100 μl of TE-saturated phenol was
added. This mixture was vortexed well and centrifuged at
14481 x g for 2 minutes. The lower phenol layer was removed
using a micropipettor and an additional 100 μl of phenol was
added. Vortexing was followed by centrifugation and phenol
removal. In order to remove any residual phenol, 500 μl of
diethyl ether was added and the resulting solution was
vortexed well. The top ether layer was removed using a drawn
out Pasteur pipet, after which the tube was left open for the
volatilization of any remaining ether. To precipitate the
DNA, 10 μl of 3.0 M sodium acetate was added followed by 350
μl of chilled 100% ethanol. The tube was vortexed and
centrifuged at 14481 x g for 10 minutes. The supernatant was
removed using a drawn out Pasteur pipet and 500 μl of chilled
70% ethanol was added to remove any residual salts in the
tube. The tube was mixed by gentle inversion and centrifuged
at 14481 x g for 5 minutes. The supernatant was removed and
discarded as before. The open tube was then placed in a
Savant Speed Vac™ vacuum concentrator for five minutes. The
resulting dry, clean DNA pellet was resuspended in sterile
distilled water at a desired concentration. The DNA solution
was stored at -20°C until use.

**Ligation of DNA for cloning into plasmids and RF phage**

Once the target and vector DNAs were digested and purified, a ligation of the target into the vector, forming a recombinant DNA molecule, was performed. A typical plasmid ligation reaction consisted of a one to one molar ratio of vector to target DNA. Usually 100 ng of vector DNA per ligation was used and the target DNA was adjusted accordingly. The total amount of DNA should never be above 1 µg. Adjustments to the molar ratio were made when problems occurred. The molar ratio was changed as needed from a 10/1 to a 1/10 vector/target, depending on the situation. Four microliters of 5x ligation buffer (250 mM Tris-HCl pH 7.6; 25% w/v PEG 8000 (polyethylene glycol, av. Mol. Wt. 8000); 50 mM MgCl₂; 5.0 mM ATP; 5.0 mM DTT) and 1 µl of T4 DNA ligase (1 unit) were added to a microfuge tube containing the DNA and water. Sterile distilled deionized water was used to bring the total volume up to 20 µl. The resulting reaction mixture was incubated at room temperature for five hours. Upon conclusion of incubation, two volumes of TE buffer were added and mixed well. This mixture was then stored at -20°C until ready for transformation.

Recombinant plasmid DNAs had three basic uses. First, recombinant DNAs were used as templates for DNA sequencing. Second, recombinant DNAs were necessary for the formation of
the novel strains of bacteria used for the expression of cloned enzyme products. The vectors used in this study facilitated the expression of proteins encoded by the cloned DNAs by the host *E. coli*. Finally, recombinant DNAs were used in the formation of site-directed mutants. Specific DNA sites were mutated. The effect of these mutations was studied through the expression of the encoded enzyme products.

Cloning into M13 vectors was done similarly with one exception, the total volume was usually 10 \( \mu l \) instead of 20 \( \mu l \). M13 recombinant DNA molecules were used for DNA sequencing purposes only.

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

Preparation of competent *E. coli* strains DH5\( \alpha \) and DH5\( \alpha F' \) was done according to the procedure developed by Chung et al. (1989). DH5\( \alpha \) and DH5\( \alpha F' \) were used for transformations with ColE1-derived plasmids (pUC18/19) and M13 phages respectively. A five milliliter portion of YT broth was inoculated with a single colony of the appropriate bacterial strain from a fresh YT agar plate. This culture was placed into a 37\(^\circ\)C New Brunswick G25 incubator shaking at 250 rpm overnight. After the overnight incubation, 500 \( \mu l \) was aseptically removed and placed into a sterile 250 ml Erlenmeyer flask containing 50 ml of YT broth. All steps involving work with the culture from this point forward were
done using aseptic techniques. The new inoculum was placed into the shaker incubator with the same conditions as above. At regular intervals, the culture was monitored spectrophotometrically at 550 nm using a Beckman DU-40 spectrophotometer. The culture was grown until the absorbance at 550 nm was equal to 0.45. The contents of the flask was transferred to a sterile 50 ml conical tube and placed into an ice water bath for 20 minutes. The tube was centrifuged at 1000 x g for 10 minutes at 4°C. The supernatant was discarded and the cellular pellet was resuspended by short vortexes in 5.0 ml of TSS solution (85% LB broth, 10% wt/vol PEG 8000, 5% vol/vol DMSO and 50 mM MgCl₂, pH 6.5). The tube was placed on ice until ready for use. For best results, the cells were used for transformation within 6 hours. Any unused portion of the competent cells was stored in 300 µl aliquots at -70°C for up to one month.

**Transformation of *E. coli* strains DH5α and DH5αF' with recombinant ColEl plasmids and M13 phage DNA**

*E. coli* strain DH5α was used for transformations with recombinant ColEl-derived plasmids (pUC18/19). A 300 µl volume of competent cells was added to a sterile 1.5 ml microfuge tube containing 1/3 of a ligation reaction. The tube was gently mixed and placed into an ice water bath for 30 minutes. The tube was then transferred to a 42°C water bath for 2 minutes to heat shock the cells. After heat
shocking was complete, 1 ml of sterile YT medium was added to the tube and it was placed in a 37°C gyratory shaker at 250 rpm for 45 minutes. This out growth period allowed for the expression of the ampicillin resistance gene found on the ColEl-derived plasmids used in this study. Upon completion of the out growth period, 20 and 200 μl aliquots of the mixture were spread onto YT agar plates containing 50 μg/ml ampicillin, 10 μl of 100 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 50 μl of 2% 5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-Gal). The suspended cell mixture was allowed to soak into the agar plates at room temperature for 30 minutes before placing them into a 37°C incubator overnight. Only cells which were transformed into ampicillin resistant E. coli DH5α formed colonies on the plate. Those colonies which contained a recombinant plasmid were white, while those which contained a nonrecombinant plasmid were blue.

E. coli strain DH5αF' was used for transformation with M13mpl8/19 recombinant DNA molecules. A 300 μl volume of competent cells was added to a sterile 1.5 ml microfuge tube containing 1/3 of a ligation reaction. The tube was gently mixed and placed into an ice water bath for 30 minutes. The tube was then placed into a 42°C water bath for 2 minutes to heat shock the cells. 30 μl and 250 μl portions of this mixture were then added to separate tubes (at 42°C) containing the following:

3.0 ml  YT soft agar
50.0 µl 2% X-Gal
10.0 µl 100 mM IPTG
200.0 µl DH5αF' lawn cells

The contents of each tube were mixed gently and quickly poured onto a prewarmed YT agar plate. Tilting the plate in several directions allowed for even spreading of the soft agar. The plate was left undisturbed for 20 minutes at room temperature in order to allow the soft agar to solidify. Once solidified, the plates were placed into a 37°C incubator overnight. Recombinant phage produced clear plaques, while nonrecombinant phage produced blue plaques.

Screening of recombinant plasmids and phage

The transformants infected with the recombinant M13mp18/19 DNAs were initially screened on the basis of color formation. As stated earlier, those plaques portraying a clear color are the recombinant DNA phage particles and those having a blue color resulted from nonrecombinant DNA phage particles. The absence of color was due to the interruption of the α-subunit of the β-galactosidase gene located in the multiple cloning site by the target DNA. This interruption causes the formation of a dysfunctional β-galactosidase gene product, which cannot cleave its substrate X-Gal. An active β-galactosidase gene cleaves X-gal resulting in the formation of a blue precipitate. Further screening of the phage particles by horizontal agarose gel electrophoresis of DNAs will be described later.
Fig. 10. Screening grid for master plating of recombinant strains.
The screening for recombinant plasmids uses the same color determination as described above. In order to maintain a library of different recombinant clones, a master plate was made using a screening grid as shown in Figure 10. A YT/ampicillin (50 μg/ml)/X-Gal/IPTG agar plate was attached to the grid. Using a sterile toothpick, each recombinant colony was carefully touched and streaked onto a numbered area. The resulting master plate was then placed into a 37°C incubator overnight. After incubation, the master plate was stored at 4°C for future reference.

**Rapid plasmid isolation using alkaline lysis technique (Birnboim and Doly 1979)**

Cells from the master plate were inoculated into culture tubes containing 5 ml of YT medium with 50 μg/ml ampicillin. The tubes were placed in a shaker incubator at 37°C and 250 rpm overnight. Upon completion of the incubation, 1.5 ml of each culture was transferred into sterile 1.5 ml microfuge tubes and the tubes were centrifuged at 14481 x g for 2 minutes. The supernatant was removed and another 1.5 ml of culture was added to each tube. The tubes were centrifuged as before and the supernatant was removed by aspiration in order to leave the bacterial pellet as dry as possible. The pellet was then resuspended in 100 μl of ice-cold solution A (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0, 6.0 mg/ml lysozyme) by vigorous vortexing. The resulting suspension was incubated at room temperature for 5 minutes. This was
immediately followed by the addition of 200 μl of freshly prepared solution B (0.2 N NaOH, 1.0% SDS). The tubes were mixed by inversion and placed into an ice water bath for 5 minutes. After this incubation, 150 μl of ice-cold 5 M potassium acetate, pH 4.8, was added, the tubes were mixed by inversion and placed into an ice water bath for 5 minutes. The tubes were then centrifuged at 4°C for 5 minutes at 14481 x g. The supernatant was removed and transferred to a new sterile 1.5 ml microfuge tube. An equal volume of phenol/chloroform (1:1) was added and the tubes were vortexed well. In order to separate the layers, the tubes were centrifuged for 2 minutes at 14481 x g. The upper aqueous layer was then transferred to a new sterile microfuge tube and 750 μl of -20°C 100% ethanol was added. The tubes were mixed well by vortexing and centrifuged at 4°C for 10 minutes at 14481 x g. The supernatant was removed and the DNA pellet was washed with 1.0 ml of -20°C 70% ethanol. The tubes were centrifuged as before for 5 minutes and the supernatant was removed by aspiration. In order to remove any residual ethanol, the tubes were placed in a vacuum desiccator for 3 minutes. Thirty microliters of TE buffer (pH 8.0) containing 25 μg/ml heat-treated RNase A was added to the dried DNA pellet. The mixture was vortexed and placed at 65°C for 10 minutes. The tubes were vortexed once again and then were stored at -20°C. The DNA was now ready for DNA sequence analysis, restriction endonuclease digestion or electrophoresis.
Preparation of M13 templates (Schreier et al. 1979)

Culture tubes containing 5 ml of YT broth were inoculated with 100 μl of an overnight *E. coli* DH5αF’ culture. Each of these tubes was in turn inoculated with an agar plug containing a single clear isolated recombinant M13mp18/19 plaque. The transfer was completed using a sterile 5 1/4” Pasteur pipette. The resulting tubes were placed into a rotary incubator at 37°C and 250 rpm for 5 hours. After the incubation time was complete, 3.5 ml of each inoculum was transferred into a 5 ml polypropylene tube. These tubes were centrifuged at 14481 x g for 8 minutes. The M13 phage particle containing supernatant was poured into a new 5 ml polypropylene tube. Care was taken not to disturb the cellular pellet. The supernatant was once again centrifuged under the same conditions to ensure that all possible cells were removed. The supernatant was again poured into a new 5 ml polypropylene tube and 900 μl of 20% PEG/2.5 M NaCl was added. The mixture was vortexed thoroughly and left at room temperature for 15 minutes. Upon completion of incubation, the tubes were centrifuged for 8 minutes at 14481 x g to pellet the phage particles. The supernatant was carefully removed. Care was taken not to disturb the phage pellet. The tubes were inverted and the lips of the tubes were wiped with a paper towel. Phage pellets were resuspended in 100 μl of TES buffer (20 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.1 mM Na₂EDTA). The phage suspension was transferred to a 1.5 ml centrifuge tube, where
100 μl of TE-saturated phenol was added to extract the phage protein coats. Tubes were vortexed well and centrifuged at 14481 x g for 2 minutes. The phenol layer was removed using a micropipettor. The phage solution was then extracted with 100 μl of chloroform. The mixture was vortexed and centrifuged as before. This time the upper aqueous phase was removed with a micropipettor and transferred to a new 1.5 ml microfuge tube. It was very important to remove only the aqueous phase, for contamination by the organic phase can inhibit DNA sequencing reactions. Nine microliters of 3 M sodium acetate was added to each tube, followed by 250 μl of ice-cold 100% ethanol. The tubes were vortexed briefly and centrifuged for 10 minutes at 4°C and 14481 x g. The supernatant was removed and the DNA pellet was washed with 1 ml of ice-cold 70% ethanol. The tubes were gently mixed by inversion to ensure the DNA pellet was not disturbed. After mixing, the tubes were centrifuged for 5 minutes using the same conditions as above. The supernatant was removed using a drawn out Pasteur pipet and the tubes were placed into a vacuum desiccator for five minutes. The dried DNA pellets were resuspended in 33 μl of TE buffer. The presence of DNA was verified by loading a sample onto a 1% agarose gel followed by electrophoresis.

**Horizontal agarose gel electrophoresis**

Agarose gel electrophoresis was the technique used to analyze a variety of different DNA samples. These samples
include, but are not limited to, rapid plasmid preparations, M13 preparations, restriction endonuclease digestions and CsCl-purified DNA preparations. The preparation of the agarose solution was as follows:

5X TBE Agarose Gel Electrophoresis Running Buffer

- 54.0 g tris base
- 27.5 g boric acid
- 20.0 ml 0.5 M EDTA, pH 8.0
- distilled deionized water up to 1 liter
- total volume

Agarose Gel Solution (1%)

- 1.0 g electrophoretic grade agarose
- 20.0 ml 5X TBE agarose gel electrophoresis running buffer
- 80.0 ml distilled deionized water

The percent agarose added was often altered depending on the size of DNA fragments analyzed. For example, if the DNA sample contained fragments ranging in size from 200 base pairs (bp) to 1000 bp, a higher percentage agarose was used in order to separate the smaller fragments more efficiently.

The contents of the agarose gel solution were placed into a 250 ml Erlenmeyer flask. The flask was then weighed and placed into the microwave for heating. The mixture was
heated until no undissolved agarose crystals remained in the flask. The flask was weighed again and distilled deionized water was added until the original weight was reached. The loss of weight was due to evaporation of the contents of the flask. The solution was allowed to cool slightly before pouring the gel into its form. Once poured, the gel solution cooled and solidification occurred within approximately 45 minutes. After solidification, the well forming comb was removed and the gel was placed into an electrophoresis tank. The tank was filled with 1X TBE buffer until it was 1-2 mm above the top of the gel. The gel was now ready for loading.

A 5X agarose gel loading buffer (25.0% glycerol, 0.5% SDS, 0.1% bromophenol blue, 0.1% xylene cyanol and 50.0 mM EDTA) was added to the samples so that the final concentration of the loading buffer was 1X. The samples were then loaded into the wells using a micropipettor. Once loaded, electrophoresis was performed at 80 V for 1 hour. Upon completion of electrophoresis, the gel was placed into a 0.5 μg/ml solution of ethidium bromide (EtBr) for staining. After staining for 20-30 minutes, the gel was placed into deionized water for 10-20 minutes to destain. Destaining removes residual EtBr from the gel, therefore reducing the background. The gel was now ready to be examined under ultraviolet illumination and photographed (if necessary) using an MP-4 transilluminator system.

Sanger dideoxyribonucleotide DNA sequence analysis of
single- and double-stranded templates (Sanger et al. 1977)

The procedure for sequencing double-stranded DNA templates differs from that for single-stranded templates only at the early stages of the protocol. Both methods followed the USB step-by-step protocols (1989) and reagents were part of the Sequenase™ version 2.0 DNA sequencing kit.

Approximately 4 μg of clean plasmid DNA (isolated by the alkaline lysis method discussed earlier) was added to a 500 μl microcentrifuge tube and sterile water was added to a total volume of 30 μl. Three microliters of 2N NaOH was added, followed by mixing and incubation at room temperature for five minutes. The alkali denatures the double-stranded DNA such that an oligonucleotide primer may later hybridize to a complementary region of the template under the proper conditions. Upon completion of the incubation, 120 μl of -20°C 100% EtOH was added followed by the addition of 5 μl 3 M sodium acetate. The mixture was mixed well and centrifuged at 4°C for 10 minutes at 14481 x g. The supernatant was removed and the denatured DNA pellet was washed with 500 μl of -20°C 70% EtOH. The supernatant was removed and the pellets were dried in a vacuum desiccator for 5 minutes. The resulting dried DNA pellets can be stored for up to two weeks at -20°C. DNA pellets were resuspended in 5 μl of sterile water immediately before use. From this point forward, the procedures for sequencing double- and single-stranded templates are the same.
The following description was designed for the processing of twelve templates at one time. Proportional adjustments to the volumes were made if fewer than twelve templates were processed. Thirteen microliters of 
-40 °C "universal" M13 forward sequencing primer solution, 26 μl of reaction buffer and 26 μl of sterile distilled water were added to a 0.5 ml microcentrifuge tube and mixed. Five microliters of the resulting mixture was dispensed into each of twelve microfuge tubes containing 5 μl of the templates (either single- or double-stranded) to be sequenced. The template/primer mixtures were then placed into a 65°C water bath for 2 minutes. The mixtures were removed and placed into a glass tray filled with the 65°C water and allowed to cool slowly until the water was less than 35°C. Alkali-denatured double-stranded template/primer mixtures were often placed at 37°C for 45 minutes instead of the above procedure, but both procedures produced similar consistent results. While the primers were annealing to the templates, sets of "termination" tubes were prepared for each of the templates. Each set contained four color coded 0.5 ml microfuge tubes, where each color tube would contain a particular dideoxyribonucleotide for termination of synthesis. Two and one half microliters of ddGTP mixture was added to the G termination tube (orange), 2.5 μl of ddATP was added to the A termination tubes (green), 2.5 μl of ddTTP was added to the T termination tubes (blue) and 2.5 μl of ddCATP was added to the C termination tubes (yellow). The use of colored tubes
helped minimize the chance for mistakes. After the termination tubes were ready, it was important to prepare the labeling solution. The labeling solution was composed of 13.0 μl of 0.1 M DTT (dithiothreitol), 22.8 μl sterile distilled water, 5.2 μl labeling mix and 6.5 μl 35S-dATP (81.25 μCi, provided by NEN). Once the primer/template mixtures had cooled below 35°C, 3.5 μl of the labeling mixture was added to each tube. The tubes were mixed well and briefly centrifuged to collect any condensation which may have occurred during the primer annealing. At this point, it was time to dilute the Sequenase™ enzyme 8-fold with Sequenase™ dilution buffer. This was accomplished by adding 3.3 μl of Sequenase™ enzyme and 1.6 μl of pyrophosphatase to 21.5 μl of ice-cold Sequenase™ dilution buffer. It was now time to begin the actual DNA sequencing reactions by first placing the termination tubes at 37°C. Two microliters of the diluted enzyme was added to each of tubes containing the template/primer/labeling mixture. These tubes were allowed to incubate at room temperature for 4 minutes. Three and one half microliters of the incubated mixture was transferred to each of the four G/A/T/C termination tubes. Great care was taken to ensure a complete mixing of the solutions. The tubes were placed back into the 37°C water bath and were allowed to incubate for approximately 20 minutes. Upon completion of the incubation, 4.0 μl of a formamide stop solution was added to each tube (an overview of the timed sequencing reactions can be seen in Fig. 11). The tubes were
### Sequencing Summary

<table>
<thead>
<tr>
<th>Time</th>
<th>Action</th>
<th>Temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Warm termination tubes for templates 1-4 at 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>1</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 1. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>2</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 2. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>3</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 3. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>4</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 4. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>5</td>
<td>Add 3.5 μl template 1 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>6</td>
<td>Add 3.5 μl template 2 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>7</td>
<td>Add 3.5 μl template 3 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>8</td>
<td>Add 3.5 μl template 4 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>9</td>
<td>REST or use this time to catch up!!!</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Warm termination tubes for templates 5-8 at 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>11</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 5. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>12</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 6. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>13</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 7. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>14</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 8. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>15</td>
<td>Add 3.5 μl template 5 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>16</td>
<td>Add 3.5 μl template 6 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>17</td>
<td>Add 3.5 μl template 7 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>18</td>
<td>Add 3.5 μl template 8 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>19</td>
<td>REST or use this time to catch up!!!</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Warm termination tubes for templates 9-12 at 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>21</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 9. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>22</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 10. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>23</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 11. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>24</td>
<td>Add 2 μl &quot;Enzyme&quot; to template 12. Mix.</td>
<td>Room</td>
</tr>
<tr>
<td>25</td>
<td>Add 3.5 μl template 9 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>26</td>
<td>Add 3.5 μl template 10 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>27</td>
<td>Add 3.5 μl template 11 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>28</td>
<td>Add 3.5 μl template 12 to termination tubes (GATC) 37°C</td>
<td>Room</td>
</tr>
<tr>
<td>29</td>
<td>REST or use this time to catch up!!!</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Add 4 μl Stop Solution to termination tubes 1-4</td>
<td>Room</td>
</tr>
<tr>
<td>35</td>
<td>Add 4 μl Stop Solution to termination tubes 5-8</td>
<td>Room</td>
</tr>
<tr>
<td>40</td>
<td>Add 4 μl Stop Solution to termination tubes 9-12</td>
<td>Room</td>
</tr>
</tbody>
</table>
Fig. 11. Summary of the DNA sequencing reaction protocol.
This protocol was developed to handle twelve reactions.
mixed and centrifuged briefly. The samples were stored at -20°C until they were ready for loading onto a sequencing gel. The samples were heated to 90°C for two minutes and immediately placed on ice prior to loading to help reduce the secondary structure formation of the newly synthesized DNA.

Denaturing polyacrylamide sequencing gels (Sanger and Coulson 1978)

Denaturing polyacrylamide sequencing gels were used to separate the DNA fragments generated by the dideoxyribonucleotide sequencing procedure. These gels have the ability to resolve radioactively-labeled DNA fragments that differ by only one nucleotide in size to a range of up to 300-400 nucleotides in length.

Two glass plates (52 cm X 41 cm X 0.6 cm) were cleaned using a glass cleaner followed by ethanol. The insides of both plates were treated with a 5% solution of dichlorodimethylsilane dissolved in heptane. It was important to coat thoroughly each plate to prevent a gel catastrophe later. The plates were then placed into a drying oven for 10 minutes at 100 °C. The plates were removed and placed onto a lab bench for cooling (siliconized surface up). After cooling to room temperature, the plates were placed on top of one another such that the siliconized surfaces were on the inside. The plates were offset by 0.5 inches lengthwise and Delrin™ spacers (51 mm X 13 mm X 0.25 mm) were introduced between the plates along each side. Another
spacer was inserted between the plates at the bottom of the cassette and made flush with the other two spacers. Once assembled, the plates were clamped together using medium sized binding clips on the sides and four large binding clips at the bottom to prevent leakage of the gel solution. At the top end of the cassette, a small spacer was inserted in the middle of the cassette approximately 1 cm deep and clamped with a large binding clip. This was an added measure to ensure uniform thickness of the polyacrylamide gel. Assembly of the gel cassette was now complete.

One-hundred ml of 6% denaturing polyacrylamide gel solution was prepared in a 250 ml beaker by adding 42.4 grams urea, 20 ml of 30% acrylamide/1% bisacrylamide and 10.0 ml 10X TBE sequencing buffer. Only top quality, molecular biology grade reagents were used since contaminants can affect the resolution of the fragments into distinct bands. Distilled deionized water was added to just below 100 ml total volume. The solution was allowed to stir until all of the ingredients were dissolved. The solution was transferred to a graduated cylinder and distilled deionized water was added to a total volume of 100 ml. Undissolved particles were removed by filtering the solution through a Buchner funnel containing one piece of Whatmann No. 1 filter paper. The filtrate was collected into an Erlenmeyer filtration flask to which 0.12 grams ammonium persulfate was added. The flask was covered with a rubber stopper and allowed to degas in vacuo until all dissolved gases were removed. Twenty-five μl of TEMED
(N,N',N" tetramethylenediamine) was added and the contents of the flask mixed by gentle swirling. The gel solution was now ready for immediate pouring.

In order to pour the gel solution without introducing air bubbles, a modeling clay reservoir was constructed to one side at the top of the cassette. The gel cassette was placed at a 45° angle and slightly tilted towards oneself for pouring. The gel solution was quickly and evenly poured into the reservoir. It was important for the reservoir not to become dry, for discontinuous pouring would introduce air bubbles into the gel. After the cassette was approximately 80% full, the cassette was laid onto four robber stoppers and allowed to polymerize. Polymerization occurred usually within one hour.

Sequencing gels can be electrophoresed the day of pouring, or saved until the next day. If the gel was not used the day of pouring, it was important to cover the exposed area of the cassette with plastic wrap so that the gel did not dry out. What was the bottom of the gel cassette when pouring was now the top of the gel for running. The binding clips and spacer were removed to expose the loading surface of the gel. The gel cassette was now installed vertically into a DNA sequencing electrophoresis stand and clamped into place. The upper and lower chambers of the sequencing apparatus were filled with 1X TBE sequencing buffer (10.0 mM tris, 1.2 mM boric acid, 0.1 mM EDTA). Care was taken to remove all air bubbles between the plates to
insure an unobstructed path for the current. The gel was pre-electrophoresed for one hour at 1400 volts. Upon completion of pre-electrophoresis, the power supply was turned off and air bubbles formed during this time were removed. Shark’s teeth well forming combs were inserted between the plates until the teeth of the comb just touched the surface of the gel. The gel was now primed for loading.

Samples to be loaded were placed at 90°C for 2 minutes and then immediately placed into an ice bath. This helped to eliminate any secondary DNA structure which could have affected migration rates. Using a 10 μl Hamilton® syringe fitted with a 32G needle, 1.0 μl - 2.5 μl aliquots of the samples were loaded into the wells in the following order: GATC. A well was skipped between each set for ease in interpreting the gel later. Samples were also loaded asymmetrically so that the correct left and right sides of the gel could be readily identified. Samples were loaded in sets of four followed by brief electrophoresis to minimize diffusion of the DNA samples and leakage of the samples between wells. If diffusion of the samples occurred, resolution was affected. Diffusion was minimized by loading the gel quickly and electrophoresing at high voltages (1700 - 2400 V). The time for electrophoresis varied, depending upon the distance from the primer the desired sequence to be analyzed was located. Xylene cyanol, a dye found in the stop solution, was used to estimate the distance traveled.
Autoradiography of sequencing gels

Upon completion of electrophoresis, the gel cassette was removed from the electrophoresis unit. All clips and spacers were removed from the cassette and the plates were pried apart using a thin spatula. A piece of Whatmann 3MM™ filter paper, cut to 43 cm X 35 cm, was placed over the gel such that it was centered and covered the bottom of the gel. Excess acrylamide was trimmed away using a razor blade. The gel adhered to the filter paper and was removed from the glass plate. A piece of plastic wrap was stretched over the gel surface such that no wrinkles were present. Excess plastic wrap was trimmed away and the resulting “sandwich” was then transferred to a gel slab dryer. The gel was dried under vacuum for one hour at 80°C. The plastic wrap was removed and the gel/filter paper was placed in a 12% methanol, 10% acetic acid solution for 15 minutes. This solution leaches the urea from the dried gel. If the urea is not removed, two problems can result. First, urea is hydroscopic and in humid conditions, the dried gel will become tacky and stick to the X-ray film. Second, urea quenches the low energy β− particles emitted from 35S, lowering the rate of film exposure. After soaking in the methanol/acetic acid solution, the gel/filter paper was prepared and dried as before. The gel/filter paper was placed into an autoradiography cassette followed by a piece of Kodak XAR-5 X-ray film (done in the darkroom) directly on top of the gel surface. The cassette was closed and placed
in a drawer to expose at room temperature for 1-3 days.

**Preparation of dihydrodiols (Kunz and Chapman 1981a)**

A single colony of *E. coli* strain DH5α carrying a recombinant *xylXYZ* plasmid was inoculated from an Amp\(^{50}\) LB agar plate into 5 ml of LB medium containing 50 µg/ml ampicillin. This culture was incubated at 37°C overnight shaking at 250 rpm. Five-hundred microliters of the overnight culture was inoculated into a flask containing 50 ml of LB medium and 50 µg/ml ampicillin. The flask was incubated as above until the absorbance reached 0.5 at 550 nm. The culture was transferred to a sterile centrifuge tube and centrifuged at 1000 x g in a Dupont RT600 swinging bucket benchtop centrifuge. The supernatant was removed and the resulting bacterial pellet was resuspended in 10 ml of *E. coli* minimal medium. Once again the tube was centrifuged as before and the bacterial pellet was resuspended in 1.0 ml of minimal medium. After transferring the bacterial suspension to a flask containing 50 ml of minimal medium, the flask was incubated at 37°C for 30 minutes and 250 rpm. Upon completion of this 30 minute incubation, meta-toluate was added to a final concentration of 5 mM. The resulting suspension was incubated at 37°C and 250 rpm. The production of dihydrodiol was measured spectrophotometrically by scanning the absorbance of the supernatant from this culture. One milliliter samples were taken at regular intervals and centrifuged for 2 minutes at 10,000 x g. The supernatants
were diluted (1:20) and the absorbance scanned from 220 nm to 320 nm. Upon completion, the remaining culture was transferred to a 50 ml conical tube and centrifuged for 15 minutes at 2000 x g. The dihydrodiol containing supernatant was stored in 1 ml aliquots at -20°C for use in enzymatic assays.

**Preparation of cell-free extracts (Kunz and Chapman 1981a)**

A single colony of *E. coli* strain DH5α carrying the recombinant plasmid for study was inoculated into 5 ml LB broth containing 50 µg/ml ampicillin. The tube was allowed to incubate overnight at 37°C and 250 rpm. The next morning, the contents of the tube were transferred to a 2 liter Erlenmeyer flask containing 500 ml of LB broth and 50 µg/ml ampicillin. The flask was incubated at 37°C and 250 rpm until the absorbance reached 0.7 at 550 nm. The contents of the flask were transferred to a 500 ml centrifuge tube and centrifuged at 9268 x g for 20 minutes. The supernatant was removed and the pellet was resuspended in 25 ml of SP-50 buffer (0.03 M Na₂HPO₄ and 0.02 M KH₂PO₄). The cell suspension was transferred to a preweighed Oak Ridge™ style centrifuge tube and centrifuged at 9268 x g for 10 minutes. The supernatant was discarded and the centrifuge tube with the cell pellet was weighed. The cell pellet was resuspended in a volume of SP-50 buffer equal to two times the weight of the cells in grams. The cell suspension was then transferred
to a French Press cell and pressed at 1000 lbs/in². The extract was slowly removed from the French Press and placed into an Oak Ridge™ style T1270 ultracentrifuge tube. A few crystals of pancreatic DNase I was added and the extract was incubated at room temperature for 10 minutes. The tube was centrifuged at 30230 x g for 30 minutes in an ultracentrifuge. The cell-free supernatant containing expressed enzyme product(s) was transferred to a 1.5 ml centrifuge tube and was placed on ice for immediate use or at -70°C until ready for use.

**Modified Lowry procedure for determining protein concentration (Lowry et al. 1951)**

The Lowry method for the determination of protein concentration is sensitive to detect as little as 0.2 µg of protein. This procedure involved the use of two reagents, alkaline copper reagent and Folin reagent. In order to make the alkaline copper reagent, 0.5 ml 2% NaK Tartrate was added to a tube containing 0.5 ml 1% CuSO₄·5H₂O. The solution was mixed well and 49 ml of 2% Na₂CO₃ in 0.1 N NaOH was added. Folin reagent (1 N phenol) was prepared by adding equal amounts of water and 2 N phenol (Folin reagent). Standard tubes, using 1 mg/ml bovine serum albumin (BSA), were set up as follows:

<table>
<thead>
<tr>
<th>µg</th>
<th>µl ddH₂O</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>190 + 10 µl 1 mg/ml BSA</td>
</tr>
</tbody>
</table>
A 1:10 dilution of the cell-free extract was made with ddH$_2$O and unknown tubes were set up as follows:

- 5.0 μl of 1:10 unknown + 195 μl ddH$_2$O
- 10.0 μl of 1:10 unknown + 190 μl ddH$_2$O
- 20.0 μl of 1:10 unknown + 180 μl ddH$_2$O

At timed intervals, 0.8 ml of alkaline copper reagent was added to each tube. Tubes were incubated at room temperature for ten minutes. While vortexing, 0.1 ml of 1 N Folin reagent was added to each tube once again at timed intervals. The tubes were incubated at room temperature for 30 minutes. After the final incubation, absorbance values at 660 nm were taken. The standard values were graphed and estimations of the total protein concentrations of the cell-free extracts were made. An example of the standard curve can be seen in Fig. 12.
Absorbance 660 nm

µg Bovine Serum Albumin
Fig. 12. Sample standard curve used for the Lowry method of determining protein concentration.
1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase

\[ \text{NAD}^+ \rightarrow \text{NADH} + \text{H}^+ \]

\( \text{3,5-cyclohexadiene-1,2-diol-1-carboxylic acid} \)

Catechol-2,3-dioxygenase

\( \text{Catechol} \rightarrow \text{2-hydroxymuconic semialdehyde} \)
Fig. 13. Enzymatic reactions catalyzed by DHCDH and C230. The *xylL* region encodes the enzyme 1,2-dihydrocyclohexa-3,5-diene carboxylate dehydrogenase (DHCDH), while the *xyle* region encodes the enzyme catechol-2,3-dioxygenase (C230).
Enzyme assays (Kunz and Chapman 1981a)

All enzymatic reactions were done at room temperature in quartz cuvettes using a Perkin-Elmer dual beam spectrophotometer. The activity of 1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase was determined by spectrophotometrically monitoring the formation of NADH from NAD$^+$ at 342 nm (Fig. 13). The enzymatic reaction mixture was prepared as follows:

<table>
<thead>
<tr>
<th>Solution</th>
<th>Reference Cuvette</th>
<th>Experimental Cuvette</th>
</tr>
</thead>
<tbody>
<tr>
<td>diol (5 mM)</td>
<td>0.0 µl</td>
<td>100.0 µl</td>
</tr>
<tr>
<td>tris (50 mM, pH 8.0)</td>
<td>500.0 µl</td>
<td>500.0 µl</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>300.0 µl</td>
<td>200.0 µl</td>
</tr>
<tr>
<td>NAD$^+$ (10 mM)</td>
<td>200.0 µl</td>
<td>200.0 µl</td>
</tr>
<tr>
<td>total volume</td>
<td>1000.0 µl</td>
<td>1000.0 µl</td>
</tr>
</tbody>
</table>

No more than 10 µl of diluted cell-free extract was added to the above mixtures.

Catechol-2,3-dioxygenase (C230) activity was measured by monitoring the formation of ring fission product 2-hydroxymuconic semialdehyde spectrophotometrically at 375 nm (Fig. 13). The enzymatic reaction mixture was prepared as follows:

<table>
<thead>
<tr>
<th>Solution</th>
<th>Reference Cuvette</th>
<th>Experimental Cuvette</th>
</tr>
</thead>
<tbody>
<tr>
<td>catechol (10 mM)</td>
<td>0.0 µl</td>
<td>33.0 µl</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>33.0 µl</td>
<td>0.0 µl</td>
</tr>
<tr>
<td>KH$_2$PO$_4$ (pH 7.5)</td>
<td>967.0 µl</td>
<td>967.0 µl</td>
</tr>
</tbody>
</table>
Once again, no more than 10 μl of diluted cell-free extract was added to the above mixtures.

**Site-directed mutagenesis**

Site-directed mutagenesis was performed using the Transformer™ Site-Directed Mutagenesis Kit from Clontech Laboratories, Incorporated. The kit was based upon the method developed by Deng and Nickoloff, 1992, as a method of introducing specific mutations into almost any double-stranded plasmid. This method offered several advantages over other conventional methods of mutagenesis, since it did not require single-stranded vectors, viral transductions or additional subcloning. The only prerequisites for this method are that the vector must contain a unique restriction endonuclease site and a selectable marker, such as antibiotic resistance. This method required annealing two oligonucleotide primers at the same time to one strand of a double-stranded plasmid. The mutagenic primer introduced the desired mutation in the particular area of interest, while the selection primer introduced a mutation in the unique restriction endonuclease site rendering it unrecognizable. The overall strategy of the Transformer™ Site-Directed Mutagenesis Kit is shown in Fig. 14.

The first step was to denature 100 ng of plasmid DNA using the alkaline method described earlier for denaturing
1. Denature dsDNA

2. Anneal Primers

3. Synthesize second strand with T4 DNA Polymerase and seal gaps with T4 DNA Ligase; primary digestion with selection restriction enzyme.

4. Transform mutS E. coli
   FIRST TRANSFORMATION

5. Isolate DNA from transformant pool

6. Restrict DNA with selection enzyme

7. Transform E. coli
   SECOND (FINAL) TRANSFORMATION

8. Isolate DNA from individual transformants to confirm presence of desired mutation.
Fig. 14. Transformer™ site-directed mutagenesis strategy. The unique characteristic of the Transformer™ Mutagenesis strategy is that in addition to the mutagenic primer, it employs an additional selection primer that mutates a unique restriction site. The two oligonucleotides are simultaneously annealed to one strand of the denatured double-stranded plasmid. After standard DNA elongation, ligation and primary selection by digesting with the restriction endonuclease for the second primer, the mixture of mutated and unmutated DNAs are transformed into a mutS E. coli strain defective in mismatch repair. Transformants are pooled and plasmid DNA is isolated from the mixed bacterial population. The DNA is digested with the selective restriction endonuclease, which the mutated DNA is immune. The parental DNA is sensitive to digestion and will be linearized. A final transformation using the digested DNA will result in recovery of the desired mutated plasmids (Clonetech product protocol instructions).
double-stranded DNA for sequence analysis. A many-fold (100+) excess of oligonucleotide primers was used to minimize the parental DNA rehybridization. 100 ng of each phosphorylated oligonucleotide primer along with ddH$_2$O to total 18 μl was added to the dried denatured DNA pellet. The solution was mixed well and 2 μl of 10X annealing buffer was added. The tube was incubated in a 37°C water bath for 10 minutes, followed by a room temperature incubation for 10 minutes. At this point both primers had hybridized to the same strand of the DNA duplex. The next step in the reaction was DNA elongation of the primers and ligation. To the primer/plasmid annealing reaction, the following was added:

3.0 μl  10X Synthesis Buffer  
1.0 μl  T4 DNA polymerase (2-4 units/μl)  
1.0 μl  T4 DNA ligase (4-6 units/μl)  
5.0 μl  ddH$_2$O

The tube was mixed well and placed in a 37°C water bath for 1 hour. The elongation and ligation enzymes were inactivated by placing the tube in a 70°C water bath for 5 minutes. After the tube cooled to room temperature, primary selection of the mutants was done by restriction endonuclease digestion. This digestion selectively linearized any rehybridized parental DNA and thus increased the percentage of mutant plasmids present in the pool for the first transformation. The restriction endonuclease digestion was
prepared as follows:

<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0 μl</td>
<td>10X Annealing Buffer (100 mM NaCl final concentration for Afl III)</td>
</tr>
<tr>
<td>2.0 μl</td>
<td>20X BSA</td>
</tr>
<tr>
<td>2.0 μl</td>
<td>ddH₂O</td>
</tr>
<tr>
<td>1.0 μl</td>
<td>10 units of Afl III (unique restriction site in parental DNA)</td>
</tr>
</tbody>
</table>

The contents were mixed well and the tube was incubated in a 37°C water bath for 2 hours. The tube was placed at -20°C until ready for use.

The sample was now ready for the first transformation. The purpose of the first transformation was to allow replication/amplification of the mutated and parental strands in a repair-deficient mutS strain of *E. coli*. One-fourth (10 μl) of the digestion mixture was transformed into competent BMH mutS *E. coli* as described earlier. After heat shocking the cell/DNA mixture, it was placed into a sterile 5 ml tube of YT medium containing 50 μg/ml ampicillin. The tube was placed in a gyratory shaker overnight at 37°C and 250 rpm. The next morning a rapid plasmid preparation was completed using the alkaline lysis method (described earlier). DNA retrieved from this method was further purified using a SpinBind™ DNA Extraction Unit purchased from FMC. It was important to have DNA as pure as possible to increase the effectiveness of the final selective restriction endonuclease digestion. The success of finding mutants was directly correlated with the degree of restriction endonuclease
digestion of the parental plasmids. The final digestion was prepared as follows:

- 2.0 µl spin bound DNA
- 4.0 µl 10X Afl III buffer
- 2.0 µl 20X BSA
- 3.0 µl Afl III (30 units) restriction endonuclease
- 29.0 µl ddH₂O
- 40.0 µl total volume

The tube was placed in a 37°C water bath for 2 hours. An additional 10 units of Afl III was added and the tube was incubated for an additional 2 hours. After the restriction endonuclease digestion was complete, the tube was stored at -20°C until ready for use.

The purpose of the final transformation was to amplify and clone the mutated plasmid. 5 µl of digested sample was transformed into competent E. coli DH5α cells and plated onto YT agar plates containing 50 µg/ml ampicillin (described earlier). Probable success of the mutagenesis procedure was visibly evident in the control plates (control plasmid and primers were supplied) since the mutation caused a change in phenotype. Efficiencies of 60-70% were observed. If the control plates were promising, individual colonies from the experimental plates were transferred to master plates and further characterized by DNA sequence analysis over the mutated area.
CHAPTER III

RESULTS

**Nucleotide sequence determination**

Nucleotide sequence analysis of the *xylXYZ* region of the pDK1 TOL plasmid was expedited by the cloning of the meta-cleavage pathway from pDKR1 (Fig. 15) into the cloning vector pBR322. A *Hind*III fragment containing all 13 genes of the meta-pathway was cloned into the tetracycline resistance gene of pBR322 and the recombinant plasmid was given the name pBK489 (Azadpour 1991). A further subclone encoding most of the *xylXYZ* region was constructed by inserting a 4.2 kbp *Hind*III-*Kpn*I fragment from pBK489 into the cloning and expression vector pUC19 (Azadpour 1991). The resulting subclone was designated pBK789 (Fig. 16). Although pBK789 was initially believed to encode the entire *xylXYZ* region, DNA sequence analysis of the *xylZ* region revealed that pBK789 contained all of *xylXY* but only a portion of *xylZ* (Khedairy 1990).

A frozen culture of *E. coli* DH5α containing the plasmid pBK789 was obtained from Azadpour. A large scale preparation of pBK789 was performed followed by final purification on CsCl gradients. The purified plasmid was subjected to various restriction endonuclease digestions and the resulting fragments were subcloned into compatible restriction
**Fig. 15.** Genetic organization of the xyl genes found on the cointegrate pDK1 TOL plasmid pDKR1. The pDKR1 plasmid (Shaw and Williams, 1988) carries both the "upper" and "meta-cleavage" pathways of the pDK1 TOL plasmid isolated from *Pseudomonas putida* HS1 strain (Kunz and Chapman, 1981b).
* pBR322
4.4 kbp
Ap<sup>+</sup>

* pDKR1
100 kbp

* pUC19
2.7 kbp
Ap<sup>+</sup>

* pBK489
20 kbp

* pBK190
7.7 kbp
Ap<sup>+</sup>

* pBK789
6.9 kbp
Ap<sup>+</sup>

** HindIII**

** Kpnl**
**Fig. 16.** Cloning strategy for the construction of subclones pBK190 and pBK789. pBK789 contains all of *xylXY* and a portion of *xylZ*. pBK190 contains the entire *xylXYZ* region and a functional toluate-1,2-dioxygenase can be expressed under the *lac* operator/promoter region of the pUC19 cloning/expression vector (Azadpour 1991).
endonuclease sites found within the Ml3mpl8/19 multiple cloning sites. The resulting recombinant Ml3mpl8/19 phage were purified and the single-stranded templates were subjected to sequence determination by the Sanger dideoxyribonucleoside method (Sanger et al. 1977). A set of subclones was designed to encompass both strands of DNA such that the entire sequence of both strands was determined. The sequencing strategy used to achieve this result can be seen in Fig. 17. Custom oligonucleotide primers were used to complete the sequence of those subclones with inserted fragments of more than 400 bp in length. An example of a typical autoradiogram derived from a DNA sequencing gel is shown in Fig. 18. The complete nucleotide sequence of the xylXYZ region was obtained and can be seen in Fig. 19.

**DNA sequence analysis**

DNA sequence analysis of the xylXYZ region, using the program MacDNAsis™, revealed three individual open reading frames (ORF). The first ORF, designated xylX (1,353 bp; nucleotides 1681 to 3033 as numbered in Figure 19), encodes a predicted protein product of 51370 Da. The ATG translational start codon of the xylX ORF is preceded by a putative ribosome binding sequence, 5′-GGAGG-3′. The second ORF, designated xylY, starts at nucleotide 3037 and continues through nucleotide 3525. The 486 bp ORF has a deduced protein product of 19368 Da. The first two ORFs are only separated by the TAA termination codon of xylX. An identical
The portion within this box was previously determined by Khedairy, 1990.
Fig. 17. DNA sequencing strategy and restriction endonuclease cleavage map of the xylXYZ region. Subclones were generated such that the nucleotide sequence of both strands of the entire xylXYZ region could be determined. Custom oligonucleotide primers were used to complete the nucleotide sequence of those subclones longer than 400 bp in length. "Shotgun sequencing" was also employed using restriction endonuclease four base pair cutters Rsal, BstUI and HaeIII (not shown). "Shotgun" subclones allowed for the filling in of gaps, joining of sequences, and verification of sequences in areas prone to error (e.g. compressions). The consensus nucleotide sequence of xylXYZ region was determined by the comparison of both strands, where each strand was sequenced independently a minimum of two times (xylZ was previously determined by Khedairly, 1990).
Fig. 18. Autoradiograph of a representative 6% denaturing polyacrylamide DNA sequencing gel. Each "rung of the ladder" is exactly one nucleotide larger than the preceding one. By starting at the bottom and reading upward (a 5' to 3' direction), the nucleotide sequence is elucidated. Samples were loaded in the following order: dideoxyguanosine triphosphate, dideoxyadenosine triphosphate, dideoxythymidine triphosphate and dideoxycytidine triphosphate (GATC).
Fig. 19. Nucleotide and predicted amino acid sequence of the xylXYZ region on TOL plasmid pDK1. Important restriction endonuclease cleavage sites are denoted above the sequence cleavage site. Putative ribosome binding site sequences are underlined. The sequence shown also includes the meta-operon promoter (Pm), where important sequences are underlined as proposed by Gallegos et al. (1996b), and approximately 1600 base pairs of sequence upstream of the promoter. Numbered nucleotide positions refer to the numbers assigned in this figure, beginning with the HindIII site.
putative ribosome binding site (5′-GGAGG-3′) for xylY is located within xylX. The third ORF (xylZ) starts at position 3535 and continues through position 4542. The 1008 bp ORF encodes a predicted protein of 36256 Da. The translational start codon is preceded by the putative ribosome binding sequence of 5′-GAGGTG-3′.

The G+C content analysis of xylX, xylY, xylZ gave values of 58.9%, 56.9% and 60.9%, respectively. The average G+C content for the entire xylXYZ region was 58.9%. It is of interest to note that the remaining genes of the pDK1 TOL meta-cleavage operon have a significantly higher average G+C content of 64.5%.

The nucleotide sequence of the xylXYZ region was compared to other known sequences using the NCBI BLAST algorithm tool (Altschul et al. 1990). High nucleotide sequence homology was found between xylXYZ pDK1 and three other sequences; xylXYZ (toluate 1,2-dioxygenase) from the archetype TOL plasmid pWW0 (Harayama et al. 1991a), benABC (benzoate dioxygenase) from Acinetobacter calcoaceticus (Neidle et al. 1991) and cbdABC (2-halobenzoate-1,2-dioxygenase) from Burkholderia cepacia (formerly Pseudomonas) (Haak et al. 1995). Comparisons of these nucleotide sequences are shown in Figures 20, 21 and 22. Comparisons of the predicted encoded polypeptides are shown in Figures 23, 24 and 25.

As expected, nucleotide sequence comparisons between xylXYZ pDK1 and archetype xylXYZ pWW0 demonstrated the
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-50</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>51-100</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>101-150</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>151-200</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>201-250</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>251-300</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>301-350</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>351-400</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>401-450</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>451-500</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>501-550</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>551-600</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdA</td>
<td>1101</td>
<td>GCACGGATCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1102</td>
<td>GCCGCCGCTTA</td>
<td>TCAGGTCGTT</td>
<td>CAATGATGTT</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1107</td>
<td>GCCGCCGCTTA</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1107</td>
<td>GCCGCCGCTTA</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1160</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1161</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1167</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1167</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1206</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1207</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1207</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1207</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1260</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1261</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1261</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1261</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1313</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1314</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1314</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1314</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1365</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1366</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1366</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1366</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>cbdA</td>
<td>1414</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>benA</td>
<td>1415</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pWWO</td>
<td>1415</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
<tr>
<td>xylX pDKl</td>
<td>1415</td>
<td>GCCGGATGCC</td>
<td>GCCATATGCG</td>
<td>ACAACGGAGGG</td>
</tr>
</tbody>
</table>
Fig. 20. Comparison of the pDK1 xy1x nucleotide sequence with xy1x (pWW0), benA and cbdA. Nucleotide sequence homologies to xy1x pDK1 are 92.7% (xy1x pWW0), 56.9% (benA), and 55.2% (cbdA).
**Fig. 21.** Comparison of the pDK1 xy1Y nucleotide sequence with xy1Y (pWW0), benB and cbdB. Nucleotide sequence homologies to xy1Y pDK1 are 96.1% (xy1Y pWW0), 57.5% (benB), and 61.6% (cbdB).
<table>
<thead>
<tr>
<th></th>
<th>560</th>
<th>570</th>
<th>580</th>
<th>590</th>
<th>600</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>639</td>
<td>CACGCGGCCGC TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 688</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>651</td>
<td>GATGTCGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>542</td>
<td>GATGTCGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>591</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>542</td>
<td>GATGTCGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>591</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>610</th>
<th>620</th>
<th>630</th>
<th>640</th>
<th>650</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>589</td>
<td>CTATCCCGGC TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 638</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>601</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>552</td>
<td>GATGTCGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>552</td>
<td>GATGTCGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>641</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>660</th>
<th>670</th>
<th>680</th>
<th>690</th>
<th>700</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>739</td>
<td>GCATGCTGATT TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 788</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>751</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>742</td>
<td>TACGCTGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>791</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>742</td>
<td>TACGCTGATT GCATGCGAGT ATGCCTGATT Turbo</td>
<td>791</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>710</th>
<th>720</th>
<th>730</th>
<th>740</th>
<th>750</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>789</td>
<td>GCATGCTGATT TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 738</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>801</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>792</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 641</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>792</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 641</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>860</th>
<th>870</th>
<th>880</th>
<th>890</th>
<th>900</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>839</td>
<td>CTATCCCGGC TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 863</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>851</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>890</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>842</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 891</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>842</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 891</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>910</th>
<th>920</th>
<th>930</th>
<th>940</th>
<th>950</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>889</td>
<td>GTGCGCGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>901</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>950</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>892</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 941</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>892</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 941</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>960</th>
<th>970</th>
<th>980</th>
<th>990</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>839</td>
<td>GTGCGCGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 988</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>851</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>842</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 991</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>842</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 991</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1010</th>
<th>1020</th>
<th>1030</th>
<th>1040</th>
<th>1050</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>989</td>
<td>GCAATGAGCG TACGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 1038</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>1001</td>
<td>TAAATGCCTG ACCATGCTGATT Turbo</td>
<td>1050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>952</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 1041</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>952</td>
<td>GCATGCTGCGT AATAGCGACG GCAATGAGCG GTAATGAGCG 1041</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1060</th>
<th>1070</th>
<th>1080</th>
<th>1090</th>
<th>1100</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbdc</td>
<td>1039</td>
<td>AG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benC</td>
<td>1051</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pW0</td>
<td>1042</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xylZ pDK1</td>
<td>1042</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 22. Comparison of the pDK1 xylZ nucleotide sequence with xylZ (pWW0), benC and cbdC. Nucleotide sequence homologies to xylZ pDK1 are 96.6% (xylZ pWW0), 44.5% (benC), and 43.1% (cbdC).
Fig. 23. Comparison of the pDK1 XylX amino acid sequence with XylX (pWW0), BenA and CbdA. Predicted protein sequence homologies to XylX pDK1 are 93.0% (XylX pWW0), 63.2% (BenA), and 55.5% (CbdA).
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbdB</td>
<td>-4</td>
<td>.....MTSLES SYLDVCAAPFL REARLLDDDRS WDEWLECYDP EAVFWMPGWD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BenB</td>
<td>1</td>
<td>MNATALLDTI SIEQISQPFLY SEARFLDDEQ WDDWLECYAP QASFWMPAWD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XyY pWW0</td>
<td>-7</td>
<td>.....MTI SYEAVRDPFLY REARYLDDKQ WESWLEMYAP DATFWMPAWD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XyY pDK1</td>
<td>-7</td>
<td>.....MTI SYEAVRDPFLY REARYLDDKQ WESWLEMYAP DATFWMPAWD</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| CbdB | 47  | DADTLVDQDP HRVLSILLYS RQGLEDVFR LRSEBSGAST PEPRTNHNIA |
| BenB | 51  | DNDQLTEDPQ TEISLLYDP RQGLEDVFR IKTESSATM PDRRTSHDIS |
| XyY pWW0 | 44 | DRDQLTEDPQ SQISLLYWGN RQGLEDVFR IKTESSATI PDRRTSHDIS |
| XyY pDK1 | 44  | DRDQLTEDPQ SQISLLYWGN RQGLEDVFR IKTESSATI PDRRTSHDIS |

| CbdB | 97  | NVEILERTER QIRARRNHWG MNRYKDLDDH YFGTSFYVPLK VSSGSLSILN |
| BenB | 101 | NVELESRDLG QITVRFNMT LSERYKNSYS YFCSRVYVID FSGEQPKILS |
| XyY pWW0 | 94 | NLELLEQSDG VCKLEYNWHT MNRYKDHTD FFGTNPCITLD TCGEPLIIA |
| XyY pDK1 | 94  | NLELLEQSDG VCKLEYNWHT MNRYKDHTD FFGTNPCITLD TCGEPLIIA |

| CbdB | 147 | KKVLKNDPLI HQVIDVYHV |
| BenB | 151 | KYVNLKNDYI HQVIDIYHI |
| XyY pWW0 | 144 | KKVLKNDYI HQVIDVYHV |
| XyY pDK1 | 144 | KKVLKNDYI HQVIDVYHV |
Fig. 24. Comparison of the pDK1 XylY amino acid sequence with XylY (pWW0), BenB and CbdB. Predicted protein sequence homologies to XylY pDK1 are 93.3% (XylY pWW0), 60.1% (BenB), and 54.6% (CbdB).
...
**Fig. 25.** Comparison of the pDK1 XylZ amino acid sequence with XylZ (pWW0), BenC and CbdC. Predicted protein sequence homologies to XylZ pDK1 are 94.4% (XylZ pWW0), 54.0% (BenC), and 47.2% (CbdC). Conserved amino acid residues are indicated by asterisks. Possible cofactor binding domains are shown where amino acids in bold are believed to be critical in cofactor binding (Neidle et al. 1991, Correll et al. 1992, Haak et al. 1995). FIB, FAD-isoalloxazine ring-binding domain.
highest nucleotide sequence homology. Homologies of \textit{xylX}, \textit{xylY}, \textit{xylZ} were 92.7\%, 96.1\% and 96.6\%, respectively. Of the non-homologous nucleotides within the predicted codons for the \textit{xylXYZ} region, 59.4 \% (79 total) were third position differences, 15.0\% (20 total) were second position differences and 25.6\% (34 total) were first position differences. These 133 total base differences resulted in 61 predicted amino acid changes, where 15 of those were "equivalent" amino acid changes (K=R, I=L=V, S=T, and D=E). Predicted protein homologies were 93.0\%, 93.3\% and 94.4\% for \textit{XylX}, \textit{XylY} and \textit{XylZ}, respectively. It is worth noting in Figure 26 that the proposed deletion within \textit{xylX} p\textit{WWO} at nucleotide 1062 is followed by the two more deletions at nucleotides 1072 and 1073. The difference in nucleotide sequence suggests that the nucleotide sequence of \textit{xylX} (p\textit{WWO}) may be incorrect, or it may be result of a recent mutation due to the near absence of base substitutions in the out-of-frame amino acid codons between deletions.

Two other DNA sequences exhibited high nucleotide sequence homology to \textit{xylXYZ} p\textit{DK1}, the plasmid-encoded \textit{cbdABC} genes from \textit{Burkholderia cepacia} and the chromosomally-encoded \textit{benABC} genes from \textit{Acinetobacter calcoaceticus}. Homologies for \textit{xylX}, \textit{xylY}, \textit{xylZ} to \textit{cbdA}, \textit{cbdB}, \textit{cbdC} were 55.2\%, 61.6\% and 43.1\%, respectively. Predicted protein homologies for the same genes were 55.5\%, 54.6\% and 47.2\%, respectively. When "equivalent" amino acids are included, the level of amino acid sequence homology increase to 65.2\%, 64.8\% and
Fig. 26. Possible nucleotide deletions within XylX of pWWO. Due to the relatively high G+C content of these genes, it is not uncommon for artifacts to occur as result of the DNA sequencing reaction or the DNA sequencing gel. These artifacts often result in the accidental elimination or misplacement of nucleotides. Another possibility is that these differences may be the result of a recent mutation. In the nine nucleotides between the proposed deletions, all but one are conserved between pDK1 and pWWO.
57.2%, respectively. Homologies of xylX, xylY, xylZ to benA, benB benC were 56.9%, 57.5% and 44.5%, respectively. Predicted protein homologies for the same genes were 63.2%, 60.1% and 54.0%, respectively. When "equivalent" amino acids (e.g. I=L=V) are included for benABC, the levels of amino acid sequence homology increase to 68.5%, 66.7% and 62.2%, respectively.

The nucleotide sequence was also analyzed for potential regulatory signals. As seen in Figure 19, a non-coding gap of 116 bp exists between the end of xylZ and the beginning of xylL. This gap may be significant for several reasons. First, it is unusually large when compared to other non-coding gaps in the pDK1 meta-cleavage operon (Table 3). When the distance between the termination codon of one gene is compared to the ribosome binding site of the adjacent gene, seven of the twelve intergenic regions are two nucleotides or less, three are between 16 and 37 nucleotides, and only two are larger than one-hundred nucleotides in length (xylZ/xylL and xylL/xylT). The G+C content within this xylZ/xylL region is 52.2%, much lower than the overall average of 63.1% G+C of the pDK1 meta-cleavage operon. Within this xylZ/xylL region, a nucleotide sequence with substantial homology to an E. coli transcriptional terminator (Platt 1986) is observed. The observed sequence is predicted to form a G+C rich stem and loop (inverted repeat) structure followed by several U residues in the transcribed RNA as seen in Figure 27. This predicted structure is also conserved in the pWW0 TOL
Table 3. Intergenic spaces between genes of the pDK1 TOL plasmid. Numbers represent the distance between the termination codon of one gene and the ribomome binding site of the other.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Gene Spacing (Nucleotides)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xylX-xylY</td>
<td>-12</td>
</tr>
<tr>
<td>xylY-xylZ</td>
<td>-2</td>
</tr>
<tr>
<td>xylZ-xylL</td>
<td>103</td>
</tr>
<tr>
<td>xylL-xylT</td>
<td>127</td>
</tr>
<tr>
<td>xylT-xylE</td>
<td>16</td>
</tr>
<tr>
<td>xylE-xylG</td>
<td>21</td>
</tr>
<tr>
<td>xylG-xylF</td>
<td>-2</td>
</tr>
<tr>
<td>xylF-xylJ</td>
<td>-2</td>
</tr>
<tr>
<td>xylJ-xylQ</td>
<td>2</td>
</tr>
<tr>
<td>xylQ-xylK</td>
<td>-2</td>
</tr>
<tr>
<td>xylK-xylI</td>
<td>-18</td>
</tr>
<tr>
<td>xylI-xylH</td>
<td>37</td>
</tr>
</tbody>
</table>
(A)

T

T   G

C ≡ G
A ≡ T
T ≡ G
T ≡ A
C ≡ G
G ≡ C
C ≡ G
C ≡ G
T ≡ G
G ≡ C

[xyz]→ 5'CTCCTTTG≡CTGTTTTATT 3'
Fig. 27. Putative transcriptional terminator found immediately downstream of xylZ. It is located in the 116 bp non-coding gap between xylZ and the adjacent downstream gene in the meta-cleavage operon, xylL. The nucleotide in parenthesis is a non-homologous nucleotide found at the same position in the pWWO TOL plasmid. For comparison purposes, the structure is drawn as DNA.
plasmid, with only a single loop nucleotide substitution. A more detailed analysis of the possible functions of this predicted structure will be discussed later.

Immediately upstream of xylX lies the operator/promoter region for the meta-cleavage operon. Comparison of the nucleotide sequence of this region to the corresponding area of pWW0 revealed a 90% nucleotide sequence homology. Figure 28 shows the comparison between pDK1 and pWW0 and the important proposed regulatory regions. pDK1 contained the identical TACAN4TGCA motif upstream of the -35 region found to be the minimum DNA segment required for transcription stimulation by XylS in the $\sigma^{70} - \sigma^{S}$- dependent Pm promoter of pWW0 (Gallegos et al. 1996b). Interestingly, the region found to be the least significant in XylS stimulation of transcription (-78 to -75), contained two nucleotide differences out of a possible four. These differences were two of only three differences found within the proposed motif sequences located between -46 and -57 and -67 and -78 as suggested by Gallegos et al. (1996b).

Approximately 1600 bp of nucleotide sequence was determined upstream of the Pm promoter and was compared to other known nucleotide sequences using the BLAST algorithm tool (Altschul et al. 1990). Only a 131 bp segment of this region (nucleotides 740 through 870) was found to exhibit high homology (83%) with any other known sequence (Fig. 29). The function of the homologous sequence is not known, but the source organism is Pseudomonas putida (Genbank accession
pDK1

5'   CTGAGCTT[ ]T[ ]CA[ ]GACT[ ]GCCAAGAAAATGGCTATCTCTAGAAAGGCTAG

*   *   **   *   *******************                    

pWW0

5'   AGGGATAACL[ ]CCT[ ]GCCAAGAAGGCTAGCTAGAAAGGCTAC

-78 -73 -70 -67 -57 -54 -49 -46 -35 [ ]                     

XbaI

pDK1

CTCT'TAGGCT'TTATGCAACTGAACAAATAATTGGAAGCAAGAAATG 3'

*   *******************            **  **  ***

pWW0

CCCT'TAGGCT'TTATGCAACGAACAAATAATTGGAAGCTATGACCCTG 3'

-10            PBR  xy/IX--> +1
Fig. 28. Nucleotide sequence comparison of the meta-cleavage Pm promoter between TOL plasmids pDK1 and pWWO. pDK1 contained the identical TACAN$_4$TGCA motif upstream of the -35 region found to be the minimum DNA segment required for transcriptional stimulation by XylS of the $\sigma$70-$\sigma$S-dependent Pm promoter of pWWO (Gallegos et al. 1996b). Interestingly, the region found to be the least significant for XylS stimulation of transcription (-78 to -75), contained two nucleotide differences out of a possible four. These differences were two of only three differences found within the proposed motif sequences located between -46 and -57 and -67 and -78 as suggested by Gallegos et al. (1996b). This figure was adapted from Gallegos et al. (1996b) and Kessler et al. (1993).
Query:  740 ATTCACCGTGTGTGCAAACATCTACCTGCACTACGTGATGCGACTTTGTGGATCAGGCAA 799
       ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^
Sbjct: 19675 ATTCACCCACTGCTGCAATATTTACTGCTACGTGATGGATTTATGGTAAGACAA 19734

Query:  800 TGGCGCGAGCGTGACTGAGGCGTGCTAGTGATTTGCTGTGGGATACGACGACAGTGTG 859
       ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^
Sbjct: 19735 TGGCGTAGCGCTACGCGAAGGCGAATGATTGCGCTGCGGTACGACGACAGTGTGA 19794

Query:  860 CTTGGTTTCA 870
       ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^
Sbjct: 19795 CTTGGTTTGA 19805
Fig. 29. Comparison of the nucleotide sequence upstream of the Pm promoter from pDK1 to a *Pseudomonas putida* nucleotide sequence of unknown function. 109 of the 131 nucleotides nucleotides are homologous (83%). Query represents the pDK1 nucleotide sequence and Sbjct represents the *Pseudomonas putida* nucleotide sequence. The numbers for pDK1 correspond to the same numbers seen in Fig. 19. The numbers for *P. putida* correspond to the numbers found in Genbank accession number AF006691. This segment is found near the *ibp* operon for the degradation of isopropylbenzene in *P. putida*. 
number AF006691). This upstream nucleotide sequence was also analyzed for potential open reading frames (ORFs). No significant ORFs were found, the largest being only 234 bp. This small sequence shows no significant homology to any other known sequence. Thus, we conclude this rather large interoperonic segment has no known function at this time.

**Subcloning of pBK190 and expression of toluate 1,2-dioxygenase**

DNA sequence analysis of the pBK190 recombinant plasmid revealed that it included 1680 bp of DNA immediately upstream of xylX. Azadpour (1991) found conversion of 5mM benzoate, m-toluic, p-toluic, or 3,4-dimethyl benzoate to its corresponding dihydriodiol took a culture of E. coli carrying pBK190 an average of 36 hours. Since our ability to carry out this conversion is important for producing the pathway intermediates necessary for the further studies of subsequent meta-cleavage enzymes (particularly dihydriodiol dehydrogenase (XylL)), it was of interest to construct additional subclones capable of performing the conversion at a faster rate. A detailed restriction endonuclease map elucidated from nucleotide sequence analysis expedited the creation of three subclones containing various deletions of the upstream region (Fig. 30). pBK192 (SmaI/KpnI) and pBK292 (XbaI/KpnI) were created by subjecting pBK190 to HindIII digestion followed by ligation of a phosphorylated HindIII/SmaI or HindIII/XbaI oligonucleotide adapter. The resulting constructs were then
Fig. 30. Diagrammatic representation of the strategy used to construct pBK192, pBK292 and pBK392. A detailed restriction endonuclease map of the xylXYZ region and resulting toluate-1,2-dioxygenase subclones. These subclones were used to determine the effect of this upstream region, which included the pDkl TOL meta-cleavage operon operator/promoter region, on toluate-1,2-dioxygenase (TO) expression from the pUC19 lac promoter. The minimum DNA sequences required for transcription as described by Gallegos et al. (1996b) are not present in the clone pBK292, but are present in the clones pBK192 and pBK392. The pBK192 construct contains approximately 400 bp of upstream sequence, while pBK392 contains only a few extra base pairs of additional DNA upstream of the putative operator/promoter region. The fine structure of the resultant subclones was verified initially by restriction endonuclease cleavage and subsequently by direct double-stranded sequence analysis of rapid plasmid isolates.
treated with the appropriate restriction endonuclease (SmaI or XbaI) followed by ligation. Subclones were verified by agarose gel analysis of restriction endonuclease cleavage products. The construction of pBK392 was complicated by the presence of three PstI restriction endonuclease sites within the TOL portion of pBK190. First, pBK190 was subjected to HindIII digestion. This was followed by partial digestion with PstI. It was important to obtain products from the partial PstI reactions where the majority of cleavages were within the upstream sequence (particularly the one closest to xylX) and not within xylX. The sample from the time point most closely meeting these criteria was selected by analyzing products from various times of digestion using agarose gel electrophoresis to size the product distribution. The chosen sample was religated in the presence of a phosphorylated HindIII/PstI oligonucleotide adapter, deleting the region from HindIII to the cut PstI site.

These subclones were used to determine the effect of this upstream region, which included the pDK1 TOL meta-cleavage operon operator/promoter region, on toluate-1,2-dioxygenase (TO) expression from the pUC19 lac promoter. The minimum DNA sequences required for transcription as described by Gallegos et al. (1996b) are not present in the pBK292 construct, but are present in the clones pBK192 and pBK392. The pBK192 construct contains approximately 400 bp of upstream sequence, while pBK392 contains only a few extra base pairs of additional DNA upstream of the putative
operator/promoter region. The fine structures of the resultant subclones were verified initially by restriction endonuclease cleavage and subsequently by direct double-stranded sequence analysis of rapid plasmid isolates. Comparison of the expression of these three clones provided the necessary information to determine if there is any influence from the TOL operator/promoter region and other upstream sequences on the expression of TO from the pUC19 lac promoter in E. coli.

TO expressed from the lac promoter of pBK190 was previously shown to convert benzoic acid, 3,4-dimethyl benzoic acid, meta-toluic acid and para-toluic acid to their corresponding dihydrodiols (Azadpour 1991). For the purpose of comparing the rates of TO expression by these clones, only meta-toluic acid was used in this study. Overnight cultures of E. coli carrying pBK190, pBK192, pBK292 and pBK392 were collected by centrifugation. Bacterial pellets were weighed and equivalent number’s of cells from each culture were transferred to minimal media containing 5 mM meta-toluic acid. The cultures were incubated at 37°C and 1 ml samples were taken at regular intervals. The accumulation of dihydrodiol was monitored by scanning diluted cell-free 10,000 x g supernatants spectrophotometrically from 220 nm to 320 nm using a Beckman DU-40 spectrophotometer. Results showed that there was a 4- to 5-fold increase in the rate of dihydrodiol accumulation by strains carrying pBK192, pBK292 and pBK392 over that of the original clone pBK190 (see Figure
These results also showed that there was no apparent effect on transcription by the *Pseudomonas* TOL promoter of pDK1 in *E. coli*, as deletion of this had no substantial effect on the level of gene expression.

**Evolutionary possibilities of the TOL meta-cleavage operon**

The entire nucleotide sequence of the pDK1 TOL meta-cleavage operon has been completed by our laboratory. Thirteen open reading frames were identified corresponding to those seen in Figure 1. Nucleotide sequence analysis of this region revealed relatively short intergenic spaces between genes, with two notable exceptions, the region between *xylZ* and *xylL* (116 bp), and the region between *xylL* and *xylT* (139 bp) (Table 3). Further analysis of these two intergenic regions revealed the presence of palindromic sequences with the potential to form stem and loop structures (Fig. 32). The existence of these regions raises questions and insight to the possible evolutionary formation of the "modern" TOL meta-cleavage pathway.

Upon the completion of the nucleotide sequence of the *xylL* region, a small open reading frame between *xylL* and *xylE* was discovered (Baker 1992, Harayama and Rekik 1990). This open reading frame encoded a predicted protein size of 11.1 kDa and was given the name *xylT*. The function of the XylT protein had not yet been determined, but some potential functions were postulated. Harayama et al. (1991b) suggested...
Fig. 31. Relative levels of toluate-1,2-dioxygenase activity from pBK190, pBK192, pBK292 and pBK392 constructs. Accumulated dihydrodiol was measured spectrophotometrically by scanning the absorbance of samples from 220 nm to 320 nm using diluted cell-free supernatants. The absorption maximum for each of the samples in the spectra was 270 nm. These subclones express TO activity under the regulation of the pUC19 E. coli lac promoter. Structural details for each of the individual constructs are given in Figure 30.
(A)

T

\[ \text{[\text{xy}]L} \rightarrow 5' \text{CTCCTTTG} = \text{CTGTTTTATT} 3' \rightarrow \text{[\text{xy}]L} \]

\[
\begin{array}{c}
\text{C} \\
\text{G}
\end{array}
\]
\[
\begin{array}{c}
\text{T} \\
\text{A}
\end{array}
\]
\[
\begin{array}{c}
\text{G} \\
\text{T}
\end{array}
\]

\[
\begin{array}{c}
\text{A} = \text{T} \\
\text{C} = \text{G} \\
\text{C} = \text{G} \\
\text{C} = \text{G} \\
(\text{-}) = \text{G} (\text{-}) \\
\text{A} = \text{G} \\
\text{C} = \text{G} \\
\text{T} = \text{A} \\
\text{A} = \text{T} \\
\text{C} = \text{G} (\text{A}) \\
\text{G} = \text{C} \\
\text{G} = \text{T} \\
\text{A} = \text{T} \\
\text{(G)} = \text{C} \\
\text{G} = \text{C}
\end{array}
\]

\[
\begin{array}{c}
\text{\text{xy}L} \rightarrow 5' \text{CGGGGTGCG} = \text{GTCATGTTTAAGGCATCCTGAGGTGAACCATG} \]
\]
Fig. 32. Putative stem and loop structures found in the xylZxylL and xylLxylT intergenic regions. The nucleotides in parenthesis are non-homologous nucleotides found at the same positions in the pWWO TOL plasmid (Baker 1992). For comparison purposes, the structures are drawn as DNA.
the function of XylT to be similar to chloroplast-type ferredoxins because of a characteristic Cys-XXXX-Cys-XX-Cys motif found in each case. Another possibility was one which was based upon the observations made in our laboratory by Ron Baker (1992). He observed two palindromic nucleotide sequences located in the intergenic regions flanking either side of the xylL gene. The first palindromic sequence, found in the xylZ/xylL intergenic region, is the one described earlier and is predicted to form a structure similar to an E. coli rho-independent transcriptional terminator (Fig. 27). Through the utilization of a series of subclones which included or excluded various areas within the xylLTE region, Baker obtained evidence suggesting that a transcriptional terminator may in fact be located immediately upstream of xylL and a positive acting element appeared to be present in the intergenic region downstream of xylL (see Figure 33 for clone diagrams and XylE specific activities). Enzymatic analyses showed that E. coli strains carrying the xylTE (pBK189) subclone expressed catechol-2,3-dioxygenase (C230) activity at levels five to eight times the C230 activity of E. coli strains carrying the xylE (pBK591) subclone (Baker 1992, and personal communication). The xylLTE (pBK692) subclone, which includes the xylZ/xylL intergenic (terminator) region, was shown to possess as little as 10% of the C230 activity of the xylE (pBK591) subclone (Baker 1992). These preliminary results were variable, but showed a trend that E. coli strains carrying xylTE expressed C230 activity
Cloned Genes                  C230 Specific Activity
                           nmol/min/mg protein

  xylLTE (pBK692)         223
  xylTE (pBK189)          25435
  xylE (pBK591)           3044

R. Baker (dissertation, 1992)
Fig. 33. Diagrammatic representation of TOL constructs pBK692, pBK189 and pBK591, and their expressed levels of C230 activity. Cultures of *E. coli* carrying plasmids pBK692 (xylLTE), pBK189 (xylTE), and pBK591 (xylE) were grown for 48 hours in LB medium and collected by centrifugation. Cell lysates were made using a French Press at 1000 lbs/in². Expression of cloned genes was controlled by the lac promoter found within pUC 19. Possible RNA secondary structure regions are designated by stem and loop.
several fold higher than strains carrying \textit{xylE}, and \textit{E. coli} strains carrying \textit{xylE} expressed C230 activity several fold higher than strains carrying \textit{xylL}\textit{TE}. He also observed no synergistic effect on C230 activity when an extract of \textit{xylT} was added to an extract of \textit{xylE}. The \textit{xylT} gene product may have some direct influence on gene expression within this region possibly acting as a type of internal regulator either at the level of transcription or translation, and a transcriptional terminator may be present in the \textit{xylZ}/\textit{xylL} intergenic region.

Combining these findings with a previous observation of weak promoter activity upstream of the \textit{xylE} gene (Keil et al. 1987), known as the \textit{xylLT} region, raised the possibility that one or more functional regulatory elements may be present in this region. Comparisons of the TOL pathway to other related aromatic catabolic pathways (Harayama and Rekik 1990, Platt et al. 1995) suggested the meta-cleavage operon could be the product of a fusion between two DNA modules. These transcription units would have been \textit{xylXYZL} and \textit{xylTEGFJKLM}, where each unit originally possessed its own promoter/operator regions.

Analysis of the G+C content within the TOL pDK1 meta-cleavage operon further supports the fusion hypothesis. As seen by the difference in G+C contents below (Table 4), the pDK1 meta-cleavage operon may have been formed from two or three separate expression units, at least one of which may have evolved in a different genetic background. The genes
encoding \textit{xylXYZ} are 58.9\% G+C while the genes encoding \textit{xylTEGFJQKIH} are 64.5\% G+C (Table 4).

\textbf{Table 4.} G+C content comparisons between \textit{xylXYZ} and the rest of the pDK1 TOL meta-cleavage operon.

<table>
<thead>
<tr>
<th>Gene</th>
<th>% G+C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pm operator/promoter region</td>
<td>39.1%</td>
</tr>
<tr>
<td>\textit{xylXYZ}</td>
<td>58.9%</td>
</tr>
<tr>
<td>\textit{xylZ/L} intergenic region</td>
<td>52.2%</td>
</tr>
<tr>
<td>\textit{xylL}</td>
<td>64.5%</td>
</tr>
<tr>
<td>\textit{xylL/T} intergenic region</td>
<td>53.0%</td>
</tr>
<tr>
<td>\textit{xylTEGFJQKIH}</td>
<td>64.5%</td>
</tr>
</tbody>
</table>

The first fusion could have occurred between \textit{xylL} and \textit{xylTEGFJQKIH}, followed by a second fusion of \textit{xylXYZ} to \textit{xylLTEGFJQKIH}. \textit{xylL} does not exhibit any difference in G+C content between itself and the G+C content of the remaining genes within the meta-cleavage operon. Therefore, this fusion could have occurred much earlier in the evolution of the pathway or the genes may simply have evolved in a similar genetic background. The \textit{xylXYZ} region, on the other hand, demonstrates a rather large difference in the G+C content from the remainder of the pathway and may have been added to the pathway more recently and from a different genetic background. Creation of this operon by fusions may also explain why the only two intergenic regions of any size (over 100 bp in this case) within the operon are the ones between
xylZ/L and xylL/T. Some "original" regulatory regions may have been included in the final form of these "modern" operons and any of these might still function. We therefore decided to further analyze these possible regulatory areas using the technique of site-directed mutagenesis.

Site-directed mutagenesis of possible regulatory elements

The initial area addressed using site-directed mutagenesis was the xylT region. At the time this research was initiated, the function of xylT was not known. Why did a clone which included xylT and xylE express C230 activity five to eight times greater than xylE alone, especially when XylT is not required for the conversion of toluate to Kreb's cycle intermediates? What function, if any, does the 139 bp non-coding intergenic region between xylL and xylT have? If XylT was somehow responsible for such an enhancing effect either at the level of transcription or translation, then a site-directed mutation within its proposed metal binding motif or creating a nonsense mutation within xylT might provide insight to the function of XylT. The complete nucleotide sequence of xylT along with two separate site-directed mutations introduced can be seen in Figure 34.

E. coli strains carrying plasmids pBK692 (xylLTE), pBK189 (xylTE) and pBK591 (xylE) were obtained from Ron Baker. A large scale plasmid preparation, followed by two CsCl gradient purifications, was completed on each strain.
**xyIT**

\[ \text{(A)} \]

1. ATGCAGACAGTTATGAGGTGCCGTAGCAGGATAGCCGAAAAATGGTTGGCCCTGCTGCCG
   MetAspSerSerTyrGluValArgGluArgIleSerGlyGlnValPheArgCysLeuPro
   (stop)

2. GAGCAGTGCTCCTGCCGCCCATGGAGGACAGGGAACGGTGCCGTCCGCGGTCGGTTGC
   GluGlnSerValLeuArgAlaMetGluGlnGlyLysArgCysValProValGlyCys

\[ \text{(C)} \]

3. CGTGCTGTCGCTGGGGTTCGCTCAAGGTGGCGTCTCAGCGGCGACTACCAGTGCGGC
   (Arg)

4. AGGATGAGCTAGCTACGTCGCTCTCCACCCGAGCCGCGAAGCAAGGCTCTGGGCTGGCTGT
   ArgMetSerCysSerGlnValProProGluAlaAlaAlaLysGlnGlyLeuAlaLeuAlaCys

5. CAACTGTATCCAGCGTCACTCAGCTGCAATGCGCTCCGCTAGTGGACAAGAATCCC
   GlnLeuTyrProArgAlaAspLeuTyrIleGluCysLeuArgGlnValArgThrAsnPro

6. TGA
   ***
Fig. 34. Nucleotide and predicted amino acid sequence of xylT. The underlined amino acids show the conserved Cys-XXX-Cys-XX-Cys motif described as a possible the metal binding site by Harayama et al. (1991b). Nucleotides and amino acids in bold were those targeted for site-directed mutagenesis. Nucleotides and amino acids located above and below the wild-type sequence in parenthesis indicate the mutation incorporated. The first cysteine, amino acid number 18 in the sequence, was changed to a stop codon in pBK189 (TE) and was designated pTEmut3. The last cysteine in the conserved motif (amino acid number 48) was changed to an arginine residue in pBK189 (TE) and was designated pTEmut5. This figure was adapted from Baker (1992).
Verification of the recombinant plasmids was done by restriction endonuclease cleavage followed by agarose gel electrophoresis.

Site-directed mutagenesis was accomplished based upon the method developed by Deng and Nickoloff (1992). The Transformer™ Site-Directed Mutagenesis Kit produced by Clonetech Laboratories, Incorporated was used. The key advantage to this method was that the mutagenesis could be done using the previously constructed clones, therefore eliminating the necessity of subcloning into another vector (see Fig. 14 for strategy). Specific oligonucleotide primers were designed to incorporate the desired single base substitutions described in Figure 34. The nucleotide sequences of all oligonucleotide primers used in this study can be seen in Table 5. For additional ease in identification of mutants, the unique restriction endonuclease site primer was designed to change the unique AflIII site in the cloning and expression vector pUC19 to another unique restriction endonuclease site, BglIII. If a resulting plasmid is sensitive to the restriction endonuclease BglIII, then at least one of the primers was incorporated into the new construct and quite possibly the desired mutation is also present. Another advantage was if additional mutations were desired, a unique site within the vector was still available and the process could be repeated.

The first three target mutations were to create a nonsense mutation early within xylT, a missense mutation
Table 5. Nucleotide sequence of oligonucleotide primers used with the Transformer™ Mutagenesis Kit. The nucleotides in parentheses represent the wild-type nucleotide sequence at that location. The nucleotides in bold represent the mutagenic nucleotide. No successful mutant constructs were created with the primers mut2, mut4 and mut8.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Nucleotide Sequence</th>
<th>Region Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>mut1</td>
<td>CAGGAAAGAAGATCTGAGCAAAG</td>
<td>pUC18/19 unique AflIII-&gt;BglII</td>
</tr>
<tr>
<td>mut2</td>
<td>GTGTCCGTTGAGCTGCCGGAG (C)</td>
<td>premature stop codon cys18 (xylT)</td>
</tr>
<tr>
<td>mut3</td>
<td>CTCCGGCAGCTCAACGGAACAC (C)</td>
<td>premature stop codon cys18 (xylT)</td>
</tr>
<tr>
<td>mut4</td>
<td>GCACAACCCACGGCACCACAC (A)</td>
<td>2nd cys in motif to arg (xylT)</td>
</tr>
<tr>
<td>mut5</td>
<td>CGCACCTTGCGCAACCACAGC (A)</td>
<td>3rd cys in motif to arg (xylT)</td>
</tr>
<tr>
<td>mut6</td>
<td>CACCAAGTAAGCAAAACAGAGGCC (GG)</td>
<td>stem and loop between xylZ/xylL</td>
</tr>
<tr>
<td>mut7</td>
<td>GGAAGCATCCCTTCAACAGCTCTGG (CC)</td>
<td>stem and loop between xylL/xylT</td>
</tr>
<tr>
<td>mut8</td>
<td>CCAAGTAAGCGGATTAAAAGGCCCTC (CC)</td>
<td>stem and loop between xylZ/xylL</td>
</tr>
<tr>
<td>mut9</td>
<td>GCTCACCAGTAAAATGGACCAAAGGAGGCC (GC)</td>
<td>stem and loop between xylZ/xylL</td>
</tr>
</tbody>
</table>
within the second cysteine (Cys→Arg) of the conserved CysXXXXCysXXCys residues within xylT and a mutation in the palindromic stem and loop structure upstream of xylT in the clone pBK189 (xylTE) (see Fig. 34). The mutagenic primer designed to create the nonsense mutation was given the name mut3, the primer for the missense mutation was named mut5 and the primer for the palindromic stem and loop structure was named mut7. Colonies obtained from the final transformation in the Transformer™ mutagenesis procedure (described earlier) were transferred to master plates and plasmid DNA was isolated from each using the Alkaline-Lysis method. Samples were then subjected to restriction endonuclease digestion with BglII followed by 1% agarose gel electrophoresis. Samples sensitive to BglII digestion (see Figure 35) were further characterized by DNA sequence analysis of the target area. Twelve colonies each from xylTEmut3, xylTEmut5, and xylTEmut7 were analyzed for mutations. Six plasmids from xylTEmut3, eight from xylTEmut5, and 9 from xylTEmut7 transformants were sensitive to BglII digestion. Three plasmids from each were subjected to DNA sequence analysis. Subclone number three for xylTEmut3 contained the desired mutation and was designated pTEmut3 (Fig. 36). The remaining samples for xylTEmut5 and xylTEmut7 were subjected to DNA sequence analysis in which clone number 24 and clone number 30 contained the targeted mutations and were designated pTEmut5 and pTEmut7, respectively (Fig. 36 and Fig. 37). A diagrammatic
Fig. 35. Agarose gel showing sensitivity of mutant clones to BglII restriction endonuclease digestion. Samples were electrophoresed in 1X TBE (pH 8.0) buffer for 90 minutes at 80 V. The size marker λDNA digested with HindIII was placed in lane nine. Identical DNA samples were loaded side by side beginning in lane one in which the left lane was uncut DNA and the right lane was DNA subjected to BglII digestion. An example of a subclone sensitive to BglII digestion can be seen in lanes 1 and 2. An example of a subclone not sensitive to BglII digestion can be visualized in lanes 12 and 13.
Fig. 36. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence changes for pTEmut3 and pTEmut5. Arrows designate the nucleotide changed by site-directed mutagenesis as compared to the wild-type nucleotide sequence. Refer back to Figure 34 for exact placement within xylT.
representation of the proposed stem and loop structure between \textit{xyl}L and \textit{xyl}T, and the position of the nucleotides changed as result of the mut7 primer can be seen in Fig. 38.

**Enzyme assays**

\textit{E. coli} DH5\(\alpha\) strains carrying various wild-type and mutant subclones were cultured in LB medium and cell-free extracts were made as described in Materials and Methods. Assays were performed for C230 activity as described earlier by monitoring the conversion of catechol to ring fission product at 375 nm. An example of such an assay can be seen in Figure 39. Protein concentrations of cell-free extracts were calculated using the Lowry method (Lowry \textit{et al}. 1951). The combined results of eleven different experiments can be seen in Table 6.

As seen from the data in Table 6, it was difficult to reproduce the results originally obtained by Baker (1992). The initial four trials were completed using the same conditions as described by Baker. Of those four experiments, only the first attempt approximately reproduced the initial results observed by him, and even then, the same \textit{E. coli} strain carrying pBK189 (\textit{xyl}TE) produced C230 activity only 3 times that of one carrying pBK591 (\textit{xyl}E), not the reported 5 to 8 times more. In the other three instances, the strain carrying pBK189 (\textit{xyl}TE) produced 2 to 10 times less C230 activity than the one carrying pBK591 (\textit{xyl}E). These data prompted attempts to better control the growth conditions of
Fig. 37. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence change for pTEmut7. Arrows designate the nucleotides changed by site-directed mutagenesis as compared to the wild-type nucleotide sequence. This sequence is located within the 139 bp intergenic region between immediately preceding xyIT.
Free-energy predictions (ΔG)

pTE(wt)  -23.5 kcal/mol
pTEmut7  -17.2 kcal/mol
Fig. 38. Putative xylLxylIT intergenic region stem and loop structure and the calculated free-energy predictions for wild-type and mutant constructs. The nucleotides changed as result of the mutagenic primer mut7 are boxed. Nucleotides in parentheses are non-homologous nucleotides located at identical positions within the TOL plasmid pWWO. Free-energy predictions were calculated using improved parameters set forth by Freier et al. (1986). This figure was adapted from Baker (1992). For comparison purposes, the structure is drawn as DNA.
Fig. 39. Representative C230 assay. Catechol 2,3-dioxygenase (C230) activity was monitored spectrophotometrically by determining hydroxymuconic semialdehyde formation (375 nm). Reactions were completed in a 1.00 ml total volume containing 48.35 mM KH$_2$PO$_4$ and 0.33 mM catechol. No more than 10 µl of undiluted or diluted cell-free extract was added to each reaction.
Table 6. C230 activity table showing the specific activities (nmol/min/mg protein), slopes of the assay (absorbance/sec), and the absorbance of the cell cultures at time of collection. Conditions refer to the growth medium, how long they were cultured and whether the assays were performed with undiluted or diluted cell-free extracts. No data were available for those areas left blank.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Date</th>
<th>Lag phase growth (48 hours)</th>
<th>LB + 30mM Glucose grown to specific Absorbance, no dilution of cell-free extracts</th>
<th>Cell culture Abs.</th>
<th>Specific activity</th>
<th>Slope</th>
<th>Abs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>2/4</td>
<td>4900</td>
<td>0.0062 0.011 0.033 0.022 0.051 0.11 0.15 0.011 0.014 0.048 0.022</td>
<td>0.72</td>
<td>0.74 0.78 0.70 0.66 0.62 0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/7</td>
<td>8170</td>
<td>0.0068 0.0272 0.0621 0.0068 0.011 0.040 0.0000 0.029 0.0000 0.0004 0.048 0.022</td>
<td>0.76</td>
<td>0.79 0.76 0.66 0.60 0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/10</td>
<td>20345</td>
<td>0.0005 0.0061 0.014 0.0065 0.017 0.079 0.030 0.001 0.0003 0.0000 0.033 0.022</td>
<td>0.72</td>
<td>0.79 0.70 0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/14</td>
<td>17561</td>
<td>462 295</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/28</td>
<td>368</td>
<td>243 2817 9898 5911 374 462 295</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T non-sense</td>
<td>3/3</td>
<td>1002</td>
<td>200 6962 2628 595 303 375 356 459</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/9</td>
<td>1040</td>
<td>595 609 2628 595 303 375 356 490</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T mis-sense</td>
<td>3/11</td>
<td>102</td>
<td>143 60 2628 595 303 375 356 459</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stem pTE</td>
<td>3/25</td>
<td>102</td>
<td>0.023 0.0078 0.27 0.14 0.68 0.66 0.60 0.064</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/31</td>
<td>0.69 0.66 0.60 0.064</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the cultures. Since the initial four trials were performed on cultures grown late into log or early stationary phase (48 hours), the next seven trials were conducted on cultures still in the mid-log phase of growth. Absorbance of cultures was monitored spectrophotometrically at 550 nm to a reading between 0.6 and 0.8 absorbance units. In only one of these seven trials did the strain carrying xylTE produce C230 activity greater than that of the strain carrying xylE, and it was only two times greater. The other six experiments varied from 4 to 168 times less C230 activity in xylTE when compared to xylE. This is exactly opposite to the previous findings. Due to the inconsistency of the data obtained, it is not plausible to draw any conclusions about the effect of mutations within xylT or the palindromic intergenic region upstream of xylT on the production of C230 activity.

Additional literature searches prompted the discovery of an article published three months earlier by Polissi and Harayama (1993). They discussed a novel mechanism to expand TOL substrate specificity by a XylT-dependent regeneration of inactivated catechol 2,3-dioxygenase. This allowed for sufficient metabolism of 4-methyl-catechol to support the growth of host cells on p-methyl-substituted compounds. This discovery of a possible xylT function along with the inconsistencies of my results caused me to cease work with xylT mutants and to concentrate further investigations on the E. coli-like terminator structure found within the xylZ/xylL intergenic region.
Site-Directed mutagenesis of terminator-like structure

A diagrammatic representation of the proposed terminator-like structure, and the position of the nucleotides to be changed as result of mutagenic primers can be seen in Fig. 40. As you can see from the figure, the calculated free-energy ($\Delta G$) of the wild type stem and loop structure is -12.0 kcal/mol. The improved free-energy predictions for duplex stability set forth by Freier et al. (1986) were used for the calculation. The $\Delta G$ of the terminator structure found within the attenuation region of the E. coli trp operon (Platt 1982) was calculated to be -12.2 kcal/mol by the same method. The $\Delta G$ of the terminator structure found downstream of the azurin gene from Pseudomonas aeruginosa (Arvidsson et al. 1989) was calculated to be -16.9 kcal/mol by the same method and could form a structure similar to the one seen in Figures 27 and 32. Several studies have shown that point mutations within the stem region of a hairpin structure that strengthen the free energy of formation also increase termination (Gardner 1982, Rosenberg et al. 1983, Stroynowski et al. 1983). Therefore, point mutations that lower the free energy of formation should decrease termination. The two proposed mutations should lower the wild type $\Delta G$ by approximately one third.

Purified pBK692 (xyILTE) was subjected to the mutagenesis procedure described earlier. Twenty-four colonies from each xyILTEmut6 and xyILTEmut9 transformation
Free-energy predictions ($\Delta G$)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pLTE (wt)</td>
<td>-12.0 kcal/mol</td>
</tr>
<tr>
<td>pLTE mut6</td>
<td>-8.3 kcal/mol</td>
</tr>
<tr>
<td>pLTE mut9</td>
<td>-7.9 kcal/mol</td>
</tr>
</tbody>
</table>
Fig. 40. Putative xylZxylL intergenic region stem and loop structure and the calculated free-energy predictions for wild-type and mutant subclones. Nucleotides in parentheses are non-homologous nucleotides located at identical positions within the TOL plasmid pWWO. The nucleotides changed as result of the mutagenic primers mut6 and mut9 are enclosed in boxes. Free-energy predictions were calculated using the improved parameters set forth by Freier et al. (1986).
were analyzed for mutations. Twenty xylLTEmut6 samples and sixteen xylLTEmut9 samples were sensitive to BglII restriction endonuclease digestion. Clone number 17 for xylLTEmut6 and clone number 49 for xylLTEmut9 were found to contain the desired mutation through DNA sequence analysis (Fig. 41) and were given the names pLTEmut6 and pLTEmut9, respectively.

**Enzyme assays**

Wild-type and mutant strains of *E. coli* DH5α were cultured and cell-free extracts were prepared as described earlier. Assays were performed for DHCDH (XylL) activity as established in Materials and Methods by monitoring the conversion of NAD$^+$ to NADH + H$^+$ at 342 nm. An example of such an assay can be seen in Figure 42. C230 (XylE) activity was measured as described earlier. The combined results of seven different experiments are given in Tables 7 and 8.

Consistency and reproducibility of results continued to pose a serious problem. An area of concern related to the inability to control the lac promoter associated with the pUC expression vector due to the high copy number of the plasmid titrating out the available lac repressor. Azadpour (1991) noted she saw no apparent affect on growth rate or enzyme induction by similar cultures in the absence vs. presence of the gratuitous lac inducer Isopropylthiogalactopyranoside (IPTG). Also, under what conditions did the CAP-cAMP activator complex participate in the expression? Such
Fig. 41. Autoradiograph of a DNA sequencing gel showing the nucleotide sequence change for pLTEmut6 and pLTEmut9. Arrows designate the nucleotides changed by site-directed mutagenesis as compared to the wild-type nucleotide sequence.
Fig. 42. Representative DHCDH assay. 1,2-dihydroxycyclohexa-3,5-diene carboxylate dehydrogenase (DHCDH) activity was monitored spectrophotometrically by NADH⁺ formation (342 nm). Reactions were completed in a 1.00 ml total volume containing 25 mM tris-HCl pH 8.0, ~0.5 mM biologically prepared diol, and 0.5 mM NAD⁺. No more than 10 µl of undiluted or diluted cell-free extract was added to each reaction.
Table 7. Results from seven DHCDH (XylL) assays including slopes (absorbance/sec) and the absorbance of the cultures are shown. Due to the similarities between the slopes and calculated specific activities observed in the earlier assays, protein concentrations and the resulting specific activities were not determined. No data were available for those areas left blank. Conditions refer to the growth medium, how long they were cultured and whether the assays were completed with undiluted or diluted cell-free extracts.

<table>
<thead>
<tr>
<th>XylL Assays</th>
<th>7/15</th>
<th>7/19</th>
<th>7/20</th>
<th>7/25</th>
<th>7/28</th>
<th>8/1</th>
<th>8/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth and Assay Conditions</td>
<td>L2</td>
<td>L2</td>
<td>L2</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>100 fold dilution</td>
</tr>
<tr>
<td>Z Stem</td>
<td>10 fold dilution</td>
<td>24 hrs in minimal growth medium</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mut6</td>
<td>24 hrs in minimal growth medium</td>
<td>10 fold dilution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLTE</td>
<td>47 hrs in minimal growth medium</td>
<td>10 fold dilution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut9</td>
<td>10 fold dilution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wt</td>
<td>10 fold dilution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>0.00067</td>
<td>0.0060</td>
<td>0.00058</td>
<td>0.00033</td>
<td>0.00023</td>
<td>0.0021</td>
<td></td>
</tr>
<tr>
<td>Abs.</td>
<td>0.62</td>
<td>0.59</td>
<td>0.61</td>
<td>0.0029</td>
<td>0.0071</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Stem</td>
<td>Mut9</td>
<td>Slope</td>
<td>0.0095</td>
<td>0.0023</td>
<td>0.00079</td>
<td>0.0018</td>
<td>0.0071</td>
</tr>
<tr>
<td>Abs.</td>
<td>0.59</td>
<td>0.59</td>
<td>0.734</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8. Results from seven C230 (XylE) performed on the same extracts as in Table 7. Due to the similarities between the slopes and calculated specific activities observed in the earlier assays, protein concentrations and the resulting specific activities were not determined. No data were available for those areas left blank. Conditions refer to the growth medium, how long they were cultured and whether the assays were completed with undiluted or diluted cell-free extracts.

<table>
<thead>
<tr>
<th>XylE Assays</th>
<th>Growth and Assay Conditions</th>
<th>7/15</th>
<th>7/19</th>
<th>7/20</th>
<th>7/25</th>
<th>7/28</th>
<th>8/1</th>
<th>8/10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LB</td>
<td>LB</td>
<td>LB</td>
<td>LB</td>
<td>LB</td>
<td>LB</td>
<td>LB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>lag phase</td>
<td>24 hrs</td>
<td>minimal growth</td>
<td>47 hrs</td>
<td>minimal growth</td>
<td>medium</td>
<td>10 fold dilution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>10 fold dilution</td>
<td>100 fold dilution</td>
<td>100 fold dilution</td>
<td>dilution</td>
</tr>
<tr>
<td>wt</td>
<td>slope</td>
<td>0.006</td>
<td>0.003</td>
<td>0.003</td>
<td>0.004</td>
<td>0.002</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>abs.</td>
<td>0.62</td>
<td>0.594</td>
<td>0.585</td>
<td>0.584</td>
<td>0.595</td>
<td>0.595</td>
<td></td>
</tr>
<tr>
<td>pLTE</td>
<td>slope</td>
<td>0.017</td>
<td>0.001</td>
<td>0.028</td>
<td>0.004</td>
<td>0.001</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>abs.</td>
<td>0.59</td>
<td>0.59</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Z Stem pLTE mut6</td>
<td>slope</td>
<td>0.025</td>
<td>0.027</td>
<td>0.033</td>
<td>0.005</td>
<td>0.002</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>abs.</td>
<td>0.59</td>
<td>0.59</td>
<td>0.734</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Stem pLTE mut9</td>
<td>slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>abs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.59</td>
<td></td>
</tr>
</tbody>
</table>
inconsistencies may be eliminated by using a vector that contains a truly or at least a more inducible promoter.

Subcloning of pLTE(wt), pLTEmut6 and pLTEmut9 into cloning/expression vector pSP73

Efforts to find an expression vector with a truly inducible promoter led to the use of the cloning/expression vector pSP73 (Promega). This particular vector contained a T7 RNA polymerase promoter not recognized by the RNA polymerase found in E. coli. When used in conjunction with the E. coli strain JM109(DE3) (Promega), which contains a chromosomal copy of the the gene for T7 RNA polymerase under the control of the lacUV5 promoter, an inducible system is obtained (Studier and Moffatt 1986). Due to the single chromosomal copy of the gene, sufficient lac repressor is available in the cell for repression. The addition of IPTG induces the lacUV5 promoter to produce T7 RNA polymerase, which in turn initiates the high-level expression of the target gene cloned into pSP73. The lacUV5 promoter contains a mutation in the CAP-cAMP region which allows for the full transcription from the lacUV5 promoter in the absence of CAP and cAMP (Silverstone et al. 1970).

pLTE (pBK692), pLTEmut6 and pLTEmut9 are PvuII/EcoRI subclones in pUC19 cloning/expression vector. Transferring the cloned xylLTE TOL region from these constructs to the new pSP73 vector was necessary. Excision of xylLTE DNA from each construct was accomplished by restriction endonuclease
digestion with HindIII and EcoRI. HindIII was chosen to ease subcloning into pSP73, resulting in 22 pUC19 multiple cloning site nucleotides being added to the size of the insert. The mixtures were also subjected to ScaI restriction to help prevent the cloning of the pUC19 vector into pSP73. Next, pSP73 was subjected to restriction with HindIII and EcoRI, followed by an additional digestion with BamHI. BamHI digestion was completed to help prevent the cloning of the small excised pSP73 multiple cloning site DNA piece. A schematic representation can be visualized in Fig. 43. The resulting mixtures were subjected to DNA ligation followed by transformation into E. coli JM109(DE3). Seventy-two colonies were chosen for verification. First, liquid cultures were grown in the presence 0.4 mM IPTG, allowing cloned genes to be expressed. 0.5 ml of each overnight culture was transferred to a tube containing 0.5 ml of 10 mM catechol. If the culture turned bright yellow, that culture was expressing catechol-2,3-dioxygenase (XylE). Second, the plasmid DNA from positive scoring cultures was isolated and subjected to restriction with EcoRI and HindIII followed by agarose gel electrophoresis. The size of pSP73 is approximately 200 bp smaller than pUC19 and this difference in size was readily visible by gel analysis. Finally, DNA sequence analysis on selected samples was completed using two different oligonucleotide primers. One primer was specific to pSP73 located approximately 120 bp downstream of the insertion site, and the other primer identified the mutated
Digestion with
EcoRI
HindIII
BamHI

Ligation

Digestion with
EcoRI
HindIII
ScaI

Multiple Cloning Site
Fig. 43. Diagrammatic representation of the cloning strategy used to construct pSLTE(wt), pSLTEmut6 and pSLTEmut9. The *xylLTE* region from pBK692 was transferred to the new cloning/expression vector pSP73. The resulting subclone was named pSLTE. The same concepts were used to create pSLTEmut6 and pSLTEmut9 from pLTEmut6 and pLTEmut9.
area within the proposed terminator structure. Samples scoring positive results in all areas were designated pSLTE (clone #44), pSLTEmut6 (clone #16) and pSLTEmut9 (clone #34).

**Enzyme Assays**

Wild-type and mutant strains of *E. coli* JM109(DE3) were cultured in 1 liter YT medium to approximately 0.30 absorbance units at 550 nm. Approximately 500 ml of culture was transferred to a centrifuge tube and collected for an initial time point. IPTG was added to the remaining culture to a final concentration of 0.4 mM to enable the induction of T7 RNA polymerase. These cultures were incubated an additional 90 minutes before collection. Cell-free extracts were produced as described earlier. Enzyme assays for DHCDH and C230 were completed as described earlier. Representative results from five experiments are presented in Table 9 and Table 10. Using this expression system, the *xylZ/xylL* terminator-like structure seems to have little or no effect on the transcription of downstream genes.
Table 9. Activity table showing the calculated DHCDH specific activities (nmol/min/mg protein) in extracts from strains carrying the newly constructed plasmids in the cloning/expression vector pSP73. IPTG (0.4 mM final concentration) was added to individual cultures of E. coli JM109(DE3) cells carrying pSLTE, pSLTEmut6 and pSLTEmut9 when each culture reached approximately 0.3 absorbance units at 550 nm. Biologically produced m-toluate diol was used as the substrate.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Time w/ IPTG</th>
<th>Specific activity</th>
<th>Change in specific activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSLTE(wt)</td>
<td>0'</td>
<td>403</td>
<td>1320</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>1723</td>
<td></td>
</tr>
<tr>
<td>pSLTEmut6</td>
<td>0'</td>
<td>21</td>
<td>1130</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>1151</td>
<td></td>
</tr>
<tr>
<td>pSLTEmut9</td>
<td>0'</td>
<td>172</td>
<td>1045</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>1217</td>
<td></td>
</tr>
</tbody>
</table>
Table 10. Activity table showing the calculated C230 specific activities (nmol/min/mg protein) in extracts from strains carrying the newly constructed plasmids in the cloning/expression vector pSP73. Assays were conducted on the same cell-free extracts as those used for Table 9. Commercially produced catechol (Sigma) was used as the substrate.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Time w/ IPTG</th>
<th>Specific activity</th>
<th>Change in specific activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSLTE(wt)</td>
<td>0'</td>
<td>172</td>
<td>1148</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>1320</td>
<td></td>
</tr>
<tr>
<td>pSLTEmut6</td>
<td>0'</td>
<td>10</td>
<td>880</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>890</td>
<td></td>
</tr>
<tr>
<td>pSLTEmut9</td>
<td>0'</td>
<td>74</td>
<td>897</td>
</tr>
<tr>
<td></td>
<td>90'</td>
<td>971</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV

DISCUSSION

In 1991, the United States produced more than 800 million gallons of toluene, the majority derived from petroleum processing (Mannsville 1992). The largest use for toluene is in the production of benzene, but it is also used as an octane booster in gasoline, as a solvent, and as a raw material for the production of toluene diisocyanate. As a solvent, toluene is often found in spray paints, lacquers, inks, glues, spot removers, paint strippers and even perfumes. The 1992 Toxic Chemical Release Inventory reported that an estimated 84 thousand pounds of toluene were released to surface water, 1.6 million pounds were released to underground injection sites, and 708 thousand pounds were released to land by U.S. industries that year. This is just one example of the many man-made toxic organic compounds which contribute to the pollution of our environment.

Concern about environmental pollution has triggered the United States Government to pass the Toxic Substances Control Act of 1976, which governs the manufacture, importation, distribution and processing of all toxic chemicals.

Bioremediation, the process of using biological agents to remove toxic wastes from the environment, has been extensively studied since the 1970’s. Major contributors to
bioremediation are naturally occurring microorganisms that possess the unique ability to completely mineralize selected aromatic and other chemical pollutants. Many of the biodegradative pathways used for the conversion of aromatics to central metabolic pathway intermediates are located on plasmids within these microorganisms. The biodegradative capability of any of these naturally occurring strains is usually restricted to a single class of aromatic compounds. Unfortunately, toxic waste sites and spills are not commonly restricted to one class of compounds and many xenobiotic compounds are not degraded at all. A further understanding of the genetic and biochemical aspects of these pathways is critical for effective genetic engineering of preexisting strains to give them the ability to mineralize an even larger variety of environmental pollutants.

Efforts to better understand the genetic and biochemical aspects of TOL-encoded catabolic pathways prompted our laboratory to subclone these regions from the P. putida plasmid pDK1. The TOL meta-cleavage operon encodes genes responsible for encoding the enzymes which convert benzoates and substituted toluates to Kreb's cycle intermediates. My research was focused on the first three genes of the meta-cleavage pathway (xylXYZ), which encode the enzyme toluate-1,2-dioxygenase, and associated downstream elements with possible regulatory function(s).

Nucleotide sequence analysis
Since most of the xylZ region had been previously sequenced by Khedairy (1990) in this laboratory, the subclone pBK789 (Azadpour 1991) (Fig. 16, Fig. 17), which contained all of xylXY and only a portion of xylZ, was initially chosen for DNA sequence analysis. The nucleotide sequence of pBK789 was completed for both complementary strands, and contiguous segments of sequence were connected by at least 30 bp of overlap. Apparent discrepancies in the sequences between complementary strands and compressions were resolved by the use of additional subclones over the area, additional synthetic oligonucleotide primers, and/or the use of dITP-based sequence reactions (USB protocols 1989). The entire nucleotide sequence for the xylXYZ region was determined in this fashion.

Computer analysis of the elucidated pDK1 DNA sequence was accomplished with the help of Mac™ DNASIS. DNA sequence analysis of the 4185 bp segment revealed the presence of three open reading frames (ORFs), the Pm promoter region and approximately 1600 bp of upstream nucleotide sequence. The three ORFs correspond to xylX (1353 bp), xylY (486 bp) and xylZ (1008 bp, most sequenced by Khedairy (1990), we verified the first 650 bp), encoding predicted protein products of 51370 Da, 19368 Da and 36256 Da, respectively. The average G+C content of the xylXYZ region is 58.9%, significantly lower than the average of the remaining genes of the pDK1 TOL meta-cleavage operon (xylLTEGEFJQKIH) at 64.5%. High nucleotide sequence homology was found between xylXYZ pDK1
and three other known DNA sequences; \textit{xylXYZ} (toluate-1,2-dioxygenase) from the archetype TOL plasmid \textit{pWW0}, \textit{benABC} (benzoate dioxygenase) from \textit{Acinetobacter calcoaceticus}, and \textit{cbdABC} (2-halobenzoate-1,2-dioxygenase) from \textit{Burkholderia cepacia}. All of these enzymes are classified as Class IB hydroxylating dioxygenases, as set forth by Batie et al. (1991).

The highest nucleotide sequence homologies were found between \textit{xylXYZ} \textit{pDK1} and \textit{xylXYZ} \textit{pWW0}. These are 92.7%, 96.1% and 96.6% for \textit{xylX}, \textit{xylY} and \textit{xylZ}, respectively. It is interesting to note that 59.4% of the non-homologous nucleotides occur within the third position of codons. Of the 133 base differences, only 61 resulted in a change in predicted amino acid sequence, and 15 of these were determined to be "equivalent" amino acids (K=R, I=L=V, S=T and D=E). Predicted amino acid homologies for \textit{xylX}, \textit{xylY} and \textit{xylZ} are 93.0%, 93.3% and 94.4%, respectively. As seen in Fig. 26, three proposed nucleotide deletions within \textit{xylX} of \textit{pWW0} were observed. The ramification of these deletions would be a change of five amino acids for \textit{xylX} \textit{pWW0}. These apparent differences may be the result of DNA sequencing artifacts, or a single nucleotide deletion followed by a double deletion of nucleotides downstream in the gene. If the difference is real, we believe that this difference must be the result of a deletion of \textit{pWW0} nucleotides instead of an addition to \textit{pDK1}, because the deletions are not found in \textit{benA} or \textit{cbdA}, which have the \textit{pDK1}-like sequence in this region.
(Fig. 20). Only one out of nine intervening nucleotides (between the two deletions) are different between pWWO and pDK1, suggesting that, if real, it may be a recent mutational event. It is interesting to note that, if the change is real, its consequences are significant. It results in the loss of a negatively charged glutamic acid residue and its replacement with a positively charged arginine residue in the pWWO amino acid sequence. A proline residue, often associated with bends within a protein, is also added to this region of pWWO. We are confident in the xylX nucleotide sequence of pDK1, especially since dITP-based sequence reactions were used to resolve compressions found in the TOL segment encoding this region of the enzyme. The two compressions found here, if incorrectly resolved could have resulted in the discussed changes. These major amino acid differences, along with our nucleotide sequence data (including compressions in the area) and the lack of conservation of the proposed deletions between cbdA and benA, provide strong suggestive evidence that the deletions within xylX of pWWO may actually be an error in the DNA sequence, and not an actual change in the gene.

The chromosomally-encoded benABC genes from Acinetobacter calcoaceticus and the plasmid-encoded cbdABC genes from Burkholderia cepacia also exhibit high nucleotide and deduced amino acid sequence homology to xylXYZ. Nucleotide sequence homologies for xylX, xylY and xylZ to benA, benB and benC are 56.9%, 57.5% and 44.5%, respectively.
Predicted protein homologies for the same genes are 63.2%, 60.1% and 54.0%, respectively. When “equivalent” amino acids are included, the homologies increase to 68.5%, 66.7% and 62.2%, respectively. Homologies for xylX, xylY and xylZ to cbdA, cbdB and cbdC are 55.2%, 61.6% and 43.1%, respectively. Predicted protein homologies for the same genes are 55.5%, 54.6% and 47.2%, respectively. These homologies increase to 65.2%, 64.8% and 57.2%, respectively, when “equivalent” amino acids are considered.

Immediately upstream of xylX lies the Pm promoter region for the meta-cleavage operon. Comparisons of the pDK1 nucleotide sequence to that of pWWO revealed a 90% sequence homology. The pDK1 sequence contained the same TACAN4TGCA motif upstream of the -35 region found to be the minimum DNA segment required for stimulation of transcription by XylS in pWWO (Gallegos et al. 1996b) (Fig. 28). Further analysis of the upstream nucleotide sequence determined that 1680 bp of TOL DNA existed upstream of the xylX start codon in construct pBK789 (and pBK190). All six possible reading frames (normal and complementary) were analyzed for potential ORFs in this upstream region. Only one potential ORF was identified greater than 200 bp in length (234 bp). The entire upstream nucleotide sequence was compared to other known DNA sequences in Genbank without finding significant homologies to any known DNA sequences (only one, to a 131 bp unknown P. putida sequence). This 1600 bp segment of seemingly non-functioning DNA shows that the interoperonic regions are not
"streamlined" with regard to base pair usage as compared to the intergenic regions its within the meta-cleavage and upper operons.

**Deletion mutagenesis of upstream nucleotide sequence and effect on TO expression**

Azadpour (1991) found that the conversion of 5 mM benzoate, m-toluate, p-toluate or 3,4-dimethyl benzoate to their corresponding dihydrodiols required a culture of *E. coli* carrying pBKl90 (under control of the *lac* promoter of pUC19) an average of 36 hours. These dihydrodiols are necessary for studies of subsequent meta-cleavage enzymes, particularly XylL (DHCDH). Constructs capable of performing this conversion at a more rapid rate were made by deleting portions of the 1680 bp upstream of the *xylX* start codon, bringing the *lac* promoter of pUC19 closer to the start of *xylX*. pBK192, pBK292, and pBK392 (Fig. 30 and Fig. 19) reduce the distance from 1680 bp to 485 bp, 59 bp and 110 bp, respectively. The TACAN₉TGCA motif described above as being the minimum DNA segment required for stimulation of transcription by XylS was not present in pBK292. *E. coli* strains carrying constructs pBK190, pBK192, pBK292 and pBK392 were grown in the presence of m-toluate and rates of toluate-1,2-dioxygenase expression were compared by the monitoring of dihydrodiol accumulation spectrophotometrically. Results demonstrated there was a 4- to 5-fold increase in the rate of dihydrodiol accumulation by strains carrying pBK192, pBK292,
and pBK392 over that of the original construct, pBK190 (Fig. 31). There is no apparent effect on transcription by the Pseudomonas TOL Pm promoter of pDK1 in E. coli, as deletion of the TACAN4TGCA motif had no substantial effect on the level of dihydrodiol accumulation.

Site-directed mutagenesis of possible downstream regulatory elements

Our laboratory has completed the entire nucleotide sequence of all 13 genes of the pDK1 TOL meta-cleavage operon. Nucleotide sequence analysis revealed relatively short intergenic spaces between genes, with the two prominent exceptions being over 100 bp in length (Table 3). Analysis of the intergenic regions (IRs) between xylZ/xylL and between xylL/xylT showed each to have the potential for stem and loop structure formation at the RNA level (Fig. 32). The predicted xylZ/xylL IR secondary structure of the transcribed RNA has characteristics similar to an E. coli rho-independent transcriptional terminator (Baker 1992, and this research). Using a series of subclones in E. coli plasmids pUC18/19, which included and excluded various areas throughout the xylLTE region, Baker (1992) obtained evidence suggesting a transcriptional terminator may be located immediately upstream of xylL and a positive acting element appeared to be present in the area downstream of xylL (Fig. 33). He also speculated that the xylT gene product may have some direct influence on gene expression within this region, possibly
acting as a type of internal regulator, acting at the level of transcription or translation. The *xylL/xylT* IR was also suggested to be involved. Combining these findings with a previously reported observation of weak promoter activity upstream of the *xylE* gene (Keil et al. 1987), raised the possibility that one or more functional regulatory elements may be present in this region.

The initial area addressed was the *xylT* region, including the *xylT* coding sequence and the intergenic region between *xylL* and *xylT*. When this research was initiated, the function of *xylT* was not known. Harayama et al. (1991b) had initially suggested (from nucleotide sequence and deduced amino acid sequence only) that the function of XylT might be similar to chloroplast-type ferredoxins, because of a characteristic Cys-XXX-Cys-XX-Cys motif found in each case. Baker's (1992) preliminary results suggested another possibility. I therefore used site-directed mutagenesis to create mutant *xylT* genes, one with a nonsense mutation in the first part of *xylT* (Fig. 34) and a second with a missense mutation in the proposed metal binding motif in *xylT* (Fig. 34) (Harayama et al. 1991b). I also produced a clone into which I had introduced mutations to destabilize the stem of the proposed stem and loop structure between *xylL* and *xylT* (Fig 38). *E. coli* strains carrying the various wild-type and mutant constructs were cultured, C230 assays (*xylE* is downstream of all these sites in the mutant clones) were performed and specific activities were compared (Table 6).
Several attempts were made to obtain results similar to those observed by Baker (1992) without success. Due to the high level of variability observed, it was not possible to draw any firm conclusions about what effect if any that the mutations introduced into \textit{xylT} or the palindromic region immediately upstream of \textit{xylT} had upon the expression of C230 activity. At this time, Polissi and Harayama (1993) had new evidence to suggest that XylT served to regenerate inactivated C230 when cells were grown on \textit{p}-methyl substituted compounds. An effective analysis of a putative regulatory region really should be carried out in a proper \textit{P. putida} TOL background. This would be especially true if a regulatory protein not included with the construct is needed for proper function. If such a \textit{Pseudomonas} regulatory protein was needed for the function of the \textit{xylL/xylT} IR palindrome (and it was not \textit{xylT}), these constructs would not be expected to respond properly in \textit{E. coli}. Thus the discovery of the true \textit{xylT} function, along with the experimental inconsistencies, prompted changing the focus of my research to the second intergenic region of interest (found between \textit{xylZ/xylL}) encoding an \textit{E. coli} transcriptional terminator-like structure. It is therefore still not clear whether the \textit{xylL/xylT} intergenic region has a regulatory function in the expression of the TOL meta-cleavage operon, but these constructs may be useful later if they are introduced back into the TOL plasmid in \textit{P. putida}, and the possible significance of this region is discussed in more
detail later.

The transcriptional terminator-like structure found between xylZ and xylL attracted our attention for several reasons. Why would a terminator-like structure be present after the first three genes in the meta-cleavage operon separating them from the remaining ten genes? Is this terminator-like structure even functional in the TOL meta-cleavage operon or is it a non-functioning relic? Several studies have shown that point mutations within the stem region of a transcriptional terminator structure that strengthen the free energy of formation (ΔG) also increase termination (Gardner 1982, Rosenberg et al. 1983, Stroymowski et al. 1983). Therefore, point mutations that lower the ΔG should decrease termination. The calculated free-energy (Freir et al. 1986) of the proposed stem and loop secondary structure from the xylZ/xylL IR is -12.0 kcal/mol (Fig. 40). This calculated ΔG was similar to that for other transcriptional terminator structures previously described in E. coli and Pseudomonas aeruginosa (Platt 1981, Arvidsson et al. 1989). Two mutant constructs of pBK692 (xylLTX) were created through site-directed mutagenesis. The calculated ΔGs of the mutant stem structures were decreased by approximately one third (Fig. 40). Assays for DHCDH and C230 disclosed that problems still existed in the areas of consistency and reproducibility of results (Tables 7 and 8). Previous studies of similar cultures and conditions (Azadpour 1991) showed that there was no apparent effect on culture
growth rate or enzyme induction by the presence or absence of the lac promoter gratuitous inducer IPTG. The high copy number of the pUC plasmids quickly titrates out all of the available lac repressor (produced only from the chromosomal lacI gene). The CAP-cAMP activator complex can also be expected to participate in regulating gene expression from the pUC promoter and this may again be affected by plasmid copy number as well as culture conditions. Thus, culture media and growth phase become important conditions to control. Efforts to effectively control these conditions were attempted without reproducible success. Concern for the lack of inducibility and control over the lac promoter associated with the pUC18/19 expression vectors finally led to the search for an expression vector with a more reliably inducible/controlled promoter.

The Promega pSP72/73 E. coli expression vector utilizes a T7 RNA polymerase promoter which is not recognized by the RNA polymerase found in E. coli. Using this vector in conjunction with E. coli strain JM109(DE3) (Promega, contains a chromosomal copy of T7 RNA polymerase under control of the lacUV5 promoter), produces an inducible system regulated by IPTG. The xylLTE regions from pLTE(wt), pLTEmut6 and pLTEmut9 were transferred to the new vector pSP73 (Fig 42). As seen from the data in Tables 9 and 10, the specific activities for DHCDH and C230 in extracts from cells carrying wild-type and mutant constructs were very similar and reproducible. In fact, the mutant constructs, where less
termination would have been postulated to occur, actually had somewhat lower DHCDH and C230 activities than the wild-type construct. Using this expression system, the data do not provide any evidence that the terminator-like structure encoded by the region found between xylZ and xylL in fact functions as a transcriptional terminator. It therefore appears that T7 RNA polymerase simply does not recognize this structure as a viable terminator. It has been observed that T7 RNA polymerase terminates transcription when a stable secondary structure followed by a run of uracil residues is present (Jeng et al. 1990, Macdonald et al. 1993), but when the run of uracil residues is missing or is shortened, termination fails or its efficiency is substantially diminished (Jeng et al. 1990, Macdonald et al. 1993). It was further observed that T7 RNA polymerase requires a minimum of five uracils for termination to occur in selected structures (Jeng et al. 1990, Macdonald et al. 1993). The calculated stability of the TOL structure is well within the necessary range, but as seen Figure 27, the adjacent run of uracil residues, although 7 total, is interrupted twice by non-uracil residues. Some data is available for termination efficiencies by RNA polymerases at structures where the uracil residues are not continuous, but none precisely describe the scenario observed here (Macdonald et al. 1993). However, based upon the known negative effect of a single interruption one can certainly expect that the two interruptions could be sufficient to cause the T7 RNA
polymerase not to terminate transcription at the xylZ/xylL IR hairpin site. One cannot predict with certainty that P. putida RNA polymerase will act similarly to T7 RNA polymerase under these same conditions. However, as I discuss below, if this structure was simply the terminator left from the first of three fused operons, it would need to be inactivated to allow the expression of the downstream genes and the interruption of the continuous uracil residues might be sufficient to allow for read through by P. putida RNA polymerase (discussed in more detail later).

**Evolutionary possibilities for the "modern" TOL operon**

Several questions remain unanswered about the evolution of this TOL system and related aromatic catabolic operons. It has been suggested that the meta-cleavage pathway operon may have evolved from the merging of DNA modules. Harayama and Rekik (1990) suggested that the meta-cleavage operon and other related operons might be the products of fusions between two such DNA modules, xylXYZL and xylTEGFJQKIH in the case of TOL, where each unit would have originally possessed its own operator promoter region (Fig. 8). Based upon my findings and information from others, it is possible to speculate that the "modern" TOL meta-cleavage operon may actually be the result of the fusion of three DNA modules instead of two. These three modules could have been xylXYZ, xylL and xylTEGFJQKIH. Evidence to support this model of TOL
evolution is discussed below.

First let us identify the relevant related operons which include dioxygenases akin to toluate-1,2-dioxygenase (TO). Rosche et al. (1997) established the phylogenetic relationships of multicomponent oxygenases. As seen in Figure 44, components of 2-halobenzoate-1,2-dioxygenase, and an ancestor to TO and benzoate dioxygenase, evolved from a common ancestor before toluate-1,2-dioxygenase and benzoate dioxygenase diverged. 2-halobenzoate-1,2-dioxygenase (encoded by *cbdABC* on plasmid pBAH1) is responsible for the conversion of 2-halobenzoate to a common intermediate, catechol, in *Burkholderia cepacia*. Benzoate dioxygenase (encoded by *benABC*) is the first enzyme of the chromosomally-encoded benzoate degradation pathway in *Acinetobacter calcoaceticus*. Once again, all these enzymes are Class IB dioxygenases as defined by Batie et al. (1991).

A comparison of the operons encoding these related genes suggests that the transcriptional terminator-like element found in the intergenic region between *xylZ* and *xylL* may be or may once have been relevant to operon function. This stem and loop structure is conserved between several operons. A stem structure identical to that found in pDK1 (Fig. 27) is also seen in pWWO. The only difference in the nucleotide sequence between the two is one nucleotide located in the loop region. The *cbdABC* operon, encoding 2-halobenzoate-1,2-dioxygenase, also contains what appears to have once been a terminator-like structure just downstream of *cbdC*. It is
Fig. 44. Phylogenetic analysis of oxygenase components (α subunits) (A) and reductase components (B). The program TREE of the Genetic Computer Group software package was used to calculate the relationships (Rosche et al. 1997). Lengths of the branches along with the numbers indicate the relative phylogenetic distances among the amino acid sequences (Rosche et al. 1997). The proposed class is the same one developed by Batie et al. (1991). This figure was adapted from Rosche et al. (1997).
interesting to note that the next potential ORF after cbdC is
327 nucleotides away (Haak et al. 1995), leaving a large gap
between genes as seen in both TOL meta operons. The
phylogenetic evidence (Rosche et al. 1997) suggests that 2-
halobenzoate-1,2-dioxygenase and TO came from a common
ancestor. Since the terminator-like structure is conserved
between these operons, it is possible for one module in the
evolution of the meta-cleavage operon to be xylXYZ, where the
terminator-like structure was included as part of this
module.

Is the terminator-like structure found between xylZ and
xylL actually functional in some way? The final answer to
this question may only lie in Pseudomonas. As described
earlier, this structure is conserved between xylXYZ (pDK1 and
pWWO) and cbdABC, at least suggesting it may be relevant or
functional at some level. However, in addition to our data
using T7 RNA polymerase, a further argument does exist that
it may be non-essential as demonstrated by comparing the TOL
meta-cleavage operon to the benzoate operon. The benzoate
operon contains five genes (benABCDE) encoding proteins for
the conversion of benzoate to catechol. The function of BenE
is currently unknown. BenD is homologous in function to XylL
(Neidle et al. 1992). The interesting property of this
operon is that a terminator-like structure does not exist
between benC and benD as it does between the homologous xylZ
and xylL genes. In fact, benC and benD are only separated by
17 nucleotides (Neidle et al. 1991). The phylogenetic
analysis (Fig. 44) shows that benzoate dioxygenase, 2-halobenzoate-1,2-dioxygenase and TO all came from a common ancestor before TO and benzoate dioxygenase diverged. If this terminator-like structure is necessary, why is it not conserved between all three operons? It is certainly possible that this terminator-like structure was also present in the benzoate operon, but over time was lost due to its non-essential function. The benzoate operon could thus represent a more advanced form of the expression unit as shown by the reduction in space between these genes.

Further evidence to support the proposal that xy1XYZ may have once been an independent module is seen in the G+C content of these genes. The difference in G+C content most likely indicates that the xy1XYZ genes have evolved in a different genetic background from the rest of the operon. As seen in Table 4, analysis of the G+C content of the TOL pDK1 meta-cleavage operon shows that the genes encoding xy1XYZ are substantially lower (58.9%) in G+C content than xy1L (64.5%) and the remaining genes in the meta-cleavage operon (64.5%). As described in the introduction, genes homologous to xy1TEGFJQKIH are conserved in other operons (napthalene and phenol degradation), suggesting that this segment is older and was once a module of its own. The intergenic region between xy1L and xy1T may therefore be remnants of an old promoter region for the genes xy1TEGFJQKIH. This possibility is supported by the relatively high (for P. putida) 47% A+T composition of the region.
I propose that the first fusion may have occurred between \textit{xylL} ("out" source not known) and \textit{xylTEGFJQKIH}, followed by a second fusion of \textit{xylXYZ} to \textit{xylLTEGFJQKIH}. The first fusion could have occurred much earlier in the evolution to explain the lack of G+C content differences between \textit{xylL} and \textit{xylTEGFJQKIH}, or the genes may simply have evolved in similar genetic backgrounds. Either at time of fusion or since that time, the \textit{xylL} terminator was eliminated. The second fusion probably occurred more recently, involving genes that evolved in different genetic backgrounds as evidenced by the large differences in the G+C content. As a result of this fusion, or since that time, the \textit{xylL} promoter has also been eliminated. It would be important to the usefulness of a newly fused operon to eliminate any unnecessary control elements that would prevent the proper function the operon. The terminator-like structure between \textit{xylZ} and \textit{xylL} may have been fully functional in the \textit{xylXYZ} operon. If this structure was completely functional in the fusion product, the remaining downstream genes would never be expressed. Even if the putative \textit{xylL/xylT} IR region promoter were functional, \textit{xylL} would still not be expressed. The interruption of the uracil residues may be the result of the need for the terminator-like structure to no longer be functional and it could possibly reflect the method used to gain downstream expression. As part of the evolutionary process, pDK1 and pWW0 TOL might be in the early stages of "clean up", removing
and/or changing the function of these intergenic regions. These proposed fusions could help to explain how the two large intergenic regions between \textit{xylZ/xylL} and \textit{xylL/xylT} (each greater than 100 bp in length) possibly came to be part of the "modern" TOL meta-cleavage operon.

There also appears to be a substantial amount of excess interoperonic "baggage" associated with the TOL plasmids. As demonstrated by this research, there is approximately 1600 bp of apparently non-coding DNA sequence upstream of the \textit{Pm} promoter. Figure 1 demonstrates how many of the different TOL plasmids contain duplications of entire operons and some non-functioning pseudogenes. These plasmids definitely are not as genetically streamlined as the bacterial chromosome. At this point in the TOL evolution, there must not be as much selective pressure for encoding space or this pressure has not yet resulted in the deletion of non-functioning segments. Additionally, the TOL upper operon (\textit{xylUWCMABN}) seems to be a recent addition to this region itself. Guigneaux (1997) showed that the G+C content of the pDX1 TOL upper operon (50\%) was much lower than than the 63.4 \% G+C content of the entire meta-cleavage operon.

These few examples show that the regions of TOL plasmids encoding these catabolic pathways may not have many of the genetic characteristics of their established chromosomal counterparts. This property of rapid change may in fact be the nature of such catabolic plasmids and as more xenobiotic compounds are introduced into our environment, more
reorganization for new pathways may be expected to result.

The mutant constructs created in this study can be of further use in answering some of the remaining questions about this region. One of the lingering questions not addressed adequately with this research is what function, if any, does the intergenic region between \textit{xylL} and \textit{xylT} and its encoded palindrome have? Studies are currently being conducted to develop a method to incorporate the mutated segments back into the pDK1 TOL plasmid using the recombination abilities of \textit{Acinetobacter calcoaceticus}. Once accomplished, the TOL plasmid can be conjugated back into \textit{Pseudomonas putida} for further phenotypic studies. In this way we will determine if our findings using \textit{E. coli} systems are valid indicators of function in \textit{P. putida}. 
REFERENCES


Birnboim HC, Doly J (1977) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Research 7:1513-1515


de Lorenzo V, Herrero M, Timmis KN (1991) An upstream XylR and IHF induced nucleoprotein complex regulates the s^54-dependent Pu promoter of TOL plasmid. EMBO J 10:159-1167


Gallegos MT, Marques S, Ramos JL (1996a) Expression of the
TOL plasmid xylS gene in Pseudomonas putida occurs from a $\sigma^{70}$-dependent promoter or from $\sigma^{70}$- and $\sigma^{54}$-dependent tandem promoters according to the compound used for growth. J Bacteriol 178:2356-2361

Gallegos MT, Marques S, Ramos JL (1996b) The TACAN$_4$TGCA motif upstream from the -35 region in the $\sigma^{70}$- and $\sigma^{54}$-dependent Pm promoter of the TOL plasmid is the minimum DNA segment required for transcription stimulation by XylS regulators. J Bacteriol 178:6427-6434


Harayama S, Rekik M (1990) The meta-cleavage operon of the
TOL degradative plasmid pWW0 comprises 13 genes. Mol Gen Genet 221:113-120


Inouye S, Nakazawa A, Nakazawa T (1987) Overproduction of the


Khedairy HS (1990) Determination of the Complete Nucleotide Sequence of the xylZ Region of the Pseudomonas putida TOL Plasmid pDK1, Encoding a Subunit of the Toluate Oxidase Complex. Master of Science. University of North Texas


Kunz DA, Chapman PS (1981b) Isolation and characterization of
spontaneously occurring TOL plasmid mutants of 
*Pseudomonas putida* HSl. J Bacteriol 146:952-964

Lee JY, Jung KH, Choi SH and Kim HS (1995) Combination of the 
tod and tol pathways in redesigning a metabolic route of 
*Pseudomonas putida* for the mineralization of a benzene, 
toluene and p-xylene mixture. Appl Environ Microbiol 61:2211-2217

Lehrbach PR, Zeyer J, Reineke W, Knackmuss HJ, Timmis KN 
to extend the range of haloaromatics degraded by 

Lowry OH, Rosenbrough NJ, Farr AL, Randall RJ (1951) Protein 
measurement with the Folin phenol reagent. J Biol Chem 
193:265-275

Macdonald L.E., Zhou Y, McAllister WT (1993) Termination and 
slippage by bacteriophage T7 RNA polymerase. J Mol Biol 
232:1030-1047

Mannsville Chemical Products Corporation

induction kinetics from the promoters of the catabolic 
pathways of the TOL plasmid pWW0 of *Pseudomonas putida* 

transcription from the positively controlled Pm promoter 
of the TOL plasmid of *Pseudomonas putida*. Mol Microbiol 
18:851-857

Mason JR, Cammack R (1992) The electron-transport proteins of 

Transcription of the TOL plasmid toluate catabolic 
pathway operon of *Pseudomonas putida* is determined by a 
pair of co-ordinately and positively regulated 
overlapping promoters. EMBO J 3:2461-2466

Nakazawa T, Yokota T (1973) Benzoate metabolism in 
*Pseudomonas putida* (arvilla) mt-2: demonstration of two 


Scheirer PH, Cortese R (1979) A fast and simple method for
sequencing DNA cloned into single-stranded bacteriophage
as an aid to rapid DNA sequencing. J Mol Biol 143:161

Shingler V, Powlowski J, Markland M (1992) Nucleotide
Sequence and Functional Analysis of the Complete
Phenol/3,4-Dimethylphenol Catabolic Pathway of

Silverstone AE, Arditti RR, Magasanik B (1970) Catabolite-
insensitive revertants of lac promoter mutants. Proc
Natl Acad Sci USA 66(3):773-779

Stroynowski I, Kuroda M, Yanofsky C (1983) Transcription
termination in vitro at the tryptophan operon attenuator
is controlled by secondary structures in the leader
transcript. Proc Natl Acad Sci 80:2206-2210

Studier FW, Moffatt, BA (1986) Use of bacteriophage T7 RNA
polymerase to direct selective high-level expression of
cloned genes. J Mol Biol 189:113-130

Primary structure of xylene monooxygenase: similarities
to and differences from the alkane hydroxylation system.
J Bacteriol 173(5):1690-1695

factor composite plasmid in vitro: means for
amplification of deoxyribonucleic acid. J Bacteriol
121:354-362

mutagenesis with Escherichia coli DNA polymerase III
holoenzyme. Gene 62:135-139

protocols for DNA sequencing with Sequenase™ Version
2.0 pp 5-13

Williams PA, Worsey MJ (1976) Ubiquity of plasmids in coding
for toluene and xylene metabolism in soil bacteria:
Evidence for the existence of new TOL plasmids. J
Bacteriol 125:818-828

Williams PA, Shaw LM, Pitt CW, Vrec1 M (1997) XylUW, two
genes at the start of the upper pathway operon of the
TOL plasmid pWW0, appear to play no essential part in
determining its catabolic phenotype. Microbiology
143(Pt1):101-107