Systems Analysis and Metabolic Engineering of Biocatalytic Reaction Networks: Application to Indene Bioconversion

by

Daniel Edward Stafford

B.S.E. summa cum laude in Chemical Engineering, University of Michigan
May 1997

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2002

© Massachusetts Institute of Technology, 2002. All Rights Reserved.

Signature of Author __________________________

Department of Chemical Engineering
December 3, 2001

Certified by ____________

Gregory Stephanopoulos
Professor of Chemical Engineering
Thesis Supervisor

Accepted by __________________________

Daniel Blankschtein
Professor of Chemical Engineering
Chairman, Committee on Graduate Students
SYSTEMS ANALYSIS AND METABOLIC ENGINEERING OF BIOCATALYTIC REACTION NETWORKS: APPLICATION TO INDENE BIOCONVERSION

by
Daniel Edward Stafford

Submitted to the Department of Chemical Engineering on December 3, 2001 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Chemical Engineering

ABSTRACT

Metabolic engineering is a science concerned with the improvement of cellular properties through the application of recombinant DNA technology. To this point, the field has frequently been limited to methodological issues pertaining to gene introduction, expression and control by way of molecular biological techniques. While such methods are critical for the implementation of strategies for metabolic pathway design and modification, an equally critical activity is the quantitative analysis of metabolic networks and the evaluation of phenotype/genotype relationships. A prominent new opportunity to extend the scope of metabolic engineering has recently emerged to analyze and engineer biocatalytic microorganisms capable of catalyzing stereospecific transformations useful in chiral pharmaceutical manufacturing.

This thesis formulates and applies a general framework for the optimization of uncharacterized bioconversion strains (those biocatalysts with little a priori genetic or metabolic data available) comprising five essential steps: (i) Establishment of an experimental system for strain selection and metabolic network analysis, (ii) definition of the bioconversion network and quantification of network fluxes, (iii) target identification, (iv) flux redistribution, and (v) analysis of the modified bioconversion strain to identify areas for further development. In the case described here, a systematic evaluation of the physiology of the soil bacterium Rhodococcus, and analysis of the relative fluxes of an indene bioconversion network, facilitated the construction of a biocatalyst for production of (2R)-indandiol suitable for the manufacturing of the HIV protease inhibitor, indinavir sulfate (Crixivan®, Merck and Co., Inc.).

The chemical synthesis of (−)-cis-(1S,2R)-1-aminoindan-2-ol [−)-CAI], a key precursor in Crixivan® manufacturing, may be carried out by the asymmetric epoxidation of indene to (1S,2R)-indan oxide at up to 87% enantiomeric excess by way of hydrolytic kinetic resolution followed by crystallization to purify the (1S,2R)-indan oxide. To circumvent this technically demanding epoxidation, Merck scientists conceptualized a bioconversion process in which indene is oxidized to one of three derivatives that can serve as precursors to (−)-CAI: cis-(1S,2R)-indandiol, trans-(1R,2R)-indandiol, or (1S,2R)-indan oxide. Rhodococcus sp. were isolated that utilize complex networks of enzymatic reactions to convert indene to several oxygenated derivatives of different stereochemistries.

In previous work, prolonged cultivation of Rhodococcus sp. I24 in a continuous flow system with a novel indene air delivery led to the evolution of a mutant strain, designated KY1, with improved bioconversion properties, in particular a twofold increase in yield of (2R)-indandiol relative to I24. Induction studies with both strains indicated that KY1 lacked
a toluene-inducible dioxygenase activity present in I24 that was responsible for the formation of undesired byproducts. Flux analysis of indene bioconversion in KY1 performed using steady state metabolite balancing and labeling with [14C]-tracers revealed that at least 94% of the indene is oxidized by a monooxygenase to (1S,2R)-inden oxide that is non-enzymatically hydrolyzed to trans-(1R,2R)-indandioli and cis-(1S,2R)-indandioli.

In this study, indene bioconversion network targets were prioritized based upon the results of flux analysis and the availability of genetic tools for Rhodococcus. The primary target that was identified was the (1S,2R)-inden oxide node of Rhodococcus sp. KY1 and the need to selectively hydrolyze the epoxide to trans-(1R,2R)-indandioli to prevent enzymatic product degradation. Toward this end, the limA gene encoding limonene-1,2-epoxide hydrolase from Rhodococcus erythropolis DCL14 was constructed by template-free PCR, and constitutively expressed in Rhodococcus sp. KY1. The recombinant KY1(pDS3) strain showed enantioselective hydrolysis of the intermediate (1S,2R)-inden oxide to trans-(1R,2R)-indandioli at early culture times with subsequent attenuation of epoxide hydrolase activity. Alternatively, the direction of the first order hydrolysis of inden oxide to trans- and cis-indandioli can be manipulated through two pH-dependant mechanisms, with primarily trans-(1R,2R)-indandioli being formed at pH > 8. Indene bioconversions with KY1(pDS3) at pH 8.6 resulted in >95% selectivity of trans-(1R,2R)-indandioli compared to approximately 60% in KY1 at pH 7.0. These approaches to directing inden oxide hydrolysis toward trans-(1R,2R)-indandioli were validated in biphasic indene bioconversions using silicon oil to solubilize the aqueously immiscible indene substrate.

Subsequent analysis of the regulatory structure of the KY1 indene bioconversion network using radiolabeled tracers revealed that the enzymatic activity catalyzing the oxidation of indene to (1S,2R)-inden oxide is inhibited in vivo by concentrations of trans-(1R,2R)-indandioli as low as 1.0 g/L. Further experiments exploring the loss of limonene-1,2-epoxide hydrolase activity in the recombinant KY1(pDS3) strain determined that epoxide hydrolase activity in KY1(pDS3) is not affected by accumulation of inden oxide or trans-(1R,2R)-indandioli. These results suggested several courses of action for industrial implementation of a bioconversion process for (2R)-indandioli production, including directed evolution studies to remove inhibitory effects on the indene monooxygenase and/or improve epoxide hydrolase turnover, product removal using hydrophobic resins, or the use of high-pH bioconversion conditions.

This project has successfully applied the concepts and tools of metabolic engineering to the directed production of trans-(1R,2R)-indandioli. The methods described here for improving (2R)-indandioli biosynthesis may be applicable to other bioconversion systems that oxidize its substrate through an epoxide intermediate. The enantiomeric purity obtained in other systems using these methods is dependant both on the stereospecificity of the relevant monooxygenase and the stereochemistry of the epoxide hydrolysis reaction(s). This work demonstrates that the systematic application of the analytical and genetic methods of metabolic engineering can improve significantly the yield and selectivity profiles of industrial production strains.

Thesis Supervisor: Professor Gregory Stephanopoulos
ACKNOWLEDGMENTS

Numerous people deserve many thanks for helping to make this research possible. First, I must recognize the valuable guidance of my thesis advisor, Prof. Gregory Stephanopoulos. He exhibited the patience to allow me to learn independently to approach research problems intelligently, and provided many opportunities for me to add breadth beyond laboratory research to my graduate education, all of which I appreciate a great deal. As my predecessors in this laboratory have often recognized, he has tremendous motivational abilities that allow us, his students, to enthusiastically pursue our goals. I also thank Prof. Anthony Sinskey for providing me the time to learn the "language" of molecular biology with his research group and for providing important research guidance. I would also like to recognize the helpful contributions of the additional members of my thesis committee: Prof. Boris Magasanik, Dr. Stephen Drew, and Prof. George Stephanopoulos.

I was privileged in my time at MIT to draw upon the collective expertise of both the Stephanopoulos and Sinskey laboratories, in addition to many scientists at Merck Research Laboratories in Rahway, New Jersey. Perhaps the most critical help I received was from Phil Lessard, who patiently assisted me with the molecular biology aspects of the "$\textit{limA}$ work." In addition, several other members of the Sinskey laboratory were very helpful, including Greg York, Jefferson Parker and Laura Willis. In the Stephanopoulos lab, I would like to thank two "elders" who helped me get started, Sushil Rijhwani and Mattheos Koffas. Their encouragement during the difficult early days of research was very valuable. A special thanks goes to my direct predecessor, Kurt Yanagimachi, who set a great example for performing efficient research that was quite a challenge to follow. Also, from Merck I would like to thank Kodzo Gbewonyo, Bruce Burgess, Joe King, and Leonard Lister for making their fermentation facility available for biphasic indene bioconversions, Pat Gailliot and Christine Ingram for their help in purifying $\textit{trans}$-indandiol, Sheri Nidosito for making her laboratory available to me during my time at Merck, and Barry Buckland for facilitating all of the work done in Rahway.

I want to thank my parents, Joe and Joan Stafford, and my brothers, Andy and Dave, for their unconditional support and encouragement, and quiet confidence in my abilities through the years that has helped make this and much else attainable. Finally, I thank my wonderful wife, Jennifer, for her love and unwavering patience over the past 4+ years. Her impeccable laboratory skills (at least for a mechanical engineer!), but more importantly her companionship were invaluable during my time at MIT. I look forward to a long, happy lifetime together with her.

Financial support from the NIH Biotechnology Training Grant and Merck Research Laboratories are gratefully acknowledged.
# TABLE OF CONTENTS

ABSTRACT 2

ACKNOWLEDGMENTS 4

TABLE OF CONTENTS 5

LIST OF FIGURES 9

LIST OF TABLES 15

CHAPTER 1 - INTRODUCTION 16

1.1 Metabolic engineering: a new frontier of chemical reaction engineering 16

1.2 Extension of metabolic engineering principles to biocatalytic reaction networks 19
   1.2.1 Biocatalytic synthesis of chiral compounds 20
   1.2.2 Protease inhibitors for treatment of HIV infection 21
   1.2.3 1,2-Indandiol: a key precursor in indinavir sulfate synthesis 21

1.3 Thesis objectives 23

1.4 Thesis organization 26

CHAPTER 2 - METABOLIC ENGINEERING: A PARADIGM FOR INDENE BIOCONVERSION ANALYSIS 28

2.1 Metabolic engineering as an integrating platform for strain development 29
   2.1.1 Introduction 29
   2.1.2 Drivers of metabolic engineering 31
       2.1.2.1 Genomics 31
       2.1.2.2 Applied molecular biology 32
       2.1.2.3 Renewable resource utilization 33
       2.1.2.4 Emerging opportunities 34
   2.1.3 Tools of metabolic engineering 34
   2.1.4 Future prospects for metabolic engineering 37
2.2 TECHNOLOGIES FOR BIOCATALYTIC STEREOSELECTIVE OXIDATION:
ENZYMES AND MICROORGANISMS
2.2.1 Structure of catabolic pathways for aromatic compounds 38
2.2.2 Gene organization in aromatic catabolic pathways 39
2.2.3 Key biocatalytic enzymes 40
   2.2.3.1 Oxygenases 41
   2.2.3.2 Epoxide hydrolases 42
      2.2.3.2.1 Fungal epoxide hydrolases 43
      2.2.3.2.2 Bacterial epoxide hydrolases 44
   2.2.3.2.2 General 45
   2.2.3.2.2 Limonene-1,2-epoxide hydrolase 46
2.2.4 Key microorganisms 47
   2.2.4.1 Pseudomonas 48
   2.2.4.2 Rhodococcus 49

2.3 METABOLIC FLUX ANALYSIS OF INDENE OXIDATION IN RHODOCOCCUS
2.3.1 Indene bioconversion: A promising system for improvement by systems analysis and metabolic engineering 50
2.3.2 Indene bioconversion in Rhodococcus sp. I24 51
   2.3.2.1 Isolation of Rhodococcus sp. I24 and characterization of indene bioconversion 52
   2.3.2.2 Substrate and product toxicity 53
2.3.3 Isolation of Rhodococcus sp. KY1 from chemostat cultures 54
2.3.4 Characterization of Rhodococcus sp. KY1 using radiolabeled tracers 55
2.3.5 Flux analysis of indene bioconversion in Rhodococcus sp. KY1 56
2.3.6 Summary 57

CHAPTER 3 - MATERIALS AND METHODS

3.1 CULTIVATION METHODS
3.1.1 Microorganisms and reagents 58
3.1.2 Media 59
3.1.3 Culture maintenance 60

3.2 INDENE BIOCONVERSIONS
3.2.1 Shake-flask indene bioconversions 61
3.2.2 Single phase indene bioconversions 62
3.2.3 Biphasic indene bioconversions 63

3.3 MOLECULAR BIOLOGY METHODS
3.3.1 DNA manipulation methods 64
   3.3.1.1 Preparation of plasmid DNA 65
   3.3.1.2 Preparation of genomic DNA from Rhodococcus 66

6
3.3.2 Transformation methods
  3.3.2.1 Preparation of competent *E. coli* JM109 79
  3.3.2.2 Transformation of competent *E. coli* JM109 79
  3.3.2.3 Preparation of competent *Rhodococcus* sps. I24 and KY1 79
  3.3.2.4 Transformation of competent *Rhodococcus* sps. I24 and KY1 80
3.3.3 Construction of *Rhodococcus erythropolis* limA 80
3.3.4 Plasmid construction 81

3.4 ANALYTICAL AND PREPARATIVE METHODS
  3.4.1 Biomass determination 86
  3.4.2 Off-gas analysis 87
  3.4.3 Protein analysis 87
  3.4.4 Preparation of indene and indene metabolites 88
    3.4.4.1 Synthesis of [14C]-indene 88
    3.4.4.2 Preparation of (1S,2R)-indan oxide and trans-(1R,2R)-indandiol 89
  3.4.5 Abiotic phase measurements 91
    3.4.5.1 Indene metabolites 91
    3.4.5.2 [14C]-Indene metabolites 92
    3.4.5.3 Glucose 93
    3.4.5.4 Ammonia 93
    3.4.5.5 Phosphate 98
  3.4.6 Enzymatic assays 98
    3.4.6.1 Preparation of cell lysates 98
    3.4.6.2 Indene monooxygenase assay 99
    3.4.6.3 Limonene-1,2-epoxide hydrolase 99
    3.4.6.4 Indan oxide cell-free hydrolysis 100

CHAPTER 4 - RESULTS AND DISCUSSION: DIRECTING THE STEREOSELECTIVE SYNTHESIS OF (2R)-INDANDIOL IN *RHODOCOCCUS*

4.1 STRATEGIES FOR MODIFICATION OF INDENE BIOCONVERSION IN *RHODOCOCCUS* sp. KY1 101
  4.1.1 Prioritization of targets in the indene bioconversion network 101
    4.1.1.1 Indan oxide node 102
    4.1.1.2 *cis*- (1S,2R)-indandiol dehydrogenase knockout 103
    4.1.1.3 Naphthalene dioxygenase knockout 104
    4.1.1.4 Prioritization of targets 104
  4.1.2 Heterologous expression of limonene-1,2-epoxide hydrolase 106
    4.1.2.1 *Escherichia coli* JM109 106
      4.1.2.1.1 Genetic confirmation 106
      4.1.2.1.2 Enzymatic activity 107
    4.1.2.2 *Rhodococcus* sp. KY1 108
      4.1.2.2.1 Genetic confirmation 108
4.1.2.2 Indene bioconversion screening of transformation candidates

4.1.3 Directed, non-enzymatic hydrolysis of indan oxide by changes in medium pH

4.1.4 Strategies for indene flux redistribution: Summary

4.2 INDENE BIOCONVERSION IN RHODOCOCCUS SP. KY1 AND RECOMBINANT DERIVATIVES

4.2.1 Effect of epoxide hydrolase expression on indene bioconversion

4.2.2 Effect of high-pH conditions on indene bioconversion

4.2.3 Biphasic indene bioconversions with KY1(pDS3) and high pH: Product selectivities

4.2.4 Recombinant indene bioconversions: Summary

4.3 REGULATION OF INDENE BIOCONVERSION IN RHODOCOCCUS SP. KY1 AND RECOMBINANT DERIVATIVES

4.3.1 Theory

4.3.2 Monooxygenase activity in KY1 indene bioconversion

4.3.3 Effect of trans-indandiol concentration on $k_{MO}$

4.3.4 Limonene-1,2-epoxide hydrolase kinetics

4.3.4.1 Effect of (1S,2R)-indan oxide

4.3.4.2 Effect of trans-(1R,2R)-indandiol

4.3.5 Regulation of indene bioconversion: Summary

4.4 DISCUSSION

CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS

5.1 SUMMARY

5.2 CONCLUSIONS

5.3 RECOMMENDATIONS FOR FUTURE WORK

REFERENCES

APPENDICES

APPENDIX A - NOMENCLATURE

APPENDIX B - CONSTRUCTION OF RHODOCOCCUS SP. I24(pDS3)

APPENDIX C - BIPHASIC INDENE BIOCONVERSION PROFILES

APPENDIX D - KY1 MONO ÖXYGENASE INHIBITION BY 1-KETO-2-HYDROXY-INDAN

APPENDIX E - RESINS FOR REMOVAL OF INDENE METABOLITES

APPENDIX F - PLASMID MAPS
LIST OF FIGURES

Figure 1.1 The basic premise of metabolic engineering is to introduce a complex reaction pathway that requires multiple conventional unit operations into a single microorganism or cell line. The mathematical tools to analyze the reaction pathways are similar between metabolic engineering and traditional chemical reaction engineering.

Figure 1.2 Chemical synthesis of indinavir sulfate and role of cis-(1S)-amino-(2R)-indanol [(-)-CAI], shaded throughout. (-)-CAI contributes two of the five chiral centers of indinavir sulfate (Reider 1997).

Figure 1.3 Possible synthetic routes to cis-aminoindanol from indene. The chemical synthesis of (-)-CAI (solid arrows) utilizes either a Jacobsen catalyst to give the intermediate (1S,2R)-indan oxide at 87% e.e., or hydrogen peroxide to give racemic indan oxide. Alternatively, stereospecific oxygenases (open arrows) can be used with indene to give (1S,2R)-indan oxide, cis-(1S,2R)-indandiol or trans-(1R,2R)-indandiol in a theoretical 100% yield, any of which are then reacted with acetonitrile to give (-)-CAI.

Figure 2.1 Key reactions in naphthalene and toluene degradative pathways in *Pseudomonas*. These reactions closely parallel those observed in *Rhodococcus* indene bioconversion. *Rhodococcus* sp. I24 can grow on both naphthalene and toluene, while *Rhodococcus* sp. KY1 grows on naphthalene.

Figure 2.2 Electron transport and relationship between subunits in naphthalene and toluene dioxygenase. Adapted from (Butler and Mason 1997).

Figure 2.3 Stereochemical control of geraniol N-phenylcarbamate-6,7-diol formation by *Aspergillus niger* (Zhang, Archelas et al. 1991).

Figure 2.4 Degradative pathways for (4R) and (4S)-limonene in *Rhodococcus erythropolis* DCL14. Enzymes are in bold type and stereochemistry is indicated on chemical structures. Adapted from (van der Werf, Swarts et al. 1999).

Figure 2.5 Indene bioconversion network for *Rhodococcus* sp. I24 proposed by (Chartrain, Jackey et al. 1998).

Figure 2.6 Indene metabolite profiles in the *Rhodococcus* sp. I24 chemostat at 0.10 h⁻¹ dilution rate. Behavior characteristic of strain KY1 is exhibited after 250 hours.
Figure 2.7  Indene bioconversion network in *Rhodococcus* sp. KY1. Metabolite excretion and uptake rates are denoted by $r_i$, while intracellular fluxes are written as $J_i$.  

Figure 2.8  Intracellular flux distribution for *Rhodococcus* sp. KY1 at 100 ppm indene air feed concentration and 0.065 h$^{-1}$ dilution rate. The fluxes were normalized by the indene uptake rate.  

Figure 2.9  Comparison of kinetic model predictions with experimental measurements of [14C]-indene metabolites for *Rhodococcus* sp. KY1 cells obtained from a chemostat at steady state obtained with a dilution rate of 0.065 h$^{-1}$ and 100 ppm indene air feed concentration. Reaction rate constants used in the kinetic model were determined from flux estimates as described in the text.  

Figure 3.1  Schematic of indene bioconversion system. Adapted from (Yanagimachi 2000).  

Figure 3.2  (A) Schematic of synthesis of *limA* gene by PCR (B) Correction of *limA* point mutations following ligation of *limA* PCR product into pCR-2.1TOPO. Shaded arrows indicate point errors detected through sequencing; the dashed segment of plimA8 was not sequenced.  

Figure 3.3  Diagram of pDS3, which expresses the *limA*-encoded epoxide hydrolase in *Rhodococcus* sp. KY1 from the constitutive lambda P$_{L}$ promoter. *aacC1*, gentamicin resistance marker from Tn1696; NG2 rep, origin of replication derived from pEP2.  

Figure 3.4  Typical biomass calibration curve for *Rhodococcus* sp. KY1.  

Figure 3.5  [14C]-indene synthesis scheme. Adapted from (Yanagimachi 2000).  

Figure 3.6  UV spectra for indene metabolites accumulated in *Rhodococcus* sps. I24 and KY1 cultures. (a) trans-indandiol, (b) cis-indandiol, (c) 1-keto-2-hydroxy-indan, (d) 1-indenol, (e) 1-indanone, (f) indan oxide, (g) indene.  

Figure 3.7  Zorbax RX-C8 reverse-phase HPLC chromatogram at 220 nm. Indene metabolites elute as follows: *trans*-indandiol (9.2 min), *cis*-indandiol (9.8 min), 1-keto-2-hydroxy-indan (11.1 min), 1-indenol (15.1 min), 1-indanone, (17.8 min), indan oxide (19.1 min), and indene (24.0 min).  

Figure 3.8  Elution of (a) *trans*-indandiol (1.1 min), *cis*-indandiol (1.2 min), 1-keto-2-hydroxy-indan (1.5 min), 1-indenol (2.5 min), 1-indanone, (2.8 min), and (b) indan oxide (3.6 min) and indene (4.2 min) on the YMC CombiScreen C8 HPLC column. Chromatograms shown are at 220 nm.
Figure 3.9 Elution of (a) cis-(1S,2R)-indandiol (22.2 min) and cis-(1R,2S)-indandiol (30.0 min) and (b) trans-(1R,2R)-indandiol (19.4 min) on Chiralcel OJ normal-phase HPLC column with 0.5 mL/min 9:1 hexanes:2-propanol. Chromatograms shown are at 220 nm.

Figure 4.1 Theoretical (2R)-indandiol selectivities in KY1 derivatives as a result of (a) epoxide hydrolase expression, (b) cis-(1S,2R)-indandiol dehydrogenase knockout, and (c) naphthalene-inducible dioxygenase knockout, based upon indene flux distribution metabolic flux analysis results reported previously (Figure 2.8).


Figure 4.3 Conversion of 2.0 mM (1S,2R)-indan oxide (83% e.e.) to trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol by lysates prepared from E. coli JM109 transformants. Residual indan oxide concentration not shown.

Figure 4.4 SDS-PAGE analysis of crude lysates prepared from JM109 transformants. Limonene-1,2-epoxide hydrolase corresponds to band at 17 kDa in JM109(pDS3) lane.

Figure 4.5 Restriction digests of KY1(pDS3) candidates. Only candidate “E” has correct digestion patterns with respect to intact pDS3, while candidates “B”, “D”, and “H” appear to have insertions near or in the NG2 replication origin.

Figure 4.6 Indene bioconversions by KY1(pDS3) transformation candidates, as denoted in restriction digests of Figure 4.5.

Figure 4.7 Conversion of 8.0 mM (1S,2R)-indene oxide (83% e.e.) to trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol by lysates prepared from Rhodococcus sp. KY1 transformants.

Figure 4.8 SDS-PAGE analysis of cleared lysates prepared from KY1 transformants.

Figure 4.9 First order kinetics of (1S,2R)-indane oxide non-enzymatic hydrolysis in cell-free defined medium at 30°C at (a) pH 10, (b) pH 9, (c) pH 8, (d) pH 7, (e) pH 6.

Figure 4.10 Effect of medium pH on (1S,2R)-indane oxide hydrolysis. 10 mM indane oxide was added to 10 mL culture medium adjusted to pH 5-10. Rate constants were derived from linear regression of semilog representations of indane oxide depletion data.

Figure 4.11 Proposed competing mechanisms for (1S,2R)-indane oxide hydrolysis.
Figure 4.12 (a) Indene bioconversion metabolite profiles for single-phase batch KY1 run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1, and biomass concentration of 7.0 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. The two products of non-enzymatic indan oxide hydrolysis, trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol, appear immediately upon the initiation of indene oxidation. (b) Gas and aqueous-phase indene concentrations. The gas inlet indene concentration was manually controlled.

Figure 4.13 (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pAL282) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1 and biomass concentration of 4.2 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. Indene bioconversion profiles are similar to those observed for KY1 (Figure 4.12). (b) Gas and aqueous-phase indene concentrations. The gas inlet indene concentration was manually controlled.

Figure 4.14 (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pDS3) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1, and biomass concentration of 4.9 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. At early culture times (up to approximately 25 h), the only products formed are indan oxide and trans-(1R,2R)-indandiol. (b) Gas and aqueous-phase indene concentrations.

Figure 4.15 Epoxide hydrolase activity in cell lysates made from cells removed from KY1(pDS3) bioconversion (Figure 4.14) at 90.5 h. Cells were washed during the lysing procedure to remove residual indene and metabolites. 8.0 mM indan oxide was added to the lysates at the beginning of the assaying procedure.

Figure 4.16 (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pDS3) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm) and pH 8.6 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. (b) Gas and aqueous-phase indene concentrations.

Figure 4.17 trans-indandiol selectivity profiles for batch indene bioconversions performed in (a) single-phase (aqueous) and (b) biphasic (aqueous/silicon oil) modes.
Figure 4.18  (a) Indene metabolite bioconversion profiles for KY1 fed-batch culture. *trans*-indandiol accumulates to approximately 1.2 g/L and 1-keto-2-hydroxy-indan accumulates to 1.6 g/L. (b) Gas-phase indene concentrations.

Figure 4.19  [*14C*]-indene oxidation kinetics to all detectable radiolabeled metabolites in cells removed from KY1 fed-batch culture (Figure 4.18) 22 h after [*14C*]-indene addition. [*14C*]-indene depletion data was used to calculate $k_{MO} = 10.1 \text{ h}^{-1}$ from Eq. (4.4).

Figure 4.20  Observed monooxygenase rate constant values in KY1 fed-batch culture. The concentration of active enzyme declines rapidly as *trans*-(1R,2R)-indandiol and 1-keto-2-hydroxy-indan accumulate in the culture.

Figure 4.21  [*14C*]-indene oxidation kinetics for *trans*-indandiol inhibition experiment in cells removed from indene-induced chemostat culture. Curves were fit to tracer data by non-linear regression using Eqs. (2.8-2.13).  (a) Control condition with no *trans*-indandiol added ($k_{MO} = 3.3 \text{ h}^{-1}$).  (b) 1.7 g/L *trans*-(1R,2R)-indandiol (run in triplicate) incubated with cells for 30 min prior to tracer addition ($k_{MO} = 0.82 \text{ h}^{-1}$).

Figure 4.22  Inhibition of indene monooxygenase activity by *trans*-indandiol.

Figure 4.23  Determination of kinetic parameters for limonene-1,2-epoxide hydrolase *in vitro* assay on (1S,2R)-indan oxide.

Figure B.1  (a) Southern blot of recombinant I24(pDS3). Lanes are as follows: (1) plasmid control, (2) HinD III digest of I24(pDS3) DNA, (3) Sac II digest of I24(pDS3) DNA.  (b) Restriction digest batteries of I24(pDS3): (4) BstE III digest of λ, (5-9) Hind III and four replicates, (10) Hae III digest of *φX174*.

Figure C.1  Indene bioconversion metabolite profiles for biphasic batch KY1 run in defined medium at pH 7.0 ± 0.1. Time is indicated as post-addition of indene to the cells. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.

Figure C.2  Indene bioconversion metabolite profiles for biphasic batch KY1 run in defined medium at pH 8.5 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.
Figure C.3  Indene bioconversion metabolite profiles for biphasic batch KY1(pDS3) run in defined medium at pH 7.0 ± 0.1. Time is indicated as post-addition of indene to the cells. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.

Figure C.4  Indene bioconversion metabolite profiles for biphasic batch KY1(pDS3) run in defined medium at pH 8.5 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.

Figure D.1  Inhibition of indene monooxygenase activity by 1-keto-2-hydroxy-indan.

Figure E.1  Proposed binding mechanism of cis-(1S,2R)-indandiol to phenylaminoboronate. The hydroxylation of boronate to form a tetrahedral boronate intermediate is favored under basic conditions. The cyclic ester is hydrolyzed using 25 mM HCl. X is the polyacrylamide bead.
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>Toluene degradative enzymes.</td>
<td>42</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>log $P_{OW}$ values for compounds structurally similar to indene.</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Adapted from (Mannhold, Rekker et al. 1998).</td>
<td></td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Steady state concentrations in <em>Rhodococcus</em> sp. KY1 chemostat cultures.</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Adapted from (Yanagimachi 2000).</td>
<td></td>
</tr>
<tr>
<td>Table 2.4</td>
<td>Conversion of 25 µM [14C]-indene to primary oxygenated products under</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>different inducers after 5 minutes (reported as percentage of tracer added).</td>
<td></td>
</tr>
<tr>
<td>Table 2.5</td>
<td>Steady state metabolite balance equations for the KY1 indene bioconversion network.</td>
<td>67</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Strains, plasmids, and primers used in this study.</td>
<td>83</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Components of gels used for SDS-PAGE analysis.</td>
<td>88</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>(2R)-indandiol yields and selectivities in <em>Rhodococcus</em> sps. KY1 and I24 chemostat cultures. Steady state metabolite data from (Yanagimachi 2000).</td>
<td>117</td>
</tr>
</tbody>
</table>
CHAPTER 1 - INTRODUCTION

1.1 Metabolic engineering: A new frontier of chemical reaction engineering

What is "metabolic engineering"? Though it had not been described as such at the time, the production of antibiotics in large scale as early as the 1940s, fermentation processes in breweries for centuries, and more recently the introduction of bioprocesses for recombinant insulin production are early examples of metabolic engineering. Process improvements in these applications were largely empirical - optimization of a few key parameters such as pH, dissolved oxygen, agitation, and substrate selection was sufficient to improve production adequately. What was likely recognized at the time, but not addressed explicitly, was that optimization of the aforementioned parameters did not merely affect a single $A \rightarrow B$ transformation as would presumably be the case in a non-enzymatic chemical transformation, but instead impacted a whole series of bioreactions as would become more evident through the extensive metabolic studies of the 1950s and 1960s. In light of these developments and work investigating the control structures of metabolic networks (Savageau 1976), the science of metabolic engineering was first explicitly defined as follows (Bailey 1991):

"Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology."

Together with an accompanying paper (Stephanopoulos and Vallino 1991), the need for studying entire metabolic networks for improving metabolite production was recognized in contrast to a focus on a single gene or enzyme. This network analysis approach has been applied and developed in a number of applications as reviewed recently (Stephanopoulos 1999).

Still, metabolic engineering has frequently been limited to methodological issues pertaining to gene introduction, expression and control by way of molecular biological techniques. While such methods constitute an enabling technology that is critical for the
implementation of strategies for pathway design and modification, an equally critical activity is the quantitative systems analysis of metabolic networks and the evaluation of the effects brought about by genetic modifications. In this task, metabolic engineering can benefit from concepts and methods previously established in chemical reaction engineering, a science in which the analytical tools for analysis of complex chemical reaction networks have been developed extensively. In the simplest terms, metabolic engineering is concerned with designing and harnessing complex reaction networks confined in a single “unit-operation,” i.e. performing multiple chemical transformations in a living cell, while non-biological reaction sequences may typically require multiple catalysts and/or reactor vessels (Figure 1.1). However, beyond this basic technical difference, there are several important distinctions between metabolic networks and the reaction pathways traditionally studied in chemical reaction engineering.

![Diagram of metabolic pathways](image)

**Figure 1.1** The basic premise of metabolic engineering is to introduce a complex reaction pathway that requires multiple conventional unit operations into a single microorganism or cell line. The mathematical tools to analyze the reaction pathways are similar between metabolic engineering and traditional chemical reaction engineering.
One obvious distinction of metabolic engineering is that it is concerned with reactions catalyzed by enzymes – it is the set of enzymes expressed within a particular cell that defines the overall conversion process. This enzymatic profile determines the exact steps by which the cell converts a compound to products either as part of its intrinsic metabolic functions or through newly introduced foreign genes. As a result, a conversion process comprises a set of well-defined steps that are subject to further modification and control at the genetic level toward enhancing the pathway(s) of interest. The tools of genetic engineering and protein engineering allow one to manipulate the profile of the overall cellular catalyst and the activity of individual enzymes, respectively.

Other distinct characteristics of cellular systems relative to non-enzymatic systems are the hierarchical nature of activation of various classes of molecules, and the presence of extensive feedback control loops. That is, genes are activated to transcribe mRNA messages that in turn are used to express proteins responsible for vital cellular functions such as metabolic conversions or transmission of signals. The individual small metabolite molecules thus affected by such reactions modulate the activity of enzymes and, along with other regulatory factors, regulate the activity of genes. These processes can be conceptualized as occurring at four distinct but interacting layers of hierarchy: genes, transcripts (mRNA), proteins and small metabolites. While this organization is useful in describing the dominant hierarchical structure of cellular reactions, it is important to note that there is extensive interaction among the classes of molecules at all levels, primarily through modulation of the processes of gene expression, translation and protein activity by gene products and small metabolites.

Chemical reaction engineering is similarly concerned with chemical reaction networks relying heavily on measurements of reaction intermediates to develop reaction rate expressions and models for system behavior. Until recently, systems analysis of biological reaction networks was limited by the availability of intracellular compound measurements of the resolution and magnitude required to meaningfully decipher the complexity of such systems. This situation, however, is rapidly changing as DNA microarrays will soon allow quantitative measurement of mRNA transcript levels and
sophisticated isotopic methods provide reliable estimates of the rate of enzymatic reactions in vivo (or metabolic fluxes, see Chapter 2). These measurements, coupled with similar efforts for the quantitative measurement of protein level and activity (proteomics), promise to generate data of unprecedented completeness about all levels of cellular function and dramatically aid the analytical side of metabolic engineering.

1.2 Extension of metabolic engineering principles to biocatalytic reaction networks

While there are numerous examples in the literature of exhaustive analyses of metabolic flux distributions in model systems, and many others where gene overexpression results in varied improvements in production of a given metabolite, there are few studies that seek to unify the quantitative analysis and applied molecular biology components of metabolic engineering. Furthermore, previous examples of metabolic engineering have primarily dealt with relatively well-defined metabolic reaction networks with ample information available regarding the key genes and enzymes of a given biotransformation. This is because in most cases, these systems were analyzed with respect to pathways involved in primary carbon metabolism as opposed to secondary biocatalytic pathways. Also, previous studies have been performed in strains or cell lines with important molecular biology tools developed to routinely overexpress or delete desired genes.

Unfortunately, while tremendous strides have been made to develop molecular biology tools and in genome sequencing projects, there remain far more biological systems untapped for production of fine or commodity chemicals than have been studied extensively, let alone be discovered. Consequently, a framework for metabolic engineering of a priori poorly characterized or otherwise novel systems must be developed. That is, an approach to methodically identify genetic targets to modify and improve metabolic phenotypes must be such that it can be executed with minimal information regarding the genetics of the microorganism. As such, it is imperative that one be able to prioritize network targets for modification in light of the current state of
knowledge of a given biological system. This requires a systems analysis of bioreaction pathways to (1) a priori predict the effect of network modifications on the bioconversion phenotype, (2) formulate approaches for phenotype improvement, i.e. redistribution of metabolic flux, and (3) evaluate process engineering opportunities to improve product yield(s).

1.2.1 Biocatalytic synthesis of chiral compounds

A major opportunity to extend the scope of metabolic engineering is in the development of technology to produce enantiomerically pure compounds that are important to the manufacturing of many pharmaceuticals, amino-acid derived chemicals, and polymers. Dihydroxylated compounds are of particular value as they can be used as substrates in a diverse set of synthetic applications that conserve the stereochemical configuration of the hydroxylated substrate. The versatility of these compounds and demand for them in chiral syntheses has necessitated the development of innovative strategies for production-scale oxygenation of hydrocarbon compounds.

Potassium permanganate and osmium tetroxide catalysts are established tools to synthetically produce cis-diols from unsaturated hydrocarbons, albeit with varied enantiomeric purity (Norpoth, Kniepschild et al. 1972; Kobayashi, Ishida et al. 2001). Complementing this technology is the recent development of manganese-salen(V) catalysts that can form epoxides from hydrocarbon substrates in high enantiomeric excess (Tokunaga, Larrow et al. 1997; Jacobsen 2000). These arene oxides can be utilized to produce trans-diols by acid-catalyzed hydrolysis. However, these methods frequently suffer from low overall yield due to the requirement to resolve enantiomeric mixtures, and many catalysts are toxic and/or require adverse reaction conditions. A biocatalytic approach using oxygenases potentially offers distinct advantages over synthetic organic chemistry for chiral synthesis because bioconversions often can be carried out at milder reaction conditions and with higher product yield. The stereoselectivity of many enzymes and the relative ease in making a biocatalyst makes this an appealing technology for the synthesis of chiral compounds.
1.2.2 Protease inhibitors for treatment of HIV infection

Protease inhibitors are well-characterized chiral drugs in terms of their mechanism of action. This class of drugs was created from structure-based drug design programs, and was first established by the design of the angiotension-converting enzyme (ACE) inhibitors captopril and enalapril (Bristol-Myers Squibb) (deStevens 1990; Kubinyi 1999). Although the structure of the angiotension-converting enzyme (excessive activity of which is implicated as a major contributor to hypertension) had not been solved at the time of the design, key features of potential ACE inhibitors were identified after the enzyme was modeled based on sequence homology with a zinc metalloenzyme, Carboxypeptidase A.

An important new class of protease inhibitors comprises molecules designed as transition state analogs to combat HIV infection (Beach 1998). The HIV protease is an aspartic acid protease that cleaves the peptide bond between proline residues and tryptophan or phenylalanine residues (Vacca and Condra 1997; Kakuda, Struble et al. 1998). This protease is a key viral enzyme required to cleave the nonfunctional gag and gag-pol HIV polyproteins into mature viral structural proteins and enzymes that are critical for viral maturation and replication (Kohl, Emini et al. 1988; Seelmeier, Schmidt et al. 1988; Kling 2000). Four inhibitors of HIV protease have been designed and FDA-approved for human use: saquinavir (Hoffmann-La Roche), ritonavir (Abbott Laboratories), nelfinavir (Agouron), and indinavir (Merck and Co., Inc.).

1.2.3 1,2-Indandiol: A key precursor in indinavir sulfate synthesis

Indinavir sulfate (Crixivan®) contains five chiral centers that must be of a specific orientation for the molecule to have the desired therapeutic effect. The high degree of chirality in the indinavir sulfate molecule poses a significant hurdle in devising a synthetic pathway that gives an acceptable yield. Manufacturing processes for chiral compounds involving chemical synthesis steps can be quite inefficient due to yield reduction caused by racemization at each step where a chiral center is formed. In one possible chemical synthesis of indinavir sulfate, a piperazine moiety contributing one
chiral center is joined to an epoxide containing the remaining four chiral centers (Reider 1997) (Figure 1.2). The epoxide is derived from cis-(15,2R)-1-amino-2-indanol [(-)-CAI], an indene derivative that contributes two chiral centers to indinavir sulfate.

![Chemical diagram showing the synthesis of indinavir sulfate and the role of cis-(1S)-amino-(2R)-indanol [(-)-CAI].]

**CRIXIVAN (indinavir sulfate)**

**Figure 1.2** Chemical synthesis of indinavir sulfate and role of cis-(1S)-amino-(2R)-indanol [(-)-CAI], shaded throughout. (-)-CAI contributes two of the five chiral centers of indinavir sulfate (Reider 1997).

The synthesis of (-)-CAI may be carried out by the asymmetric epoxidation of indene using a Jacobsen catalyst (Hughes, Smith et al. 1997) by way of hydrolytic kinetic resolution to (1S,2R)-indan oxide at 87% enantiomeric excess (e.e.), followed by crystallization to purify the (1S,2R) enantiomer (Senanayake, Roberts et al. 1995). (1S,2R)-indan oxide is then converted to (-)-CAI by reacting with acetonitrile in a Ritter reaction in an overall process yield of approximately 50%. Alternatively, to avoid the use of the manganese-salen(V) catalyst, racemic indan oxide can be prepared by reacting indene with hydrogen peroxide in acetonitrile. Predictably, however, the use of H2O2...
ultimately results in even lower yields of \((-\)-CAI due to the need to resolve racemic mixtures of indan oxide\(^1\).

To circumvent the technically demanding chemical synthesis of \((-\)-CAI and reduce product loss, Merck scientists conceptualized a biocconversion process in which indene is oxidized to one of three derivatives that can serve as precursors to \((-\)-CAI: \textit{cis-}\((1S,2R)\)-indandioli, \textit{trans-}\((1R,2R)\)-indandioli, or \((1S,2R)\)-idan oxide (Figure 1.3). A microbial screening program was undertaken to isolate strains able to dihydroxylate indene stereoselectively at up to 100\% theoretical yield. Oxygenases were identified in isolates of the genus \textit{Pseudomonas} and \textit{Rhodococcus} that can catalyze these transformations. Either the \textit{cis} or \textit{trans} enantiomer of \((2R)\)-indandioli can be used for the synthesis of \((-\)-CAI because of racemization at C-1 that occurs during the \((2R)\)-indandioli to \((-\)-CAI transformation (Senanayake, Roberts et al. 1995). Because of an increased ability to tolerate high concentrations of indene and indene metabolites, and more desirable indene oxidation profiles, \textit{Rhodococcus} isolates were identified as the most promising microorganisms for the biocatalytic production of \((2R)\)-indandioli. However, these strains produce a number of side products that reduce overall product yield. Thus, a program was initiated to analyze the indene bioconversion network of \textit{Rhodococcus} to enable construction of a robust bioconversion strain for \((2R)\)-indandioli biosynthesis.

### 1.3 Thesis objectives

The aim of this thesis is to develop and apply a general framework for the optimization of bioconversion systems, particularly those biocatalysts with an \textit{a priori} undeveloped database of biochemical and genetic information. This framework is comprised of five essential steps:

1. Establishment of an experimental system for strain selection and metabolic network analysis

\(^1\) The nomenclature appearing in the literature for the indene derivative epoxidized at C-1 and C-2 varies substantially. This entity has been referred to as "indene oxide," "indane oxide," "1,2-epoxy-indan," and "idan oxide." In this thesis "idan oxide" is used to refer to this compound despite the convention of a particular reference(s).
(ii) Definition of the bioconversion network and quantification of network fluxes

(iii) Target identification

(iv) Flux redistribution

(v) Analysis of the modified bioconversion phenotype to identify areas for further strain and process improvement.

The framework is developed in the context of a systematic evaluation of the physiology of *Rhodococcus* strains and analysis of the relative metabolic fluxes of the indene bioconversion network. This is designed to facilitate the improvement of (2R)-indandiol production suitable for Crixivan® manufacturing through the development of genetic and/or non-enzymatic approaches.

![Diagram](image)

**Figure 1.3** Possible synthetic routes to *cis*-aminoinanol from indene. The chemical synthesis of (-)-CAI (solid arrows) utilizes either a Jacobsen catalyst to give the intermediate (1S,2R)-inden oxide at 87% e.e., or hydrogen peroxide to give racemic indan oxide. Alternatively, stereospecific oxygenases (open arrows) can be used with indene to give (1S,2R)-inden oxide, *cis*-(1S,2R)-indandiol or *trans*-(1R,2R)-indandiol in a theoretical 100% yield, any of which are then reacted with acetonitrile to give (-)-CAI.
Previous research is used to address the first two steps in this framework (Yanagimachi 2000). As discussed in more detail in Chapter 2, a novel chemostat system was developed and used for both analysis of steady state physiology and for the generation of a mutant *Rhodococcus* strain, denoted KY1, with a drastically improved chiral indandiol product profile through the application of proper selective pressure to *Rhodococcus* sp. I24 cultures. This represented a novel application of the chemostat system that differs from the traditionally exclusive use of the chemostat for the study of cell physiology. $^{14}$C-tracer studies, used for (a) indene bioconversion network definition and (b) the direct determination of steady state bioconversion fluxes, demonstrated that the improvement observed in the mutant KY1 strain with respect to indene bioconversion was most likely due to a stereospecific monooxygenase of high activity catalyzing the conversion of indene to indan oxide of (2R) chirality, and the elimination of a competing toluene-induced dioxygenase pathway present in I24. This resulted in doubling the (2R)-indandiol product yield with no significant change in productivity.

To continue the development of the aforementioned metabolic engineering framework, three specific objectives were formulated with respect to indene bioconversion in *Rhodococcus* sp. KY1:

1. Identify the key node in the *Rhodococcus* sp. KY1 indene bioconversion network for improved (2R)-indandiol biosynthesis.

2. Develop genetic and/or chemical approaches for redistributing indene bioconversion fluxes toward (2R)-indandiol formation.

3. Analyze the resultant phenotype to confirm flux redistribution and to identify further opportunities for (2R)-indandiol productivity improvement through process engineering or the application of emerging technologies.
1.4 Thesis organization

In Chapter 2, the role of a directed metabolic engineering approach in biocatalytic systems is delineated, and the roles of genomic and proteomic technologies in metabolic engineering and systems analysis of bioreaction networks are highlighted in the context of recent advances in this field. Next, known enzymes and microorganisms for biocatalytic oxidations particularly relevant to (2R)-indandiol biosynthesis are described. Finally, previous accomplishments in the study of indene bioconversion in *Rhodococcus* are discussed. These include the isolation of the *Rhodococcus* sp. KY1 strain from chemostat cultures of *Rhodococcus* sp. I24, as well as the systematic analysis of the indene bioconversion networks of both strains using $^{14}$C-tracers and metabolic flux analysis.

Chapter 3 presents the methods used in this research, including the details of the genetic constructions used to modulate the indene bioconversion phenotype of *Rhodococcus* sp. KY1. Chapter 4 begins with analysis of the results of previous research that identified the relative flux distributions of indene oxidation in the KY1 strain, by determining the impact these data have on the selection of targets in KY1 for improving (2R)-indandiol production. The main results of the research are then presented in three parts: (a) the development of two independent strategies for modifying flux distribution in the indene bioconversion network, (b) the results of implementation of these strategies in batch indene bioconversions, and (c) analysis of the enzyme kinetics and inhibition characteristics of the native monooxygenase and a newly introduced epoxide hydrolase enzyme in the context of the resultant phenotypes. Chapter 5 then presents the conclusions of this study and recommendations for further process development opportunities for indene bioconversion in *Rhodococcus*.
Portions of the following chapter are included in two publications:


CHAPTER 2 - METABOLIC ENGINEERING: A PARADIGM FOR INDENE BIOCONVERSION ANALYSIS

Equally important to integrating emerging technologies, largely resulting from advances in genomics, into the paradigm of metabolic engineering is the further development of the analytical methods necessary to analyze the biological systems of future industrial interest. Among those systems are soil microorganisms that are known to express many important oxidative enzymes. Although a number of novel enzymes have already been characterized in these strains, screening studies that yield microorganisms capable of oxidizing unusual substrates will continue to generate many new potential industrial biocatalysts with little information established regarding their genetic and biochemical properties. This will require preliminary focused studies to unravel relevant biochemistry, and parallel development of biological tools needed to facilitate subsequent strain modification toward more favorable production phenotypes.

The indene bioconversion network that is expressed in *Rhodococcus* sp. is an ideal system for developing biocatalytic metabolic engineering principles for uncharacterized systems. While the indene bioconversion network in *Rhodococcus* is relatively complex, with parallel oxygenation reactions on a common substrate leading to indandiols of multiple chiralities, the network is decoupled from primary metabolism in terms of carbon flux since indene cannot be used as a growth substrate. Consequently, a relatively small number of reactions are under consideration for amplification or knockout with respect to the hundreds of transformations that are a part of central carbon metabolism. This chapter provides background for the study and modification of *Rhodococcus* sp. by discussing the nature of the key enzymes, the oxygenases and epoxide hydrolases, important to this research. Additionally, an overview of the key biocatalytic microorganisms *Pseudomonas* and *Rhodococcus* is presented.

Following this background discussion, the research providing support for the first two steps of the metabolic engineering framework being developed in this thesis, (i) establishment of an experimental system for strain selection and metabolic network analysis and (ii) definition of the bioconversion network and quantification of network
fluxes, is presented. These steps are accomplished through the novel application of radiolabeled tracers to biocatalytic flux analysis in concert with more conventional measurements of substrate uptake and metabolite excretion rates using a steady state chemostat.

First, however, the fundamental tenants of metabolic engineering – the measurement and calculation of metabolic fluxes and metabolic control - are described with examples of past and current research in the area. Advances in genomics, proteomics, directed evolution and other technologies are described in the context of how they can be integrated into the current metabolic engineering paradigm.

2.1 Metabolic engineering as an integrating platform for strain development

The integration of the analytical framework and experimental tools of metabolic engineering with emerging technologies such as DNA microarrays and directed evolution stand to dramatically improve the approaches by which strain improvement and biocatalyst design are pursued in the future. Progress in genomics and applied molecular biology, together with increasing emphasis on renewable resource utilization for chemical production, has advanced metabolic engineering to the forefront of biotechnological interest.

2.1.1 Introduction

Metabolic engineering has emerged in the past decade as the interdisciplinary field aiming at improving cellular properties by pathway modifications using modern genetic tools (Bailey 1991; Stephanopoulos, Aristidou et al. 1998). An impressive portfolio of applications has resulted from this activity. On the experimental side, applications have focused on pathway modifications in (primarily) prokaryotic (Vallino and Stephanopoulos 1993; Cane, Walsh et al. 1998; Stephanopoulos, Aristidou et al. 1998; Altaras and Cameron 1999; Dong-Eun, Heung-Chae et al. 1999; Stephanopoulos 1999; Farmer and Liao 2000; Ostergaard, Olsson et al. 2000) and eukaryotic systems (Fussenegger, Moser et al. 1998; Shanks, Bhadra et al. 1998; Follstad, Balcarcel et al. 1999;
Fussenegger 2001), and the development of new tools for the introduction of genetic controls (Keasling 1999) and assessment of the resulting cell physiology. Mathematical methods have been used to elucidate the structure of metabolic pathways (Stephanopoulos and Vallino 1991; Simpson, Follstad et al. 1999), distribution of kinetic control in metabolic pathways (Fell 1992; Jeneson, Westerhoff et al. 2000) and to facilitate the in vivo determination of intracellular fluxes using metabolite and isotopomer balances in conjunction with tracer enrichment and mass isotopomer measurements using NMR spectroscopy and GC-MS (Delgado and Liao 1992; Vallino and Stephanopoulos 1993; Zupke and Stephanopoulos 1995; Szyperski, Bailey et al. 1996; Schmidt, Carlsen et al. 1997; Wiechert, Siefke et al. 1997; Follstad, Balcarcel et al. 1999; Kelleher 1999; Klapa, Park et al. 1999; Stephanopoulos 1999; Wittmann and Heinzle 1999; Christensen and Nielsen 2000; Petersen, de Graaf et al. 2000). The field is poised to benefit further from developments in genomics and proteomics and new technologies that probe the expression phenotype of cells, such as DNA microarrays.

Many industrially relevant problems in biotechnology are well-addressed within a metabolic engineering framework that both identifies possibly unknown bioreaction pathways and elucidates reaction and/or transport bottlenecks and metabolic regulation of pathways (Bailey 1991; Stephanopoulos 1999). It is through a quantitative analysis and understanding of a bioreaction network that genetic modifications can be targeted to improve strain phenotype using modern tools from applied molecular biology (Stephanopoulos and Vallino 1991; Simpson, Follstad et al. 1999). The resulting recombinant strains need, in turn, to be similarly analyzed to assess the effect(s) of specific genotype modifications (Stephanopoulos, Aristidou et al. 1998). An important outcome of such genotype change-phenotype assessment cycles is the elucidation of function-genotype relationships. In this sense, metabolic engineering has been at the forefront of functional genomics, an activity that has recently attracted broad interest in response to genome sequencing activities (Bailey 1999).

In the future, the engineering of pathways in strains will be facilitated by the availability of genes and gene expression data. Presently, the sequencing of approximately 40 prokaryotic genomes have been completed with hundreds more
currently in progress (TIGR Microbial Database, http://www.tigr.org/tdbmdb/ mdbcomplete.html). Quantitative analysis of transcription patterns characteristic of different physiological states (particularly desired states), so defined by metabolic fluxes, together with characterization of metabolic networks and control exerted therein, will potentially help identify key genetic targets for improving biocatalyst phenotype. There holds tremendous promise to not only be able to improve cell properties through rational target selection and introduction of genetic controls, but also through rapidly developing combinatorial technologies such as gene shuffling and directed evolution.

This section is intended to highlight the drivers and required tools for future metabolic engineering developments, in the context of recent advances in selected areas of particular importance to industrial microbiology. The mathematical framework of metabolic engineering has been significantly developed, and now depends on the design of new experimental technologies to fuel further advancements. At the core of far-reaching applications in fine chemicals, pharmaceuticals, biopolymers, and medicine will be the key characteristics of integration and quantification that will be increasingly recognized as indispensable components of a systems biological approach to cell improvement, for use as industrial biocatalysts.

2.1.2 Drivers of metabolic engineering

2.1.2.1 Genomics

Perhaps the most important driver in further development of metabolic engineering is the field of genomics. Genome sequencing will provide access to genes and intergenic control elements important for transcriptional control of individual genes and whole pathways. Important technologies utilizing this information are being developed, most notably among them DNA microarrays (Brown and Botstein 1999). The most direct application of microarrays to metabolic engineering will be to identify discriminatory genes characteristic of desired physiological states, such as those contributing to high productivity. This will be accomplished primarily through extensive transcriptional studies under a variety of environmental conditions and
genetic backgrounds. As transcriptional data accumulates, it will also be possible to reconstruct genetic regulatory networks.

Besides the straightforward applications of microarrays described above, there are many other, as yet unconventional, applications where microarrays will play a defining role in pathway elucidation. Among them are the uses of microarrays for validation of gene function (De Backer, Nelissen et al. 2001), determination of membrane-bound vs. cytosolic proteins (Maximillian, Eisen et al. 2000), and characterization of DNA binding proteins (Ren, Robert et al. 2000). Furthermore, there is a broad spectrum of opportunities to use microarrays in a more creative way, such as in clever screening methods, much like drug discovery or strain selection. In keeping with the systems emphasis of metabolic engineering, gene expression data will play an important, though not absolute, role in enabling the elucidation of metabolic control structures. This task will also require focused protein measurements (not to be confused with proteomics) to permit one to establish a more rigorous link between genome and metabolome. Quantitative transcription pattern analysis of different physiological states, in conjunction with characterization of metabolic networks, can potentially be used to prioritize key genetic and metabolic targets for improving biocatalyst profiles. Metabolic engineering is an ideal framework for the integration of these diverse data sets.

2.1.2.2 Applied molecular biology

The second important driver of metabolic engineering is the continuing development of applied molecular biology. Such advances, the result of a strong commitment by industry and the federal government to life sciences research, will facilitate genetic constructions required for the synthesis of improved biocatalysts through controlled gene expression (Keasling 1999; Ostergaard, Olsson et al. 2000). Another significant development is the introduction of combinatorial methods, such as directed evolution and gene shuffling techniques, a notable departure from conventional target-based pathway modification (Tobin, Gustafsson et al. 2000). These methods have succeeded in dramatically improving enzymatic activities and regulatory properties, and have great promise for many further applications in their own right or in combination.
with rational approaches for the modification of entire metabolic pathways. One example, that demonstrates the application of both rational targeting and combinatorial methods for pathway optimization, is the production of lycopene and other carotenoids (Farmer and Liao 2000; Wang, Oh et al. 2000). Of particular interest, in the context of pathway engineering, is the directed evolution of enzymes to modify cofactor requirements (Joo, Lin et al. 1999) and to broaden substrate specificity (Jurgens, Strom et al. 2000) or enantioselectivity. Such a combinatorial approach may also be useful in the optimization of entire metabolic pathways as opposed to single enzymes (Schmidt-Dannert, Umeno et al. 2000). However, more robust methods are needed to extend these concepts and experimental base to truly interconnected pathways that go beyond simple linear biosynthetic structures.

2.1.2.3 Renewable resource utilization

A third driver will be the increased attention to renewable resources for the synthesis of specialty and commodity chemicals by “green” processes. Agricultural raw materials are a rich, renewable source of carbohydrates that could provide the basis for the production of many chemicals. The most efficient route for the utilization of these resources is through biotechnological methods, provided that suitable pathways are engineered. One prominent example is the continuing development of biological approaches to the synthesis of propanediols. The microbial production of 1,3-propanediol, an intermediate in the synthesis of polyesters and polyurethanes, has been a longtime focus of metabolic engineering efforts. More recently, a number of studies have focused on developing a biological route to 1,2-propanediol that is chemically derived from propylene (Altaras and Cameron 1999; Altaras and Cameron 2000). Other compounds of interest in this context are lysine (Vallino and Stephanopoulos 1993), lactate (Dong-Eun, Heung-Chae et al. 1999), and biopolymers such as polyhydroxyalkanoates (Mee-Jung, Sun et al. 2001). Additionally, much interest remains in the development of alternative fuels that can be made biologically from readily available carbohydrates.
2.1.2.4 Emerging opportunities

Finally, metabolic engineering has the potential to produce a variety of high-added value compounds for clinical and pharmaceutical use. Prominent among them are optically pure compounds as final products or precursors for chiral pharmaceutical manufacturing. One area of research that may provide a source of important structures for pharmaceutical applications is the combinatorial biosynthesis of complex polyketide structures starting from simple precursors such as acetyl CoA, proponyl CoA, and butyryl CoA and derivatives (Cane, Walsh et al. 1998; Khosla 2000). Polyketide synthases (PKSs) have a modular structure, where each module contains multiple active sites responsible for chain elongation. The nature of these enzymes, together with recent studies indicating the potential for engineering PKSs with heterologous modules provided appropriate “linker” polypeptides are present on the catalyst, makes metabolic engineering of polyketide biosynthetic pathways a promising approach to the design of biosynthetic pathways for the synthesis of novel pharmaceuticals. Work is continuing to elucidate PKS structure-function relationships (Zawada and Khosla 1999). Another example, the focus of this thesis, is the development of a biocatalyst for the stereoselective synthesis of 1,2-indandiol with (2R) chirality for use as a chemical precursor in the asymmetric synthesis of indinavir sulfate (Reider 1997). These efforts represent an attempt to develop metabolic engineering into an enabling platform technology for chiral pharmaceutical synthesis. The above are two among many possible applications of metabolic engineering for production of high-added value compounds. In summary, the driving forces described here will contribute to generating more opportunities that will intensify efforts on pathway modification and metabolic engineering in general.

2.1.3 Tools of metabolic engineering

As mentioned earlier in this chapter, metabolic engineering is concerned with modifying pathways and assessing the physiological outcome of such genetic modifications. To develop new and useful cellular traits, researchers have access to
numerous experimental and computational tools, but some additional methods are needed. Among the molecular biological tools available are efficient transformation systems, including low copy number plasmids, and methods of transcriptional control by modification of mRNA stability (Keasling 1999). The availability of efficient tunable promoters is important in metabolic engineering because one is often interested in introducing relatively small changes in gene expression as opposed to many-fold overexpressions typically reported with most systems. Attaining the goal of overproduction or cellular property improvement often requires a fine modulation of individual enzyme activity that in turn needs accurate tunable promoters. The availability of such promoters will also permit systematic studies of genotype-phenotype relationships, which is a central theme of metabolic engineering.

Equally important are the mathematical and experimental means to facilitate pathway structure analysis and the measurement of pathway fluxes and flux distributions. Fluxes are a fundamental feature of metabolic pathways and cell physiology, and a critical parameter in the evaluation of the cellular phenotype and the elucidation of flux control. Therefore, methods for flux determination (frequently referred to as metabolic flux analysis) are an indispensable component of metabolic engineering. Metabolic flux quantification methodologies are currently limited to metabolic steady state conditions through stoichiometric balances around intracellular metabolites and measurements of metabolite excretion and substrate uptake rates. These data are not sufficient to decipher the flux carried out by metabolic cycles or parallel pathways that diverge and converge later to the same metabolite, or to describe the reversibility of intracellular reactions (de Atauri, Curto et al. 1999). However, isotopic tracers can be used to address many of these limitations. Using $^{13}$C-compounds labeled on specific atom(s), metabolic fluxes can be mapped on the isotopomer fractions of a network of pathways through isotopomer balances (Schmidt, Carlsen et al. 1997; Klapa, Park et al. 1999) deduced from measurements using either $^{13}$C-NMR (carbon enrichment (Petersen, de Graaf et al. 2000), multiplet pattern (Klapa, Park et al. 1999), or 2D-NMR (Szyperski, Bailey et al. 1996)) or mass spectrometry (mass isotopomer fractions) (Wittmann and Heinzle 1999). $^{13}$C-NMR has recently been used for the elucidation of the
central carbon metabolism of *Corynebacterium glutamicum* (Wiechert, Siefke et al. 1997) and specifically for deciphering the anaplerosis in the same microorganism (Petersen, de Graaf et al. 2000). Mass spectrometry has been used in medicine for estimating gluconeogenic rates (Kelleher 1999) and in bacteria for the analysis of the network of *Penicillium chrysogenum* (Christensen and Nielsen 2000). 2D-NMR has been used in the analysis of the riboflavin-producing *Bacillus subtilis* (Sauer, Hatzimanikatis et al. 1997).

Isotopomeric methods are insufficient in cases where there is no splitting of the carbon backbone in a bioreaction sequence. To facilitate flux determination in this situation, a method employing the use of uniformly radiolabeled $^{14}$C-tracers has been developed in the context of metabolic flux analysis of indene bioconversion in *Rhodococcus* sp. KY1 (discussed later in this chapter). The branched nature of the bioreaction networks associated with indene oxidation required that some fluxes be directly determined. Since there is no change in the carbon backbone of the indene oxidation products in these strains, a procedure was developed to utilize $^{14}$C-labeled intermediates to directly measure the steady state fluxes in the KY1 indene bioconversion network. Fluxes were calculated by tracking the depletion of the corresponding radioactive tracer pulsed into the corresponding pseudo-steady state metabolite pool (Yanagimachi, Stafford et al. 2001).

While fluxes are an important component in identifying pathway flux control, they are not sufficient for the complete description of the control of metabolic pathways. This can be pursued in the framework of metabolic control analysis (MCA) (Delgado and Liao 1992; Fell 1992) provided that metabolite concentrations can be measured in a parallel fashion to complement flux information. Such methods are presently tedious and not scaleable. High-throughput probes of the metabolome will allow the deployment of the mathematical structure of MCA to the complete elucidation of metabolic networks. On the mathematical side, it is important to consider and further develop methods that utilize the concept of reaction grouping (Brown, Hafner et al. 1990; Stephanopoulos, Arisidou et al. 1998) as means of deciphering pathway flux control in a rational and systematic manner. Additionally, fast, high-throughput approaches for flux determination, perhaps utilizing transient experimental data, would tremendously
enhance the magnitude and complexity of pathway systems that are amenable to the above analytical approaches.

Finally, it is noted that microarrays will be a very important tool of metabolic engineering. Gene expression data are related to flux and metabolite measurements, although a satisfactory linkage of the two classes of measurements has yet to be reported. Important as they may be in identifying discriminatory genes and expression patterns, the real value of DNA microarray data will emerge once the expression phenotype is linked to data describing cell physiology such as metabolic fluxes. This linkage will depend strongly on the continued development of tools and technologies to increase measurement capabilities. As mentioned earlier, one clear need is the ability to routinely measure protein levels relevant to a metabolic system under analysis. The technology to do this on a large scale is still some time away.

2.1.4 Future prospects for metabolic engineering

The mind-frame in the first years of metabolic engineering was focused on strain improvement by pathway engineering. In the meantime, however, it has become abundantly clear that many new applications of the basic concepts of metabolic engineering will develop with the advent of enabling technologies such as DNA microarrays, directed evolution, tracers for flux measurement, and advances in applied molecular biology. A continuing expansion of this field will be driven in part by genomics research and technologies that make use of sequence information to generate data on important classes of intracellular molecules. The connection of such developments with metabolic engineering is based on a bi-directional interaction whereby these technologies will help develop tools and reagents needed for pathway manipulation (genes, promoters, etc.), but also because metabolic engineering provides a most convenient framework for the analysis and integration of such data.
2.2 Technologies for biocatalytic stereoselective oxidation: Enzymes and microorganisms

Chiral compounds – molecules with one or more asymmetric carbon atoms - are central to biological chemistry. Molecular geometry is a fundamental determinant of nearly all biochemical phenomena that defines how biomolecules interact with one another, which ultimately confers phenotypes that are of interest to scientists and engineers. The pervasive role of chirality throughout biology is evident, for example, by the nature of proteins that are composed solely of L-amino acids and not their D-enantiomers. Additionally, well-characterized metabolic pathways have evolved in both prokaryotic and eukaryotic organisms to metabolize many epimers of pentose and hexose (among other) carbohydrates. In light of these and many other examples, it is clear that chirality is a crucial aspect of all enzyme-substrate and protein-protein interactions that result in cellular and system phenotypes.

As such, the design and use of small chiral molecules against biological macromolecular targets is of considerable significance in the pharmaceutical industry. Increasingly important among them are chiral compounds whose therapeutic activity is due primarily to a single enantiomer. These compounds accounted for $133 billion in worldwide drug sales in 2000, up 13% from 1999 (Stinson 2001), continuing a trend of the past several years (Stinson 1997; Stinson 1999; Stinson 2000). In addition, sales of enantiomeric intermediates used in pharmaceutical manufacturing are similarly lucrative. Demand for these compounds is expected to continue to increase as more patents of drugs marketed as racemic mixtures expire, enticing pharmaceutical companies to compete for the patenting rights to the corresponding single-enantiomer active ingredient, where applicable. The selectivity of these drugs is a result of the differential binding characteristics particular stereoisomers have with the active site of a target enzyme. Different stereoisomers of a compound can have drastically reduced therapeutic activity or even toxic effects. Consequently, there is much interest in developing technologies, including biocatalysts, for the synthesis of chiral compounds. Toward this objective, numerous soil microorganisms have been identified that have evolved stereoselective enzymes catalyzing oxidative reactions on aromatic
hydrocarbons. Oxygenation reactions, which are particularly prevalent in the aromatic degradative pathways of soil bacteria, are especially difficult to carry out non enzymatically. Consequently, many of these enzymes are of substantial interest to process chemists and others. The next section briefly reviews the key aromatic degradative pathways expressed in the soil bacteria that harbor some of these important stereoselective enzymes.

2.2.1 Structure of catabolic pathways for aromatic compounds

Metabolism of aromatic compounds, in both bacterial and mammalian systems, generally begins with an oxygenation reaction on the aromatic ring that activates the substrate for subsequent ring-cleavage reactions and mineralization to glycolytic and TCA intermediates. The catabolic pathways of naphthalene and toluene have been well characterized as growth substrates in Pseudomonas putida F1 and are described as representative examples of aromatic degradative pathways that have been deduced for numerous other compounds (Williams and Sayers 1994). Both naphthalene and toluene are acted upon by a dioxygenase to form a dihydrodiol, followed by dehydrogenation to form the corresponding diol (Figure 2.1) (Bayly and Barbour 1984; Gibson and Samramanian 1984; Gottschalk 1986; Zylstra and Gibson 1989; Grund, Denecke et al. 1992; Lau, Bergeron et al. 1994). The naphthalene-1,2-diol is subsequently oxidized to salicylate, which is decarboxylated by salicylate 1-hydroxylase to form catechol. Two pathways have been found to mineralize catechol into central carbon metabolism intermediates: In the ortho pathway, catechol is cleaved by a catechol-1,2-dioxygenase and ultimately degraded to acetyl-CoA and succinate, while in the meta pathway, catechol is oxidized in a catechol-2,3-dioxygenase-catalyzed reaction and subsequently metabolized to pyruvate and acetaldehyde.

The most common toluene degradative pathway also is comprised of initial dioxygenase and dehydrogenase reactions to give 3-methylcatechol, which is primarily metabolized via the meta pathway. Not depicted in Figure 2.1 are additional, less common toluene degradative pathways that have been reported. For example, P. putida
mt-2 oxidizes toluene to benzyl alcohol, which is converted to catechol for subsequent meta cleavage; *P. mendocina* is reported to oxidize toluene to p-cresol, which is then converted to protocatechuate prior to ring cleavage. While no evidence exists to exclude the presence of these alternate pathways from *Rhodococcus* sp. I24, the detected products of indene metabolism, as detailed later in this chapter, suggest that the primary pathways of indene oxidation in I24 are initiated by naphthalene dioxygenase and toluene dioxygenase enzymes.

2.2.2 Gene organization in aromatic catabolic pathways

In *P. putida*, the genes encoding the naphthalene degradative enzymes are generally located in two distinct operons spanning 30.5 kb: *nahAaAbAcAdBCDEF*, which gives rise to the enzymes converting naphthalene to salicylate (upper pathway), and *nahGHINLJK*, encoding the catabolism of salicylate to pyruvate and acetaldehyde (lower *meta* pathway). A regulatory gene, *nahR*, is located between the two operons (Yen and Serdar 1988). The *nahAaAbAcAdBCDEF* operon is expressed constitutively at a low-level, and the upper pathway product (salicylate) then induces the expression of both *nah* operons (Yen and Serdar 1988). In some strains, addition of naphthalene or salicylate induces the enzymes of the chromosomal ortho-cleavage pathway (Barnsley 1976; Yen and Serdar 1988).

The gene organization of the toluene degradative pathway has been well characterized in *P. putida* F1. The genes are arranged as *todFC1C2BADEGIH* and span 9.3 kb (Zylstra and Gibson 1989; Lau, Bergeron et al. 1994). The genes encode the enzymes listed in Table 2.1 (in order of action in toluene catabolism):
Figure 2.1  Key reactions in naphthalene and toluene degradative pathways in Pseudomonas. These reactions closely parallel those observed in Rhodococcus indene bioconversion. Rhodococcus sp. I24 can grow on both naphthalene and toluene, while Rhodococcus sp. KY1 grows on naphthalene.
Table 2.1 Toluene degradative enzymes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>TodA</td>
<td>Flavoprotein reductase of toluene dioxygenase (TDO)</td>
</tr>
<tr>
<td>TodB</td>
<td>Ferredoxin (of TDO)</td>
</tr>
<tr>
<td>TodC1C2</td>
<td>large and small subunits of terminal ISP (of TDO)</td>
</tr>
<tr>
<td>TodD</td>
<td>cis-toluene dihydrodiol dehydrogenase</td>
</tr>
<tr>
<td>TodE</td>
<td>3-methylcatechol-2,3-dioxygenase</td>
</tr>
<tr>
<td>TodF</td>
<td>HOHD hydrolase</td>
</tr>
<tr>
<td>TodG</td>
<td>2-hydroxypenta-2,4-dienoate hydratase</td>
</tr>
<tr>
<td>TodH</td>
<td>4-hydroxy-2-oxovalerate aldolase</td>
</tr>
<tr>
<td>TodI</td>
<td>Aldehyde dehydrogenase</td>
</tr>
</tbody>
</table>

The todC1C2BADE genes are co-induced by toluene. Additionally, the todX gene located upstream of todF is proposed to encode an outer membrane protein that plays a role in toluene transport (Wang, Rawlings et al. 1995).

2.2.3 Key biocatalytic enzymes

2.2.3.1 Oxygenases

Oxygenases are useful enzymes for introducing chiral centers to prochiral compounds in a stereospecific manner. As shown above, these enzymes catalyze the initial step in the biodegradation of many aromatic compounds by a number of microorganisms. Oxygenases play a significant role in the metabolism of straight-chain alkyl and aromatic compounds, and also many halogenated hydrocarbons (Finnerty 1992; Warhurst and Fewson 1994). The broad applicability of these biocatalysts has generated strong interest in their function including the mechanisms of their activity and subunit composition of many oxygenases (Harayama, Kok et al. 1992; Butler and Mason 1997). In general, the complex nature of these enzymes, as well as their cofactor requirements, necessitates the development of whole-cell bioconversion systems to prevent enzyme degradation and facilitate cofactor regeneration (Kragl, Kruse et al. 1996; Nidetzky, Neuhauser et al. 1996).
This class of enzymes is predominantly composed of two types: dioxygenases and monooxygenases. In some cases, dioxygenases have been found to catalyze monooxygenase-type oxidation on non-native substrates (Wackett, Kwart et al. 1988; Gibson, Resnick et al. 1995; Torok, Resnick et al. 1995; Lee, Resnick et al. 1997). Dioxygenases are multi-subunit enzymes, one class of which hydroxylate adjacent alkene carbon atoms to give a cis-diol with both hydroxyl groups derived from molecular oxygen. The second major type of dioxygenases, the ring-cleaving class, are beyond the scope of this study and are not discussed here. Two of the most thoroughly characterized dioxygenases are the naphthalene and toluene dioxygenases from *Pseudomonas putida*, which are both three-component enzyme systems. In these enzymes, a reductase with the coenzyme flavin adenine dinucleotide (FAD) accepts two electrons from NADH; these electrons are then transferred to a [2Fe-2S] ferredoxin and a terminal oxygenase, which then mediates the hydroxylation of the aromatic ring (Mason and Cammack 1992; Butler and Mason 1997). The subunit organization of these dioxygenases is depicted in Figure 2.2.

![Diagram](image)

**Figure 2.2** Electron transport and relationship between subunits in naphthalene and toluene dioxygenase. Adapted from (Butler and Mason 1997).

A number of oxygenases that catalyze multiple transformations of indene have been described. Toluene dioxygenase has been found to possess both monooxygenase and dioxygenase activities. Wackett and co-workers induced *Pseudomonas putida* F39/D with toluene, which then converted indene to cis-(15,2R) indandiol in approximately 30% e.e. and (15)-indenol in 26% e.e. (Wackett, Kwart et al. 1988). In *Pseudomonas* sp. 9816-4, indene serves as a substrate for both mono- and dioxygenation reactions by a
naphthalene dioxygenase to form (15)-indenol and cis-(1R,25)-indandiol, respectively (Gibson, Resnick et al. 1995). The same products were detected in *Rhodococcus* sp. NCIMB 12038 when the cells were induced with naphthalene (Allen, Boyd et al. 1997). Both naphthalene dioxygenase and toluene dioxygenase have been shown to catalyze reactions in addition to those observed on their native substrates and indene (Torok, Resnick et al. 1995).

Most known monooxygenases are of mammalian origin and of the cytochrome P-450 class of enzymes. They generally are composed of one or two electron-transport proteins and a terminal hydroxylase, and utilize molecular oxygen to oxidize their substrate with the concomitant reduction of NADPH (Harayama, Kok et al. 1992). In some cases such as naphthalene and anthracene oxidation (van Bladeren, Sayer et al. 1985), monooxygenation of alkene substrates yield epoxide products, which are variably stable depending on the compound due to their highly reactive nature. These enzymes are primarily involved in detoxification reactions of xenobiotic substrates, including most pharmaceutical compounds that tend to have numerous aromatic constituents, as well as in the biosynthesis of steroid hormones (Guengerich 1991).

Although fewer bacterial monooxygenases have been identified, one well-studied enzyme is the methane monooxygenase (MMO) converting methane to methanol in a number of gram-negative methanotrophic bacteria (Murrell, Gilbert et al. 2000). The most commonly expressed enzyme, the membrane-bound particulate MMO (pMMO), is a three-subunit enzyme complex and requires copper as a cofactor as studied in *Methylococcus capsulatus* (Bath) (Zahn and DiSpirito 1996). The *pmoCA* genes (3.2 kb total) have been cloned in the same microorganism (Semrau, Chistoserdov et al. 1995). Of more interest is the soluble MMO (sMMO), which is capable of co-oxidation of numerous aromatic compounds and thus is a candidate for bioconversion processes. sMMO consists of an iron-sulfur, FAD-containing reductase, a small electron-transport coupling protein (protein "B"), and a terminal αβγδ non-heme hydroxylase (Lund, Woodland et al. 1985). The genes encoding each of the subunits have been cloned, also
in *M. capsulatus* (Bath) (Murrell 1994). The active-site biochemistry of sMMO has been studied in significant detail as reviewed recently (Westerheide, Pascaly et al. 2000).

A number of other P-450 bacterial monoxygenases have also been characterized to varying extents (O'Keefe and Harder 1991). Similar to sMMO, the well-studied P450CAM monoxygenase (catalyzing the stereospecific hydroxylation of (1R)-camphor) from *P. putida* is a multi-subunit enzyme containing a ferredoxin, reductase, and terminal oxygenase protein. As opposed to typical dioxygenase systems, where both atoms of molecular oxygen are incorporated into the product, in P-450CAM the excess oxygen atom is incorporated into water (Sliger and Murray 1986). Other P-450 enzymes have been identified in *Rhodococcus* (Cardini and Jurtshuk 1968; Shao and Behki 1996), *Streptomyces* (Shafiee and Hutchinson 1987; Shafiee and Hutchinson 1988), and *Pseudomonas* (Panke, Wubbolts et al. 2000).

### 2.2.3.2 Epoxide hydrolases

Epoxide hydrolases catalyze the addition of water to epoxides to form the corresponding trans-diol. This occurs with an inversion of stereochemical configuration at the carbon atom attacked by the nucleophilic hydroxyl moiety. As will be elaborated upon below, epoxide hydrolysis often occurs non-enzymatically at a significant rate, and can be influenced by pH as investigated for 1,2-tetrahydronaphthalene epoxide and the anti- and syn- tetrahydrodiol epoxides of naphthalene (Lee and Fisher 1997). Most epoxide hydrolases belong to the α,β-hydrolase fold superfamily of enzymes (Ollis, Cheah et al. 1992; Beetham, Grant et al. 1995; Rink, Fennema et al. 1997), which also includes esterases, dehalogenases, and lipases (Beetham, Grant et al. 1995). The only known epoxide hydrolases not belonging to this class are the mammalian enzymes cholesterol epoxide hydrolase (Watabe, Kanai et al. 1981) and leukotriene A4 hydrolase (Funk, Radmark et al. 1987), and the limonene-1,2-epoxide hydrolase from *Rhodococcus erythropolis* DCL14 (discussed further below). Epoxide hydrolases have been identified that play a role in the degradation of numerous xenobiotics in mammalian systems (Besse and Veschambre 1994; Archer 1997; Archelas and Furstoss 1998; Archelas and Furstoss 1999; Weijers and de Bont 1999). As oxirane moieties are highly reactive (i.e.
electrophilic), they are known for their ability to form DNA and protein adducts in vivo, and as a result are highly toxic and mutagenic (Sayer, Yagi et al. 1985; Adams, Yagi et al. 1995). In the following sections, epoxide hydrolase activities of particular relevance to biocatalytic microorganisms are discussed.

2.2.3.2.1 Fungal epoxide hydrolases

Several fungal strains have been isolated that express epoxide hydrolase activity on one or both indan oxide enantiomers (Zhang, Reddy et al. 1995). Two strains, Diplodia gossipina ATCC 16391 and Lasiodiplodia theobromae MF 5215, are able to kinetically resolve racemic indan oxide to the (1S,2R) form at 100% e.e. A preparative scale bioresolution with D. gossipina gave a 14% yield (50% theoretical maximum yield). Zhang et al. also suggested that the bioresolution of racemic indan oxide with D. gossipina is pH-depndant. Indan oxide hydrolysis kinetic data for a narrow pH range (7.5-9.0) in 1.0 M phosphate buffer showed an increase in the hydrolysis rate at pH 7.5 relative to higher pH conditions (data showing the stereochemistry of the indandiols formed from indan oxide hydrolysis were not reported). In other work, Gilmaniella humicola MF 5363 is reported to resolve (1R,2S)-indan oxide to 100% e.e. via catalytic hydrolysis of the (1S,2R) enantiomer. Several other fungal strains were also reported to resolve racemic indan oxide to one form of indan oxide in lower e.e.

A study of the dihydroxylation of geranoil-N-phenylcarbamate to geranoil-N-phenylcarbamate-6,7-diol by Aspergillus niger also showed that the stereochemistry of the final product can be influenced by both epoxide hydrolase activity and the pH of the culture (Zhang, Archelas et al. 1991). In the absence of A. niger at pH 2, geranoil-N-phenylcarbamate-(6S),7-epoxide is rapidly hydrolyzed (non-enzymatically) to the (6S)-diol, while at pH 7 this epoxide is stable (Figure 2.3). Only in the presence of A. niger at pH 7 is the epoxide intermediate enzymatically hydrolyzed to the (6R)-diol. Similar results were subsequently reported for the dihydroxylation of citronellol-N-phenylcarbamate by A. niger (Zhang, Archelas et al. 1991; Zhang, Archelas et al. 1992; Chen, Archelas et al. 1993).
A. niger and Beauveria sulfurescens have also been shown to express epoxide hydrolases that act on styrene oxide and derivatives (Pedragosa-Moreau, Archelas et al. 1996). A. niger was shown to act only on (S)-styrene oxide and a derivative with a methyl group substituted onto the oxirane ring proximal to the phenyl substituent. However, B. sulfurescens was shown to act not only on the two aforementioned compounds with similar results, but also on several other compounds, including indan oxide. When racemic indan oxide was added to B. sulfurescens, (1R,2S)-indan oxide was resolved to 98% e.e. at 20% yield, with trans-(1R,2R)-indandiol obtained at 69% e.e. and 48% yield (Boyd, Sharma et al. 1982).

2.2.3.2.2 Bacterial epoxide hydrolases

2.2.3.2.2.1 General

Although there are a limited number of cloned epoxide hydrolases in bacterial systems, a number of activities have been characterized in Nocardia, Pseudomonas, and Rhodococcus among other bacteria (Mischitz, Kroutil et al. 1995; Faber, Mischitz et al. 1996; Krenn, Osprian et al. 1999). One example of a cloned bacterial epoxide hydrolase is
that from *Agrobacterium radiobacter* AD1, a strain that utilizes epichlorohydrin as a sole carbon source (Rink, Fennema et al. 1997). Sequence homology studies indicate that this enzyme belongs to the α,β-hydrolase fold superfamily.

### 2.2.3.2.2 Limonene-1,2-epoxide hydrolase

Studies of limonene metabolism in *Rhodococcus erythropolis* DCL14 uncovered an enzyme with substantial implications for stereoselective hydrolysis of epoxides, a limonene-1,2-epoxide hydrolase (LEH). Limonene is a key component of citrus peel oils and may be used as a substrate for the synthesis of other flavor compounds. Limonene metabolism in *R. erythropolis* DCL14 proceeds in an analogous fashion to many other aromatic compounds, as limonene is oxidized to a dihydroxylated intermediate (van der Werf, Swarts et al. 1999). In contrast to naphthalene and toluene metabolism, however, the corresponding diol is formed through cleavage of an epoxide ring instead of direct dihydroxylation (Figure 2.4). LEH is a cytoplasmic enzyme involved in the metabolism of limonene by *R. erythropolis* DCL14 that lacks substantial sequence homology to previously characterized epoxide hydrolases belonging to the α,β-hydrolase fold protein family. The enzyme is a 17 kDa monomer (as determined by gel filtration and SDS-PAGE) that does not require a metal ion as cofactor (van der Werf, Overkamp et al. 1998). The *limA* gene encoding LEH was cloned and determined to be approximately 0.5 kb in size (Barbirato, Verdoes et al. 1998).

Limonene-1,2-epoxide hydrolase shows substantial differences in activity against the four stereoisomers of limonene-1,2-epoxide. The reported *in vitro* activity of LEH against (1R,2S,4R)-limonene-1,2-epoxide is 85.1 mmol min⁻¹ (mg protein)⁻¹, with the (1S,2R,4R), (1R,2S,4S), and (1S,2R,4S) stereoisomers having 7.8, 65, and 58% of that activity respectively (van der Werf, Orru et al. 1999). Also of note is that when all four stereoisomers are added simultaneously to a mixture of the enzyme, the two (1R,2S) diastereomers are hydrolyzed preferentially. Among other epoxides that are enzymatically hydrolyzed by LEH are 1-methylcyclohexene oxide (47% relative activity),
cyclohexene oxide (4%), and indan oxide (57%, \( E = 3.6 \)). When LEH was challenged with racemic indan oxide, the \((1S,2R)\) enantiomer was preferentially hydrolyzed.

2.2.4 Key microorganisms

2.2.4.1 *Pseudomonas*

A popular bioconversion host is the gram-negative soil bacterium *Pseudomonas putida* that has been found to catalyze a diverse range of biotransformations. The genome of *P. putida* is currently being sequenced, and many molecular biology tools have been developed for gene transfer and/or knockout in this species. *Pseudomonas* is of particular relevance to indene bioconversion as *P. putida* F1 expresses toluene dioxygenase (TDO) capable of oxidizing indene to the \((-\)-CAI precursor cis-\((1S,2R)\)-indandiol under toluene induction together with a number of side products (Wackett, Kwart et al. 1988). However, this host is unsatisfactory for industrial-scale \((2R)\)-indandiol production because (1) TDO gives an enantiomeric mixture of cis-indandiol, (2) *P. putida* F1 requires toluene for induction that consequently competes with indene for oxidation, and (3) indene oxidation in *P. putida* F1 is feedback inhibited by its oxidation products (Chartrain, Jackey et al. 1998; Buckland, Drew et al. 1999). Another strain, *P. putida* 9816-4, expresses a naphthalene dioxygenase capable of oxidizing indene to cis-\((1R,2S)\)-indandiol and other side products (Gibson, Resnick et al. 1995). Other examples of biocatalytic activities in *Pseudomonas* species include: ethylbenzene oxidation in *P. putida* 39/D (Gibson, Gschwendt et al. 1973), catabolism of benzene, toluene, and \(p\)-xylene mixtures by recombinant *P. putida* mt-2 (Lee, Jung et al. 1995), nitrile hydratase activity in *P. sp. B21C9* (Masutomo, Inoue et al. 1995), and lipase activity in *P. cepasia* (Itoh, Mitsukura et al. 1997). The availability of a full compliment of genetic tools makes *Pseudomonas* promising for many bioconversion applications, though

---

2 The \( E \)-ratio is defined as the ratio of \((k_{cat}/K_m)\) values for the catalytic hydrolysis of the two epoxide enantiomers. Other methods have been developed to measure \( E \) that rely upon measurements of conversion and the enantiomeric excess of the product and substrate for cases where direct determination of the aforementioned kinetic parameters is experimentally prohibitive.
catabolite repression can complicate strain development (Holtel, Marques et al. 1994; Muller, Petruschka et al. 1996; O'Connor, Duetz et al. 1996; Saier 1998).

2.2.4.2 Rhodococcus

*Rhodococcus* is a popular choice for biocatalysis and bioremediation because it has been found to be highly solvent-tolerant and lacking catabolite repression that hinders biotransformations in other systems, including *P. putida*. Bacteria of this genus are members of the nocardioform actinomycetes, closely related to *Corynebacterium*, *Nocardia* and *Mycobacterium* (Lechevalier and Lechevalier 1985). *Rhodococcus* are gram-positive, aerobic, and non-motile. Their morphology includes both rod and coccus forms with extensive filamentation (Warhurst and Fewson 1994). Isolates have been found to catalyze a wide range of transformations, including: hydrolysis of nitriles to carboxamides (Osprian, Jarret et al. 1999), desulfurization of dibenzothiophene sulfone (Wang, Humphrey et al. 1996), metabolism of styrene (Warhurst and Fewson, 1994), biphenyl (Kosono, Maeda et al. 1997), 1,2-butanediol (Nakajima-Kambe, Sawai et al. 1995), pyrene (Bouchez, Blanchet et al. 1997), anal ine (Aoki, Uemori et al. 1988), chlorophenol (Eulberg, Golovleva et al. 1997), o-xylene (DiGennaro, Rescalli et al. 2001), cyclododecanone (Kostichka, Thomas et al. 2001), benzene (Paje and Couperwhite 1996; Paje, Neilan et al. 1997), and conversion of acrylonitrile to acrylamide and acrylic acid (Hughes, Armitage et al. 1998).
Figure 2.4  Degradative pathways for (4R) and (4S)-limonene in *Rhodococcus erythropolis* DCL14. Enzymes are in bold type and stereochemistry is indicated on chemical structures. Adapted from (van der Werf, Swarts et al. 1999).
2.3 Metabolic flux analysis of indene oxidation in *Rhodococcus*

2.3.1 Indene bioconversion: A promising system for improvement by systems analysis and metabolic engineering

An effective whole-cell biocatalyst comprises a bioreaction network optimally configured for maximizing the yield and productivity of the desired product. The development of such a strain can best be performed within a metabolic engineering paradigm. This is based on a rigorous flux analysis that reveals the relative importance of the different metabolic pathways in the strain and suggests specific ways for further improvement. To enable the metabolic engineering of indene bioconversion in *Rhodococcus*, (a) the prerequisites for flux analysis of relatively "uncharacterized" strains, such as those described here, where there is little *a priori* knowledge of the bioconversion pathways of interest, and (b) the tools for controlled and efficient genetic modification, required development.

The foremost requirement is the accurate determination of an observable bioreaction network structure that describes indene oxidation in *Rhodococcus*. Based on product accumulation profiles and induction studies, indene bioconversion networks have been proposed for several isolates as described further below (Chartrain, Jackey et al. 1998). To further validate these networks for our strains and employ them for flux analysis, an experimental system that can maintain cells at steady state while allowing accurate metabolite measurements for flux determination was developed. The system comprised a chemostat with a regular feed of liquid medium and separate supply of the indene precursor through a gas phase line. Indene uptake and metabolite production rates were easily measured in this system, leading to the calculation of the unknown bioconversion fluxes. Additional measurements for system closure and further validation were obtained by using radiolabeled tracers and measuring the products of their oxidation in *Rhodococcus* cultures (Yanagimachi, Stafford et al. 2001).

In parallel to the above efforts, the biological tools required to implement at the genetic level proposed gene deletions or overexpressions were needed. For novel strains, such as the isolates described here, the genetics of bioconversion are relatively
unknown and must be developed to allow implementation of any changes deemed appropriate from flux analysis. The tools needed include plasmids that can replicate in both *Rhodococcus* and *Escherichia coli*, for carrying a genomic library and manipulating cloned genes; selectable markers that must be determined for use in plasmids; and transformation methods to facilitate gene transfer between strains. The development and application of such tools is described in Chapter 4.

2.3.2 Indene bioconversion in *Rhodococcus* sp. I24

2.3.2.1 Isolation of *Rhodococcus* sp. I24 and characterization of indene bioconversion

The quest for microorganisms capable of performing the desired biotransformation of indene led to the isolation of several strains of the genus *Rhodococcus* from soil samples contaminated with aromatic compounds that are able to oxidize indene to 1,2-indandiols of different chirality, and various other oxygenated derivatives (Chartrain, Jackey et al. 1998). Induction studies indicated that several oxygenases were present and differentially induced by naphthalene, toluene, and indene. The stereospecific nature of the enzymes expressed in *Rhodococcus* as well as their ability to tolerate indene as a substrate makes these microorganisms promising candidates for development as an industrial-scale biocatalyst for the production of (2R)-indandiol.

*Rhodococcus* sp. I24 was isolated from a toluene-contaminated aquifer and was found to oxidize indene to 1,2-indandiol and several other products, and to grow on both naphthalene and toluene. The undesired products 1-indenol and 1-indanone were formed directly from indene while racemic 1-keto-2-hydroxy-indan was formed from the indandiols. Based on product formation profiles and induction experiments, I24 was hypothesized to contain a system of oxygenase enzymes that convert indene to various enantiomers of indandiol through the proposed bioreaction network shown in Figure 2.5 (Chartrain, Jackey et al. 1998). The oxidation of indene to the indandiols followed by dehydrogenation is consistent with the degradation pathways elucidated for similarly
structured compounds naphthalene and toluene in *Pseudomonas* (Figure 2.1) (Gibson and Samramanian 1984). However, the catechol analog 1-keto-2-hydroxy-indan is not oxidized via a ring-cleaving dioxygenase as has been determined in other aromatic degradative pathways. With indene as the sole aromatic compound present, I24 produced primarily *trans-(1R,2R)-indandiol* (>98% e.e.) in shake-flask cultures and withstood significantly higher concentrations of indene than *P. putida* in a two-liquid phase cultivation system that utilized silicon oil as an indene carrier (Chartrain, Jackey et al. 1998). Based on these findings, I24 emerged as a promising strain for subsequent development using a metabolic engineering approach for the biocatalytic production of (2R)-indandiol.

![Indene bioconversion network](image)

*Figure 2.5* Indene bioconversion network for *Rhodococcus* sp. I24 proposed by (Chartrain, Jackey et al. 1998)
2.3.2.2 Substrate and product toxicity

The hydrophobic nature of indene presents at least two substantial hurdles in developing an efficient bioconversion process for the production of (2R)-indandiol. Studies in other microorganisms capable of metabolizing aromatic compounds suggest that the substrate and/or products may be toxic to *Rhodococcus* and attenuate indene bioconversion activity. Additionally, because indene is only sparingly soluble in aqueous media (~100 mg/L under standard fermentation conditions), the availability of indene to the key oxidation enzymes must be considered.

The hydrophobic (e.g. lipophilic) behavior of organic compounds is conventionally described by the partitioning of the organic compound between aqueous and non-aqueous phases, denoted by the parameter log $P_{OW}$, where $P_{OW}$ is the partition coefficient of the compound between 1-octanol and water under equilibrium conditions (Leo, Hansch et al. 1971; Smith, Hansch et al. 1975). Data on the partitioning of indene between aqueous and organic phases have not been reported previously. For cases where actual log $P_{OW}$ have not been experimentally determined, computational approaches have been developed to estimate this parameter. One well-established method is by the use of hydrophobic fragmental constants based on Eq. (2.1), where $f$ is the hydrophobic fragmental constant, $a$ is the frequency of the given fragment (i.e. functional group in a molecule), $C_M$ is a correction factor and $k$ is the corresponding frequency (Nys and Rekker 1973; Rekker and Mannhold 1992; Mannhold, Rekker et al. 1998):

$$\log P_{OW} = \sum_{i=1}^{n} f_i + \sum_{i=1}^{m} k_i C_M$$  \hspace{1cm} (2.1)

Table 2.2 shows the experimentally determined log $P_{OW}$ and the corresponding empirical values for indene and a series of structurally related compounds using the approach based on Eq. (2.1). The calculated log $P_{OW}$ value for indene, 3.07, suggests that indene will partition into non-aqueous environments where present. Studies in liposomes, prepared from *Escherichia coli* phospholipids, showed that the log $P_{OW}$ value described above correlates well with the partition coefficients measured for ten cyclic hydrocarbons, including toluene, naphthalene, phenanthrene, and benzoic acid
(Sikkema, de Bont et al. 1994). In the context of a conventional, single-phase bacterial bioconversion system such as that described below for indene oxidation in *Rhodococcus*, the lipid-rich cytoplasmic membranes are likely to harbor significant concentrations of indene. As the cytoplasmic membrane plays a major role in cellular energy transduction, signal transduction, and the regulation of solute transport, among many other functions, recognizing the potentially toxic nature of indene is key to designing a robust indene bioconversion process.

The nature of hydrocarbon toxicity has been studied extensively (Sikkema, de Bont et al. 1995) and has been shown to affect growth rates and bioconversion titers in a variety of systems via direct incorporation of the hydrocarbon into the cell (Gibson, Cardini et al. 1970), although this phenomena can be minimized by feeding the substrate through the gas phase (Gibson, Gschwendt et al. 1973). Evidence of a reduction in respiratory activity upon addition of toluene to *P. putida* (Jenkins, Stephens et al. 1987) and actual physical disruption of the cytoplasmic membrane in toluene-challenged *E. coli* (Woldringh 1973), among many other studies, support the notion that the cytoplasmic membrane is the key cellular structure directly affected by the presence of aromatic hydrocarbons. Numerous microorganisms have been identified that exhibit tolerance to relatively high solvent levels, and appear to have modified fatty acid compositions in the cell membranes (de Smet, Kingma et al. 1978; Sikkema, Poolman et al. 1992; Weber, Ooijkaas et al. 1993; Heipieper and de Bont 1994; Chen, Nijenhuis et al. 1995; Weber and de Bont 1996; Pinkart and White 1997; Ramos, Duque et al. 1997) and/or modified cell wall structure (Park, Chang et al. 1988; Jarlier and Nikaido 1994), as well as the ability in some cases to actively transport excess solvent out of the cell (Isken and de Bont 1996; Paulsen, Brown et al. 1996; Aono, Tsukagoshi et al. 1998; Kieboom, Dennis et al. 1998; Kieboom, Dennis et al. 1998; Paulsen, Chen et al. 2001). With regard to indene bioconversion, studies have recently been reported on the use of multi-channel flow cytometry to evaluate the effect of indene and its oxidation products on the physical integrity of the cell membrane and its ability to maintain a potential for respiratory activity (Leyval, Debay et al. 1997; Lopez-Amoros, Castel et al. 1997; Hewitt, Boon et al.
To minimize the toxic effects of aromatic hydrocarbon substrates and to improve product titers, two-phase (aqueous-organic) bioconversion systems have been used for many years (Lilly 1982). Examples include the oxidation of \( p \)-xylene (Cruden, Wolfram et al. 1992) and degradation of phenol in \( P. \ putida \) (Collins and Daugulis 1997), production of \((R)\)-phenylacetyl carbinol in yeast using hexane as an organic phase (Nikolova and Ward 1992), and the production of octanoic acid and styrene oxide in recombinant \( E. \ coli \) bioconversions, where octane and dodecane were the organic phases used in the respective monooxygenase reactions (Wubbolts, Favre-Bulle et al. 1996).

2.3.3 Isolation of \textit{Rhodococcus} sp. KY1 from chemostat cultures

To facilitate the quantitative analysis of the indene bioconversion network in \textit{Rhodococcus} sp. I24, an experimental system enabling the accurate measurement of indene metabolites was developed. A multi-phase fermentation system commonly employed when dealing with substrates or products of relatively low solubility was not desirable for this analysis due to the aforementioned uncertainties associated with the partitioning of the indene metabolites between the aqueous and organic phases. In addition to the difficulty in obtaining a representative liquid-phase sample to measure indene metabolite concentrations, the lack of experimental partition coefficient data for indene metabolites made a single-phase system essential. To circumvent this issue, a continuous flow system that utilized a gas-phase delivery of indene was implemented. The gas-phase concentration of indene was monitored using a photoionization detector and was manually controlled by mixing with a second air stream prior to sparging through the culture. By measuring the indene air concentration in the feed and exit gas streams of the chemostat, the indene uptake rate was calculated (Yanagimachi 2000). In combination with the measurement of the liquid phase concentration of indene and other indene metabolites in the chemostat (Chartrain, Jackey et al. 1998), the indene metabolite balances were closed (Table 2.3). Independent confirmation of the indene uptake rates calculated using the gas-phase indene concentrations was provided using \([^{14}\text{C}]\)-indene uptake experiments, as will be described below.
**Table 2.2**  log $P_{OW}$ values for compounds structurally similar to indene. Adapted from (Mannhold, Rekker et al. 1998).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Experimental log $P_{OW}$</th>
<th>log $P_{OW}$ based on Eq. (2.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene</td>
<td>2.08 (n = 10)</td>
<td>2.11</td>
</tr>
<tr>
<td>phenol</td>
<td>1.49 (n = 12)</td>
<td>1.55</td>
</tr>
<tr>
<td>toluene</td>
<td>2.59 (n = 7)</td>
<td>2.63</td>
</tr>
<tr>
<td>biphenyl</td>
<td>3.91 (n = 6)</td>
<td>4.02</td>
</tr>
<tr>
<td>naphthalene</td>
<td>3.36 (n = 8)</td>
<td>3.40</td>
</tr>
<tr>
<td>indene</td>
<td>Not determined</td>
<td>3.07</td>
</tr>
</tbody>
</table>
Table 2.3  Steady state concentrations in *Rhodococcus* sp. KY1 chemostat cultures. Adapted from (Yanagimachi 2000).

<table>
<thead>
<tr>
<th>Steady State Values</th>
<th>D=0.10 h⁻¹</th>
<th>D=0.065 h⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 ppm</td>
<td>170 ppm²</td>
</tr>
<tr>
<td><em>trans</em>-(1R,2R)-indandiol (mg/L)</td>
<td>86</td>
<td>181</td>
</tr>
<tr>
<td><em>cis</em>-(1R,2S)-indandiol (mg/L)</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td><em>cis</em>-(1S,2R)-indandiol (mg/L)</td>
<td>24</td>
<td>52</td>
</tr>
<tr>
<td>1-keto-2-hydroxy-indan (mg/L)</td>
<td>25</td>
<td>93</td>
</tr>
<tr>
<td>indan oxide (mg/L)</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>indene (mg/L)</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Biomass (g DCW/L)</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Indene uptake rate (material balance)²</td>
<td>35 ± 5</td>
<td>71 ± 5</td>
</tr>
<tr>
<td>Indene uptake rate (air measurement)²</td>
<td>40 ± 7</td>
<td>62 ± 12</td>
</tr>
</tbody>
</table>

²Determined from sum of indene metabolite excretion rates
³Determined from inlet and outlet gas-phase indene concentrations
⁴Uptake rates are in µmol/g DCW/h
⁵Data for a pseudo-steady state when the concentrations were constant for one residence time
Using this novel fermentation system, I24 was grown in a steady state chemostat culture with an indene feed concentration of 85 ppm in 1.0 vvm of air, and a dilution rate of 0.10 h⁻¹. In preliminary experiments with a continuous system, cell washout of I24 occurred when the indene concentration in the air feed exceeded approximately 200 ppm for dilution rates ranging from 0.05-0.10 h⁻¹. Thus, the indene feed utilized in the experiment described here is well under the toxicity limit of indene to I24. Upon reaching a steady state for five residence times, the primary indene metabolites detected were cis-indandiol, 1-keto-2-hydroxy-indan, and the undesirable byproducts 1-indenol and 1-indanone (Figure 2.6). The lack of trans-indandiol formed may be due to a relatively high indene affinity of the dioxygenases relative to the monooxygenase under these culture conditions.

When the indene feed concentration was increased from 85 to 120 ppm with all other parameters held constant, a significant change in indene metabolism was observed after approximately ten residence times. Formation of 1-indenol and 1-indanone ceased, and the primary oxidation products were trans-indandiol, cis-indandiol, indan oxide and 1-keto-2-hydroxy-indan. The yield of (2R)-indandiol from indene increased from approximately 30% to 60% following the metabolism shift (Table 2.3). The mutant with the altered metabolism from I24, denoted as strain KY1, was isolated from the chemostat and has shown indene metabolite profiles in steady state and batch fermentations consistent with those observed following the metabolism shift in the I24 chemostat culture (Yanagimachi 2000). Additionally, KY1 has been stable in numerous fed-batch experiments. It is believed that the KY1 strain evolved in response to the selective pressure applied by the chemostat environment to the I24 cells. The possibly toxic nature of 1-indenol and 1-indanone, especially at the high concentrations observed in the chemostat, facilitated the emergence of the new KY1 strain that is unable to oxidize indene to 1-indenol and 1-indanone. In steady state chemostat studies performed with KY1, a substantially higher biomass concentration was obtained at a dilution rate of 0.065 h⁻¹ than at 0.10 h⁻¹ at an indene feed of 100 ppm, but the biomass concentrations were similar between the same two dilution rates at 170 ppm indene (Table 2.3). This may be a result of indene toxicity, the effects of which are presumably exerted more
strongly at higher feed concentrations. The higher metabolite concentrations observed for the \((0.065 \text{ h}^{-1}, 100 \text{ ppm})\) state relative to the \((0.065 \text{ h}^{-1}, 170 \text{ ppm})\) state suggests a possible correlation between biocatalyst concentration and indene metabolite titers. These data imply that an optimal fed-batch indene biotransformation be performed at relatively low indene feed to prevent growth attenuation due to substrate toxicity.

![Indene metabolite profiles](image)

**Figure 2.6** Indene metabolite profiles in the *Rhodococcus* sp. I24 chemostat at \(0.10 \text{ h}^{-1}\) dilution rate. Behavior characteristic of strain KY1 is exhibited after 250 hours.
The emergence of the KY1 strain validates the use of a chemostat as a selective tool for strain improvement. This highlights the rather unconventional use of the chemostat for the evolution of strains with improved bioconversion properties, which contrasts with its traditional application to cell physiology studies. In this unique case, the bioconversion substrate (indene) is decoupled from the growth substrate (glucose), and cell growth is presumably limited by the toxic nature of indene and not by the absence of a critical substrate. This unusual feature of indene bioconversion allowed the biocatalyst to be maintained under well-defined environmental conditions with respect to both substrates for the purpose of strain selection, and simultaneously study the effect of the bioconversion substrate on cell physiology through the quantification of network fluxes. In this sense, the chemostat allowed (a) definition of a steady environment and application of selective pressure to I24, which resulted in the isolation of KY1, and (b) systematic study of the different steps of the bioconversion process. The details of this second application of the chemostat to indene bioconversion are provided below.

2.3.4 Characterization of *Rhodococcus* sp. KY1 using radiolabeled tracers

Induction studies that utilized [14C]-indene as a probe were used to more rigorously characterize the indene bioconversion network of I24 and elucidate the difference(s) between the KY1 and I24 strains. Cells were again grown in chemostat cultures in which naphthalene (40-70 ppm), toluene (100-200 ppm), or indene (100-110 ppm) was fed through the gas-phase until a steady state was reached (Yanagimachi 2000). The introduction of these compounds induced the activity of different oxygenases in the KY1 and I24 networks. Cells were removed from the chemostat culture and their physiology probed with [14C]-indene. Specifically, by following the kinetics of formation of the primary oxygenated derivatives of [14C]-indene following the introduction of the [14C]-indene probe, the induction characteristics of key enzymes became apparent. Because of the rapid uptake of [14C]-indene by the cells, the rate of tracer depletion was reaction-limited and provided a measure of the *in vivo* activity of these enzymes. In cases where multiple enzymes were induced, oxygenase activity was estimated using the rate of formation of the appropriate [14C]-indandiol.
Table 2.4 depicts the relative concentrations of $[^{14}\text{C}]$-labeled indene metabolites obtained after adding 25 μM $[^{14}\text{C}]$-indene to I24 and KY1 cells under different inducers. These studies demonstrated that I24 expresses a toluene-inducible dioxygenase activity that produces primarily cis-(1S,2R)-indandiol and 1-indenol, and a naphthalene-inducible dioxygenase that produces primarily cis-(1R,2S)-indandiol and 1-indenol from indene. The tracer data also revealed that KY1 lacks the toluene-inducible dioxygenase present in I24 by virtue of the inability of KY1 to oxidize indene under toluene induction. The naphthalene-induced behavior of I24 and KY1 was similar. The slightly decreased excess of the cis-(1S,2R)-indandiol enantiomer produced by I24 relative to KY1 can be attributed to possible cross-induction of the toluene-inducible dioxygenase in I24. Furthermore, tracer studies under indene induction showed that in KY1, the primary route of indene oxidation is through a novel monooxygenase activity to indan oxide presumed to be of (1S,2R) stereochemistry, while the metabolism in I24 closely resembled that observed under toluene induction.

Additional experiments were performed to further confirm the indene bioreaction network structures. Addition of indan oxide to both in vivo and cell-free systems showed that this intermediate is non-enzymatically hydrolyzed to both trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol in a 4:3 ratio, and that induction by indene had no effect on the hydrolysis rate (Yanagimachi 2000). A corollary of this result is that the trans-(1S,2S) and cis-(1R,2S)-indandiol enantiomers are not formed by hydrolysis of (1S,2R)-indan oxide. This discounted the possibility that either (2S)-indandiol enantiomer is formed (from epoxide hydrolysis) but not detected due to rapid degradation to 1-keto-2-hydroxy-indan. Also, incubation of $[^{14}\text{C}]$-labeled cis-indandiols with induced I24 and KY1 cells resulted in only 1-keto-2-hydroxy-indan being formed, while $[^{14}\text{C}]$-trans-(1R,2R)-indandiol degradation was not detected in either strain (Yanagimachi 2000). These data indicated that (a) there are no isomerization reactions occurring between the three indandiol enantiomers formed by indene oxidation in I24 and KY1, and (b) the dehydrogenase activity previously proposed to degrade trans-indandiol to 1-keto-2-hydroxy-indan is not present at a significant rate. The latter conclusion is consistent with the observation by Chartrain et al. that trans-(1R,2R)-
indandiol was dehydrogenated in I24 only at long culture times. In the context of a subsequent quantitative flux analysis of indene bioconversion, the flux supported by a trans-(1R,2R)-indandiol dehydrogenase was negligible relative to the flux through the other network reactions. Based on findings from these tracer studies, a new bioreaction network was proposed for the KY1 strain (Figure 2.7). The increased yield of (2R)-indandiol characteristic of KY1 made this the most interesting microorganism for further study using metabolic flux analysis.

Table 2.4  Conversion of 25 μM [14C]-indene to primary oxygenated products under different inducers after 5 minutes (reported as percentage of tracer added)

<table>
<thead>
<tr>
<th>[14C]-Metabolite</th>
<th>KY1</th>
<th>I24</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Toluene</td>
<td>Naphthalene</td>
</tr>
<tr>
<td>cis-(1R,2S)-indandiol</td>
<td>0</td>
<td>63</td>
</tr>
<tr>
<td>cis-(1S,2R)-indandiol</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1-indenol</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Indan oxide</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Indene (unoxidized)</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 2.7 Indene bioconversion network in *Rhodococcus* sp. KY1. Metabolite excretion and uptake rates are denoted by $r_i$, while intracellular fluxes are written as $J_i$. 
2.3.5 Flux analysis of indene bioconversion in *Rhodococcus* sp. KY1

The indene bioconversion network proposed for *Rhodococcus* sp. KY1 (Figure 2.7) using the induction studies with radiolabeled indene can be used to write five independent mass balances to describe six intracellular fluxes (Table 2.5). This yields an underdetermined system for the fluxes of the KY1 network requiring that at least one flux be directly measured to uniquely calculate the remaining network fluxes. It is further desirable to directly measure additional fluxes to generate redundancies that can be used to confirm the structure of the proposed bioreaction network, validate the flux estimates and help detect gross measurement errors, if present.

[^14]C-cis-(1S,2R)-indandiol was used to directly measure the corresponding steady state dehydrogenase flux, \( J_{RDH} \), in the KY1 indene bioconversion network (Yanagimachi, Stafford et al. 2001). By measuring the formation of [^14]C-1-keto-2-hydroxy-indan associated with the concomitant depletion of [^14]C-cis-(1S,2R)-indandiol in steady state cells, the cis-(1S,2R)-indandiol dehydrogenase flux was calculated using Eq. (2.2), where \( C2R^* \) is the radiolabeled cis-(1S,2R)-indandiol [C2R].

\[
\ln \left( \frac{[C2R^*(t)]}{[C2R^*(0)]} \right) = -\left( \frac{J_{RDH}}{C2R} \right) t
\]  

(2.2)

The direct determination of this additional flux (\( J_{RDH} \)) allowed the calculation of the remaining fluxes in the indene bioconversion network using the metabolite excretion rates (\( r_i \)) calculated from steady state metabolite concentrations in the chemostat. Figure 2.8 depicts the flux distribution through the KY1 indene bioconversion network for a representative steady state case.

The flux distribution results obtained using the directly determined \( J_{RDH} \) along with the metabolite production rates were validated using two redundant measurements. One consistency check was provided by comparing the predicted value of the indan oxide chemical hydrolysis ratio (\( J_{I} / C2R \)) with the value measured directly using the transient depletion of [^14]C-(1S,2R)-indan oxide tracer and an analogous expression to Eq. (2.2). No significant differences were found between the \( J_{I} / C2R \) ratios.
measured from the tracer experiment and that calculated as described earlier, in experiments where the tracer was added to both steady state cultures and supernatant and cell lysates of KY1 steady state cultures. A second redundancy was provided by comparing the indene uptake rate calculated from the sum of the indene metabolite excretion rates with the indene uptake rate independently determined from the indene gas-phase concentrations in the chemostat. Both redundancy checks confirmed the flux estimates obtained from the metabolite balances and the direct measurement of $J_{RDH}$. Thus, any undetected perturbation of the steady state generated by the assaying procedure (i.e. alteration of NADH/NAD$^+$ ratios) was not significant enough to alter the measured flux distribution.

Table 2.5  Steady state metabolite balance equations for the KY1 indene bioconversion network

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Mass Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indene</td>
<td>$r_{IND} - J_{MO} - J_{DO} = 0$</td>
</tr>
<tr>
<td>Indan oxide</td>
<td>$J_{MO} - J_T - J_{C2R} - r_{IO} = 0$</td>
</tr>
<tr>
<td>trans-(1R,2R)-indandiol</td>
<td>$J_T - r_T = 0$</td>
</tr>
<tr>
<td>cis-(1S,2R)-indandiol</td>
<td>$J_{C2R} - J_{RDH} - r_{C2R} = 0$</td>
</tr>
<tr>
<td>cis-(1R,2S)-indandiol</td>
<td>$J_{DO} - J_{SDH} - r_{C2S} = 0$</td>
</tr>
<tr>
<td>1-keto-2-hydroxy-indan</td>
<td>$J_{RDH} + J_{SDH} - r_K = 0$</td>
</tr>
</tbody>
</table>
Figure 2.8  Intracellular flux distribution for *Rhodococcus* sp. KY1 at 100 ppm indene air feed concentration and 0.065 h⁻¹ dilution rate. The fluxes were normalized by the indene uptake rate.
A final test of the intracellular fluxes determined by metabolite balancing was provided through comparison with the predictions of a first order kinetic model describing the oxidation of pulsed $[^{14}\text{C}]-\text{indene}$ to all detectable indene derivatives in steady state cells (Yanagimachi, Stafford et al. 2001). Assuming Michaelis-Menten kinetics for a typical reaction depicted in Figure 2.7, the rate of labeled metabolite conversion by that reaction can be expressed as:

$$\frac{d[M^*]}{dt} = \left[ \frac{v^i_{\text{max}} [M_{\text{tot}}]}{K^i_m + [M_{\text{tot}}]} \right] [M^*] = -\frac{J_i}{[M_{\text{tot}}]} [M^*]$$

(2.3)

If the concentration of $M_{\text{tot}}$ remains constant in the course of the labeling experiment, the above expression is reduced to first order kinetics with respect to the labeled metabolite concentration described by Eq. (2.4) below:

$$\frac{d[M^*]}{dt} = -k_i [M^*]$$

(2.4)

where

$$k_i = \frac{J_i}{[M_{\text{tot}}]} X$$

(2.5)

In all of the radiolabeled tracer experiments conducted, the concentrations of the corresponding indene metabolites were found to be constant so that the linear model with respect to the radiolabeled tracer is justified. However, for the mass balance on radiolabeled indene, the total metabolite concentration is not constant and the first order kinetic model is only satisfied when the total concentration is sufficiently low such that $[M_{\text{tot}}] \ll K_m$ for that respective enzyme. Here, the flux is also not constant and can be expressed as shown in Eq. (2.6):

$$J_i = \frac{v^i_{\text{max}}}{K^i_m} [M_{\text{tot}}]$$

(2.6)

Substituting Eqs. (2.5) and (2.6) into the $[^{14}\text{C}]-\text{indene}$ mass balance, the dynamics of $[^{14}\text{C}]-\text{indene}$ depletion by all active oxygenases can be described using Eq. (2.7):
\[
\frac{d[I^*]}{dt} = \left[ \sum_i \left( \frac{v_{i \text{max}}}{K_m} \right) X \right] [I^*] \tag{2.7}
\]

Thus, the dynamics of $[^{14}\text{C}]$-indene oxidation to downstream metabolites can be predicted using the flux estimates derived previously from metabolite balancing and direct flux measurement by translating these values into $k_i$ estimates using Eqs. (2.5) and (2.6). These reaction rate constants can be used in the following equations that describe indene oxidation by KY1.

\[
\frac{d[I^*]}{dt} = -(k_{MO} + k_{DO})[I^*] \tag{2.8}
\]

\[
\frac{d[IO^*]}{dt} = k_{MO} [I^*] - (k_{C2R} + k_T)[IO^*] \tag{2.9}
\]

\[
\frac{d[T^*]}{dt} = k_T [IO^*] - k_{TDH}[T^*] \tag{2.10}
\]

\[
\frac{d[C2R^*]}{dt} = k_{C2R} [IO^*] - k_{RDH}[C2R^*] \tag{2.11}
\]

\[
\frac{d[C2S^*]}{dt} = k_{DO} [I^*] - k_{SDH}[C2S^*] \tag{2.12}
\]

\[
\frac{d[K^*]}{dt} = k_{RDH}[C2R^*] + k_{SDH}[C2S^*] + k_{TDH}[T^*] \tag{2.13}
\]

Figure 2.9 compares the experimentally measured metabolite profiles resulting from the oxidation of a pulse of $[^{14}\text{C}]$-indene by steady state chemostat cells with the kinetic profiles predicted by Eqs. (2.8-2.13) using flux values independently determined for the same steady state. The excellent agreement between the actual tracer data and the predicted oxidation profiles provides an additional validation of the fluxes calculated for the KY1 network (Yanagimachi, Stafford et al. 2001).

Flux analysis of several steady states at different dilution rates and indene feed concentrations uniformly demonstrated that the key route of indene oxidation in \textit{Rhodococcus} sp. KY1 is through the novel monooxygenase enzyme. For all steady states analyzed, at least 94% of the indene was oxidized to indan oxide. This analysis also demonstrated that KY1 lacks a \textit{trans}-(1R,2R)-indandiol dehydrogenase previously
hypothesized to be present in the parent I24 strain. Additionally, the use of tracers showed a previously unidentified chemical step in the bioconversion network, namely the hydrolysis of indan oxide to cis-(1S,2R)-indandiol in addition to trans-(1R,2R)-indandiol.

![Graph showing the comparison of kinetic model predictions with experimental measurements of [14C]-indene metabolites for Rhodococcus sp. KY1 cells obtained from a chemostat at steady state obtained with a dilution rate of 0.065 h⁻¹ and 100 ppm indene air feed concentration. Reaction rate constants used in the kinetic model were determined from flux estimates as described in the text. I(t), indene; IO(t), indan oxide; T(t), trans-indandiol; C(t), cis-indandiol; K(t), 1-keto-2-hydroxy-indan.]

Figure 2.9 Comparison of kinetic model predictions with experimental measurements of [14C]-indene metabolites for Rhodococcus sp. KY1 cells obtained from a chemostat at steady state obtained with a dilution rate of 0.065 h⁻¹ and 100 ppm indene air feed concentration. Reaction rate constants used in the kinetic model were determined from flux estimates as described in the text. I(t), indene; IO(t), indan oxide; T(t), trans-indandiol; C(t), cis-indandiol; K(t), 1-keto-2-hydroxy-indan.
2.3.6 Summary

A central finding of this analysis was that indene monoxygenase is the key enzyme for indene oxidation, and the most likely candidate for overexpression if further increase of the total oxidation flux of the indene network is desired. The emergence of indene monoxygenase as the main oxidizing enzyme in KY1 is contrary to the initial hypothesis that implicated toluene-induced dioxygenase as the main route for (2R)-indandiol biosynthesis. Estimates of monoxygenase activity in KY1 suggest that it is probably satisfactory for industrial-scale production of (2R)-indandiol in fed-batch mode (Yanagimachi 2000).

As discussed in Chapter 4, this revised view of indene bioconversion emphasizes the need to express enzymes catalyzing the selective hydrolysis of indan oxide to trans-(1R,2R)-indandiol to prevent degradation by dehydrogenase(s). In terms of genetic modification, this task is more palatable than the original focus on multiple enzyme knockouts. Such secondary targets to improve (2R)-indandiol yield that were also identified by this analysis include the knockouts of multiple dehydrogenase activities and the dioxygenase producing cis-(1R,2S)-indandiol.

The work described in Chapter 2 encompasses the first two metabolic engineering steps (defined in Chapter 1) required to design a robust (2R)-indandiol biocatalyst using a metabolic engineering approach. An experimental system (chemostat) was developed and used to apply selective pressure to the Rhodococcus sp. I24 strain that led to the isolation of the improved KY1 strain. Furthermore, the chemostat was used together with a novel application of $^{14}$C-tracers to first identify the non-enzymatic nature of indan oxide hydrolysis and further define the indene bioconversion network, and also to determine the flux distributions in KY1 that identified the primary routes of indene bioconversion.
CHAPTER 3 - MATERIALS AND METHODS

3.1 Cultivation methods

3.1.1 Microorganisms and reagents

_Rhodococcus_ sp. I24 was isolated from a toluene-contaminated aquifer and can utilize naphthalene or toluene as growth substrates (Chartrain, Jackey et al. 1998; Buckland, Drew et al. 1999). _Rhodococcus_ sp. KY1 was isolated from an indene-induced chemostat culture with the I24 strain and lacks the ability to grow on toluene (Yanagimachi 2000). _Escherichia coli_ JM109 was purchased from New England Biolabs, Inc. (Beverly, MA) (Table 3.1). All chemicals were obtained from Aldrich Chemical Company (Milwaukee, WI) or Sigma (St. Louis, MO) unless indicated otherwise.

3.1.2 Media

_Rhodococcus_ sp. I24, KY1 and transformants were grown in either Luria-Broth (LB) medium or in a defined medium. The defined medium was optimized for maximum cell growth and consisted of 40 g/L glucose, 1.4 g/L (NH₄)₂SO₄, 1.0 g/L MgSO₄·7H₂O, 0.015 g/L CaCl₂·2H₂O, 1.0 g/L MOPS buffer, 1.0 mL/L A9 trace elements solution, 1.0 mL/L Stock A solution, 35.2 mL/L 1.0 M phosphate buffer. The A9 trace elements solution contained, per liter of deionized water: 0.5 g FeSO₄·7H₂O, 0.4 g ZnSO₄·7H₂O, 0.02 g MnSO₄·H₂O, 0.015 g H₃BO₃, 0.01 g NiCl₂·6H₂O, 0.25 g EDTA, 0.05 g CoCl₂·6H₂O, and 0.005 g CuCl₂·2H₂O. The Stock A solution consisted of 2.0 g/L NaMoO₄·2H₂O and 5.0 g/L FeNa·EDTA in deionized water. 1.0 M phosphate buffer contained 113 g/L K₂HPO₄ and 47 g/L KH₂PO₄ in deionized water. This medium was prepared by first adding the appropriate amounts of (NH₄)₂SO₄, MgSO₄·7H₂O, CaCl₂·2H₂O, and MOPS buffer to 882 mL of deionized water and autoclaving this solution for 35 min at 121°C. A9 trace element solution, Stock A, 1.0 M phosphate buffer, and 80 mL of a separately prepared sterile 500 g/L dextrose solution were then added. Except for shake flask cultures, 1.0 mL/L polypropylene glycol (MW 2000) was also added prior to autoclaving.
3.1.3 Culture maintenance

Strains were maintained by periodically plating onto LB-agar plates with appropriate antibiotics, where applicable. Frozen stocks were prepared by adding 500 μL of culture grown in LB with 500 μL of 50:50 LB-glycerol in sterile 2.0 cryogenic vials (Corning Costar Co., Cambridge, MA). Stocks were stored at -80°C until use.

3.2 Indene bioconversions

3.2.1 Shake-flask indene bioconversions

*Rhodococcus* sp. KY1, KY1(pDS3), and KY1(pAL282) were grown in the optimized medium described above (with appropriate antibiotic) in 250 mL screw cap shake flasks (25 mL working volume) to minimize evaporation of indene, or in 1.0 L baffled shake flasks (100 mL working volume). Shake flasks were inoculated with 1.0 mL of preculture grown in 10 mL LB with 10 μg/mL gentimcin sulfate (LB-Gent10) where appropriate. After culture reached an OD600 of approximately 10 (measured using a Uvikon 930 spectrophotometer (Research Instruments International, San Diego, CA)), 1 μL indene per 5 mL working volume was added to the shake flasks approximately every 24 h. All cultures of *E. coli* JM109(pDS3) and JM109(pAL282) were grown in LB-Gent10 similarly to KY1 and recombinants.

3.2.2 Single phase indene bioconversions

A New Brunswick 2.0 L capacity vessel (New Brunswick Scientific Co., Edison, NJ) was used with a VirTis Omni-Culture base (VirTis Company, Gardiner, NY) (Figure 3.1). The fermentation temperature was controlled at 30 ± 1°C and agitation speed at 1000 rpm. The pH was maintained at 6.9 ± 0.2 by addition of filtered 2.0 M NaOH. Filter-sterilized air feed was maintained at 1.0 vvm with Cole-Parmer gas mass flow controllers (Model no. 33116-20, 0-5 std L/min, Cole-Parmer, Vernon Hills, IL). The fermentor was autoclaved at 121°C for 35 min with the salts and antifoam in deionized water. The autoclaved stock, trace elements, and glucose solutions were added
aseptically after the fermentor cooled. The fermentor was inoculated with 1-10 mL of
inoculum grown overnight at 30°C in a shake flask from frozen stocks. The indene feed
(and in chemostat cultures, the medium feed) was initiated when the fermentor culture
reached an OD$_{600}$ of approximately 10 (mid-exponential growth of the culture). The feed
medium was the same composition as the fermentor medium unless otherwise noted. In
chemostat cultures a 1.0 L liquid volume was maintained by removing liquid overflow in
the vessel using a peristaltic pump. The indene was sparged into the fermentor using
one air stream sparged through 200 mL liquid indene contained in a flask subsequently
combined with another pure air stream prior to sparging into the fermentor (see Figure
3.1). The flow rates of each air stream were adjusted for a total flow rate of 1.0 L/min.
The gas-phase indene concentration was measured using a PE Photovac 2020
photoionization detector (PE Photovac, Markham, Ontario) calibrated with a 100 ppm
isobutylene calibration gas and set with a response factor of 0.3.

3.2.3 Biphasic indene bioconversions

Indene bioconversions performed using silicon oil as an organic phase were
carried out in the Bioprocess Research Facility at Merck Research Laboratories, Rahway,
NJ. In a 23 L Chemap fermentation vessel, 19.2 g (NH$_4$)$_2$SO$_4$ (Fisher Scientific), 14.0 g/L
MgSO$_4$·7H$_2$O (Fisher Scientific), and 14.0 g MOPS buffer were added to 11.0 L deionized
water. pH was adjusted pre-sterilization to 7.2 using 50% w/w NaOH solution (Fisher
Scientific). Following sterilization at 123°C for 35 min, 1.2 L of 50% glucose, 14.0 mL of
A9 trace elements solution, 14.0 mL of Stock A solution, 492.8 mL of 1 M phosphate
buffer, and 0.21 g CaCl$_2$·2H$_2$O was added to the vessel for a final aqueous volume of
approximately 13 L. For KY1(pDS3) bioconversions, 14.0 mL of 10.0 g/L gentamicin
sulfate was also added to the fermentor. Seed cultures were prepared by first culturing 1
mL frozen stock in 250 mL LB (Stage 1) for 34 h at 220 rpm and 30°C, and transferring 10
mL of this culture to 2.0 L LB (Stage 2), grown for 38 h at 220 rpm and 30°C. 400 mL of
the Stage 2 culture was used to inoculate the vessel. The initial fermentor operating
conditions were controlled online at 30°C, 8 L/min airflow, 5 psig backpressure, 400 rpm
agitation speed, and pH 7.0 by addition of 20% NaOH. Oxygen uptake and CO₂ evolution rates were monitored online prior to indene addition. After the fermentation broth reached an OD₆₀₀ of approximately 15.0, 2.0 L of silicon oil containing 60 mL of 90% indene was aseptically pumped into the fermentor. To prevent depletion of indene by both enzymatic oxidation and stripping into the gas phase, 15.0 mL indene was added 24 h following the introduction of silicon oil, and 60 mL indene was added 65 h post-oil addition.
3.3 Molecular biology methods

3.3.1 DNA manipulation methods

Restriction enzymes were purchased from New England Biolabs and used according to manufacturer instructions. PCRs were performed using a PTC-200 chamber (MJ Research, Waltham, MA) and according to standard protocols (Sambrook, Fritsch et al. 1989). DNA sequencing was performed at the MIT Biopolymers Facility (Cambridge, MA) using an ABI cycle sequencer by primer walking.

3.3.1.1 Preparation of plasmid DNA

Maxipreparations of plasmid DNA were performed using the Wizard Plus Maxipreps DNA Purification System (Promega, Madison, WI). Minipreparations of plasmid DNA in E. coli JM109 transformants were performed using a boiling lysis protocol (Sambrook, Fritsch et al. 1989). 1.5 mL of E. coli cells cultured overnight in LB with appropriate antibiotic were centrifuged and the pellet was resuspended in 250 μL TES (50 mM Tris pH 8.0, 50 mM EDTA, 15% sucrose). 20 μL of 10 mg/mL lysozyme was added to the suspension, and the cells were incubated at room temperature for 1-10 min. Cells were then placed in boiling water for 1 min, after which the tubes were microcentrifuged for 10 min at 13 000 rpm. The pellet was removed and the plasmid DNA was then precipitated on ice for 30 minutes by adding 150 μL ammonium acetate and 900 μL ethanol. After centrifugation and vacuum drying the pellet using a SpeedVac Evaporator/Concentrator Model S11 (Savant Instruments, Inc., Farmingdale, NY), the pellet was resuspended in 50 μL TE (10 mM Tris·HCl (Tris (hydroxymethyl) aminomethane) pH 8.0, 1 mM EDTA).

3.3.1.2 Preparation of genomic DNA from Rhodococcus

10 mL cultures were grown in LB (with appropriate antibiotic, where applicable) until the culture appeared cloudy. The cells were transferred to a 15 mL polypropylene centrifuge tube and pelleted at 6000 g for 5 min in a GSA rotor or equivalent. The
supernatant was discarded and the cell pellet was frozen at -80°C for 30-60 min or overnight at -20°C. The cell pellet was resuspended in 250 µL fresh 10 mg/mL lysozyme in TE and transferred to a fresh 1.5 mL microcentrifuge tube. 20 µL of 1.0 mg/mL mutanolysin was added to the suspension, and this was incubated at 37°C for 1-2 hr while gently shaking. 50 µL of 0.5 M EDTA, 50 µL of 10% SDS (sodium dodecyl sulfate), and 50 µL of 5 M NaCl were then added in succession and mixed gently, and the suspension was then kept at 4°C for 10 min. 10 µL of freshly prepared 20 mg/mL proteinase K was then added and the mixture was incubated at 37°C for 60 min. Following the incubation, the cell slurry was transferred to a new microcentrifuge tube containing 0.18 g of sodium perchlorate. This was mixed gently and incubated at room temperature for 15 min. DNA was extracted by adding 150 µL saturated phenol (Labscienific, Inc., Livingston, KY) and 150 µL of 24:1 chloroform:isoamyl alcohol to the suspension and vortexing thoroughly, followed by microcentrifugation at 4000 g for 5 min. The aqueous (top) phase was removed from the microcentrifuged tube and re-extracted with 150 µL of 24:1 chloroform:isoamyl alcohol in a new microcentrifuge tube. This was mixed well and recentrifuged at 4000 g for 5 min. The DNA was precipitated by removing the aqueous phase and adding it to 300 µL cold 2-propanol (Mallinckrodt Co., Paris, KY) in a new microcentrifuge tube. The DNA was pelleted by microcentrifugation at 8000 g for 15 min. The 2-propanol was discarded and the pellet was washed with 500 µL of 70% ethanol (Aaper Alcohol and Chemical Co., Shelbyville, KY) followed by microcentrifugation at 4000 g for 5 min. The ethanol was discarded, the pellet was vacuum dried, and the DNA was resuspended in 200 µL TE.

RNA was degraded by adding 1 µL of 0.5 mg/mL RNase (Boehringer Mannheim, Indianapolis, IN) to the DNA and incubating at 37°C for 30 min. 100 µL of 7.5 M ammonium acetate was added, and the DNA/ammonium acetate solution was extracted by adding 150 µL phenol and 150 µL of 24:1 chloroform:isoamyl alcohol to the suspension followed by microcentrifugation as described above. The aqueous phase was extracted as described above, and the DNA was then precipitated by adding the aqueous phase to 600 µL ethanol in a new microcentrifuge tube and microcentrifuging at 8000 g.
for 15 min. The pellet was washed with 500 μL of 70% ethanol, vacuum dried, and the DNA was resuspended in 200 μL TE.

3.3.2 Transformation methods

3.3.2.1 Preparation of competent *E. coli* JM109

Electrocompetent *E. coli* JM109 cells were prepared by culturing aerobically overnight at 37°C in test tubes containing 10 mL LB. 1.0 mL of this culture was used to inoculate 100 mL LB in a 1.0 L baffled shake flask. The culture was incubated at 37°C until the OD$_{600}$ reached 0.8-1.0. The cells were then transferred to four, 50 mL polypropylene centrifuge tubes and maintained at 4°C for at least one hour. The cells were centrifuged at 6000 g and washed three times in deionized water at 4°C, and resuspended in 1.0 mL 10% glycerol. 140 μL aliquots were placed in sterile 1.5 mL microcentrifuge tubes, where the electrocompetent cells were stored at -80°C until use.

3.3.2.2 Transformation of competent *E. coli* JM109

Transformation of competent JM109 cells was performed by first thawing the cells on ice and then adding 5-15 μL DNA to 70 μL cells in a sterile microcentrifuge tube. This was kept on ice for 10 min, and the suspension was then transferred to a chilled 2.0 mm gap electroporation cuvette (BTX, San Diego, CA). Cells were electroporated using a BioRad Gene Pulser set at 2.5 kV, 25 μF, and 200 Ω, and then permitted to recover by immediately adding 300 μL LB and transferring to a sterile microcentrifuge tube. The cells were incubated at 37°C for 1-4 h, after which they were spread onto LB-agar plates with appropriate antibiotics.

3.3.2.3 Preparation of competent *Rhodococcus* sps. I24 and KY1

Electrocompetent *Rhodococcus* cells were prepared by inoculating test tubes containing 10 mL LB with a single colony from plated I24 or KY1 cells, and incubating for 1-2 days at 37°C. 0.1-5 mL of the inoculum was transferred to 100 mL MB-Glycine (5
g/L yeast extract, 15 g/L bacto tryptone, 5 g/L bacto soytone, 5 g/L NaCl, and 15 g/L glycine) in a 1.0 L baffled shake flask. This culture was incubated at 37°C overnight or until the OD600 reached 0.6. The cells were pelleted by centrifuging for 5 min at 6000 g in a GSA rotor using sterile centrifuge bottles or 50 mL conical tubes. The cell pellet was resuspended in 30 mL ice-cold EPB1 (20 mM Hepes, 5% glycerol, pH 7.2) and recentrifuged for 5 min at 6000 g. The resuspension and centrifugation was repeated in EPB1, and the pellet was then washed in 10 mL ice-cold EPB2 (5 mM Hepes, 15% glycerol, pH 7.2). This was then centrifuged for 5 min at 8000 g and the supernatant was discarded. The final cell pellet was resuspended in 1.0 mL EPB2. 140 µL aliquots were placed in sterile 1.5 mL microcentrifuge tubes, where the electrocompetent cells were stored at -80°C until use.

3.3.2.4 Transformation of competent *Rhodococcus* sps. I24 and KY1

Transformation of competent *Rhodococcus* cells was performed by first thawing the cells on ice and then mixing 5-15 µL DNA with 70 µL cells. This was kept on ice for 5-10 min, and the suspension was then transferred to a chilled 2.0 mm gap electroporation cuvette. Cells were electroporated using a BioRad Gene Pulser set at 2.5 kV, 25 µF, and 400 Ω, and then recovered in 300 µL LB. The cells were incubated at 37°C for 1-20 h, after which they were spread onto LB-agar plates with appropriate antibiotics.

3.3.3 Construction of *Rhodococcus erythropolis limA*

The *limA* gene (Genbank accession No. Y18005) was synthesized in three fragments via template-less PCR with the primers listed in Table 3.1 (purchased from Gibco/BRL Life Technologies, Grand Island, NY), and the fragments were joined by the technique of splicing by overlap extension (SOE) (Horton, Hunt et al. 1989). First, primers limA1 and limA2; primers limA3 and limA4; and primers limA5 and limA6 were annealed in pairs and extended with thermal cycling (Figure 3.2a). The PCR profile included 19 cycles at 94°C for 1 min, 50°C for 30 sec, and 72°C for 1 min, preceded by 5 min at 95°C. The PCR products were gel-purified using the GeneClean II kit (Bio 101,
Carlsbad, CA), and 1 μl from each of these first-round PCR products was then used in a second-round PCR SOE reaction including primers limA5' and limA3', with thermal cycling including 30 cycles at 94°C for 1 min, 50°C for 1 min, and 72°C for 2 min, preceded by 5 min at 94°C. The final ca. 550 bp SOE product was gel-purified and ligated into the pCR2.1-TOPO vector from the pCR2.1-TOPO cloning kit (Invitrogen, Carlsbad, CA). Several of the resulting plasmids were isolated and sequenced, revealing that each carried a small number of point mutations relative to the published sequence for the limA gene. Error-free regions from the limA open reading frames (ORFs) of three of these plasmids were combined via standard restriction enzyme digestion and ligation reactions to create the plasmid pDS2 (Figure 3.2b). Sequencing of pDS2 determined that this plasmid carried no point mutations relative to the published sequence for limA.

3.3.4 Plasmid construction

A vector carrying a gentamicin resistance marker, the origin of replication from pEP2 and the lambda Pl promoter was prepared for constitutive expression of limA in Rhodococcus. pAL233 (Lessard, O'Brien et al. 1999), which carries the NG2 rep function from pEP2, was cleaved with SacII and NotI, blunted with T4 polymerase, and religated to remove the bulk of the sacB coding region, producing pAL241. The Pl promoter from lambda phage was first prepared as a fusion with the spectinomycin resistance marker (Spec) from the omega interposon. The cl857 gene from lambda phage was amplified from lambda 857 genomic DNA (New England Biolabs) with PCR Primer 1 and Primer 2 (Table 3.1). The lambda Pl promoter was likewise amplified by PCR with Primer 3 and Primer 4. These fragments were then joined by SOE, bringing the cl857 gene upstream of the lambda PL. The resulting product was ligated into pCR-Script SK+ (Stratagene, La Jolla, CA) with the 3' end of the lambda Pl adjacent to the KpnI-end of the polylinker, creating the plasmid pJWR1. The Spec ORF was excised as a BspHI fragment from pUT (de Lorenzo, Herrero et al. 1990), blunted with Klenow fragment of DNA polymerase, and ligated into the SmaI site of pBluescript KS+ (Stratagene) such that the start codon of Spec was adjacent to the BamHI site in the pBluescript polylinker, to produce pJWS1. Moving Spec as a BamHI-XhoI fragment from pJWS1 to the BamHI and XhoI sites of
pJWR1 produced pJWRS1, in which Spec was under the control of the lambda P_l-cl857 repressor/operator system. With this spectinomycin resistance marker as a template, the lambda P_l element was amplified by PCR with primers Spec-1 and Spec-PL; the Spec ORF was then amplified with primers SD-Spec and Spec-2a (Table 3.1). These two fragments were joined by SOE, creating a spectinomycin resistance marker under the control of lambda P_l (containing only one complete operator site, OL1), an improved ribosome binding site, and a Scal site between the promoter and Spec. This fragment was ligated into pCR-Script SK+ with the promoter adjacent to the NotI site, forming pAPE1b. The Scal-(NotI blunted) fragment from pAPE1b was then ligated to a (Ddel-BglII) blunted fragment from pBluescript SK+ carrying the origin of replication and the lacZ alpha complementation element, creating a plasmid, protoAPE2, oriented with the lambda P_l adjacent to the lacZ. Cutting protoAPE2 with FspI and Smal followed by self-ligation produced pAL242. pAL242 was linearized with NotI and inserted into the EagI site of pAL241, producing pAL243. The gentamicin resistance marker from pGM160 (Muth, Nubbamer et al. 1989) was excised as an Sphi-BspHI fragment and ligated into the Sphi and NcoI sites of pAL243 to create pAL245, which brought together the gentamicin resistance marker, the NG2 rep and the lambda P_l-Spec marker.

The limA ORF was excised from pDS2 as an EcoRV-Sphi fragment and ligated into the Scal and Sphi sites of pAL245, creating pDS3 (Figure 3.3). The negative control plasmid pAL282 was prepared by deleting the EcoRI restriction fragment that contained the limA ORF from pDS3.
<table>
<thead>
<tr>
<th>Strains</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rhodococcus</em> sp. KY1</td>
<td>Isolated from a continuous culture of <em>Rhodococcus</em> sp. I24</td>
<td>(Yanagimachi 2000)</td>
</tr>
<tr>
<td><em>Rhodococcus</em> sp. I24</td>
<td>Converts indene to indandiol by several pathways</td>
<td>(Chartrain, Jackey et al. 1998; Buckland, Drew et al. 1999)</td>
</tr>
<tr>
<td><em>Escherichia coli</em> JM109</td>
<td>*endA1, recA1, gyrA96, thi, hsdR17, (r5-, m5+), relA1, supE44, Δ(lac-proAB), [F', traID36, proAB, lacIΔZΔM15]</td>
<td>(Yanisch-Perron, Vieira et al. 1985)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDS3</td>
<td>Gent8, NG2 ori, limA</td>
<td>This study</td>
</tr>
<tr>
<td>pAL282</td>
<td>Gent8, NG2 ori</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Primer sequences (5'-3')**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>limA1</td>
<td>CACAGGGGAGCCCCTCACATGACATCAAAGATCAGAAACCTCTGCTGGCGGTCCAGAAGGA</td>
</tr>
<tr>
<td></td>
<td>CAGTGCCGCGCGGCCTGCAGGCTCCAGAAGGAAAGATCGT</td>
</tr>
<tr>
<td>limA2</td>
<td>GGTGATCTAGTCCGCTTAGGACATCGACACTGCTGGTTGACAAAAGTCTCGGGGTTACCTCGGAG</td>
</tr>
<tr>
<td></td>
<td>CGGATATGAGACGCTTCACCCCGATCGGAGGAGCTCAACCTGTCAACTTCTCCGAGTC</td>
</tr>
<tr>
<td>limA3</td>
<td>GACGAAACTCTGGCTGCGGCTTGTCACCTGCAAGATGACATCGA</td>
</tr>
<tr>
<td></td>
<td>CGAAGTATGATTTAAGGAAAGGA</td>
</tr>
<tr>
<td>limA3'</td>
<td>CCACAGGAGGCCCTCAGCATG</td>
</tr>
<tr>
<td>limA4</td>
<td>GTGGGTAGTGCCTGCTAGGACATCGACACTGCTGGTTGACAAAAGTCTCGGGGTTACCTCGGAG</td>
</tr>
<tr>
<td></td>
<td>CGGATATGAGACGCTTCACCCCGATCGGAGGAGCTCAACCTGTCAACTTCTCCGAGTC</td>
</tr>
<tr>
<td>limA5</td>
<td>CCTACGCCGACTACCCCGAAGAGCTGAGTCAACCTGTGCAACTTCTCCGAGTC</td>
</tr>
<tr>
<td></td>
<td>CGTACACCGGAGGGCAAGATACCGGGTGTTGCGGTG</td>
</tr>
<tr>
<td>limA5'</td>
<td>CCACAGGAGGCCCTCAGCATG</td>
</tr>
<tr>
<td>limA6</td>
<td>GGAAAGGGATTAGCGCGCGGAGGGAAGGTGACAGCTTTCTTCCAATTCGCGAGAT</td>
</tr>
<tr>
<td></td>
<td>CGAAGTATGATCGACCTCGAACCCTTCCTCCGAGGATGACATG</td>
</tr>
<tr>
<td>Primer 1</td>
<td>GAAACCCGGTACCTCGACGCAAGAT</td>
</tr>
<tr>
<td>Primer 2</td>
<td>CCCGCAGCTGAGGCGGGAGATGCAACCATTATCACG</td>
</tr>
<tr>
<td>Primer 3</td>
<td>GGCGGTGATAATTGGCTGCACTGCCCCGGCCCTCAGCCGCGGGG</td>
</tr>
<tr>
<td>Primer 4</td>
<td>GCTAACCCTAGTACGTCTCACCTCTGCGAGCTAATTCGCCGGA</td>
</tr>
<tr>
<td>Spec-1</td>
<td>TGCTATCTGCTGCGGACATAATATACAGATGC</td>
</tr>
<tr>
<td>Spec-2a</td>
<td>CATATTACCTCCGAGATATATATATATACGCC</td>
</tr>
<tr>
<td>SD-Spec</td>
<td>AGGAAGGCCCCTCAGGA</td>
</tr>
<tr>
<td>Spec-PL</td>
<td>TCACTGGGGGCTCCCTCCGCTGAACCGTGGATCT</td>
</tr>
</tbody>
</table>
Figure 3.2  (a) Schematic of synthesis of limA gene by PCR (b) Correction of limA point mutations following ligation of limA PCR product into pCR-2.1TOPO. Shaded arrows indicate point errors detected through sequencing; the dashed segment of plimA8 was not sequenced.
Figure 3.3 Diagram of pDS3, which expresses the limA-encoded epoxide hydrolase in *Rhodococcus* sp. KY1 from the constitutive lambda P<sub>L</sub> promoter. *aacC1*, gentamicin resistance marker from Tn1696; NG2 rep, origin of replication derived from pEP2.
3.4 Analytical and preparative methods

3.4.1 Biomass determination

Biomass concentrations were determined by calibrating OD<sub>600</sub> to dry cell weight in single-phase and shake-flask indene bioconversions. The presence of silicon oil in biphasic bioconversions precluded biomass determinations in such experiments. Dry cell weight was measured by centrifuging 10 mL culture a: 12 000 g for 15 min, washing the cell pellet twice in deionized water, and then drying the sample at 110°C for 2 days. Figure 3.4 shows a typical calibration curve used for biomass determination.

![Graph showing biomass calibration curve](image)

**Figure 3.4** Typical biomass calibration curve for *Rhodococcus* sp. KY1
3.4.2 Off-gas analysis

Respiratory data was obtained for selected single-phase bioconversions (prior to indene addition) by feeding the fermentor off-gas to a Perkin Elmer Multiple Gas Analyzer 1600 mass spectrometer. The CO₂ evolution rate (CER) and oxygen uptake rate (OUR) were calculated according to Eqs. 3.1 and 3.2, respectively:

\[
CER = \frac{v_{in}}{V} \left( \frac{y_{N_2}^{in} - y_{N_2}^{out}}{y_{N_2}^{out}} \right) \left( \frac{y_{CO_2}^{out} - y_{CO_2}^{in}}{y_{CO_2}^{in}} \right)
\]  

(3.1)

\[
OUR = \frac{v_{in}}{V} \left( \frac{y_{O_2}^{in} - y_{O_2}^{out}}{y_{O_2}^{out}} \right) \left( \frac{y_{N_2}^{out} - y_{N_2}^{in}}{y_{N_2}^{in}} \right)
\]  

(3.2)

where \( y_i \) are the nitrogen, oxygen, and carbon dioxide mole fractions in the inlet and outlet gas streams, \( v_{in} \) is the molar flow rate of the gas in the bioconversion, and \( V \) is the reactor volume.

3.4.3 Protein analysis

Protein concentrations were determined with the Bradford assay using bovine serum albumin (BSA) as the standard (Bio-Rad Laboratories, Hercules, CA). SDS-PAGE analysis of proteins was performed using a Protean IIxi analytical protein electrophoresis unit (Bio-Rad Laboratories, Hercules, CA). First, a 5.0 mL separating gel was prepared (Table 3.2) and added to the unit, and 5.0 mL of deionized water was applied to the top of the separating gel until solidification to prevent bubble formation. The water was subsequently discarded, and a 3.0 mL stacking gel (Table 3.2) was applied above the separating gel. The electrophoresis unit was filled with running buffer (0.2 M glycine, 0.1% SDS, and 50 mM Tris-HCl pH 8.3), and 10 μL sample was mixed with an equal volume of loading buffer (125 mM Tris-HCl pH 6.8, 10% 2-mercaptoethanol, 10% SDS, 10% glycerol with bromophenol blue), prior to application on the gel. Gels were run at 80 V for approximately 2 h using a Bio-Rad Power Pac 300 Unit (Bio-Rad Laboratories, Hercules, CA). The gels were stained for 5 min in a microwave at high power in a mixture consisting of 0.2% coomassie brilliant blue solution in 45% methanol, 45% water
and 10% acetic acid. Gels were then placed in destaining solution (25% methanol, 65% water and 10% acetic acid) overnight.

<table>
<thead>
<tr>
<th>Table 3.2</th>
<th>Components of gels used for SDS-PAGE analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>15% Separating gel</td>
</tr>
<tr>
<td>Deionized water</td>
<td>3.50 mL</td>
</tr>
<tr>
<td>30% acrylamide (National Diagnostics, Atlanta, GA)</td>
<td>7.5 mL</td>
</tr>
<tr>
<td>Tris+SDS</td>
<td>3.75 mL (1.5 M, pH 8.8)</td>
</tr>
<tr>
<td>10% ammonium persulfate</td>
<td>50 μL</td>
</tr>
<tr>
<td>TEMED</td>
<td>10 μL</td>
</tr>
</tbody>
</table>

3.4.4 Preparation of indene and indene metabolites

3.4.4.1 Synthesis of [14C]-indene

[14C]-indene was synthesized at Merck Research Laboratories, Rahway, NJ. Details of the synthesis have been published previously (Yanagimachi 2000). Briefly, the synthesis of [14C]-indene was carried out beginning with [14C]-benzene as shown in Figure 3.5. The tracer was supplied as an ethanol solution (29 mCi, 425 μCi/mg, in 7 mL of ethanol, >99.6% radiochemical purity by HPLC). A second stock solution was prepared separately (2.22 mCi, 435 μCi/mg, in 1.81 mL of ethanol).
Figure 3.5  $[^{14}C]$-indene synthesis scheme. Adapted from (Yanagimachi 2000).

3.4.4.2 Preparation of (1S,2R)-inden oxide and trans-(1R,2R)-indandiol

(1S,2R)-inden oxide was provided as a heptane solution (93% pure, 84% e.e., 1075 µg/mL) by Merck Research Laboratories (Rahway, NJ). trans-(1R,2R)-indandiol was purified using two different methods. The first method is a variation of that described previously (Chartrain, Jackey et al. 1998). Chirality of the trans-indandiol in *Rhodococcus* sp. KY1 indene bioconversion broth was first determined by chiral HPLC using the methods described below. Broth was initially centrifuged at 14 000 g in 200 mL centrifuge bottles to remove macroscopic particulates, and then filtered using AquaPrep 600 groundwater filtration capsules (0.45 µm, Gelman Sciences, Ann Arbor, MI). The filtrate was stored at 4°C until use. A Bio-Rad Econo-Column (2.5 cm ID x 30 cm, Bio-Rad, Hercules, CA) was packed with 100 mL Amberchrom CG-161S chromatographic resin (Tosoh Biosep, LLC, Montgomeryville, PA) according to manufacturer specifications. After equilibration with 200 mL of 10% ethanol, 200 mL of filtered broth was applied to the column at approximately 1.5 mL/min. trans-indandiol was eluted using a step-gradient of ethanol, starting at 200 mL of 30% ethanol, proceeding to 200 mL of 40% ethanol to 200 mL of 50% ethanol (2.0 mL/min). The column effluent was collected in 50 mL fractions. trans-indandiol eluted in the last three fractions of the 40% ethanol step. The column was re-equilibrated with 200 mL of 100% ethanol followed by
200 mL of 10% ethanol. The 40% ethanol fractions were mixed with an equal volume of ethyl acetate, centrifuged, and the organic layer was removed and dried under an air stream. The \textit{trans}-(1R,2R)-indandiol was found to be at least 99% pure by reverse-phase and chiral HPLC.

Alternatively, batches from five 15 L biphasic bioconversions were loaded into two 50 L tanks (with broth volumes of 40 L and 35 L) and allowed to settle overnight to separate the aqueous and silicon oil phases. The aqueous phase was cut from the silicon oil, and the settling was repeated. The silicon oil was washed with two volumes of water to retrieve remaining \textit{trans}-indandiol. The aqueous phase was filtered using a Millipore multipurpose filtration skid equipped with two ultrafiltration membranes (A/G Technology Co., Model #UFP500E55, 2.1 m² area each). Before use, the membranes were stored in 0.1 N NaOH and washed with NaOCl, pH 10.5 and 0.5 N NaOH followed by a water flush until the system pH was 8.75. The batches were filtered at a feed pressure of approximately 10 psig. The retentates from the first settled batches were washed with 2 x 2 L of water before the addition of the second lot of aqueous phase cuts to the system. Finally, the solids from all batches were rinsed with 4 x 2 L of water. These additions were filtered off and pooled with the main filtrate. Subsequently, an adsorption column was run employing a SP-207 resin (1081 cm³ volume, 3.5 cm ID x 112.5 cm, Mitsubishi Co., prepared by washing in 3 volumes of acetone and rinsing in water). Loading was carried out at a flowrate of approximately 10 column volumes per hour (~190 mL/min), with a total volume of 56 L filtrate loaded. The loaded column was washed with 2 L water and eluted with 20-25% acetonitrile. The product was detected in fractions collected after 0.5 L of eluent was collected in a combined fraction volume of approximately 3.5 L. The \textit{trans}-indandiol product was extracted by freeze drying and determined to be >99% pure by reverse-phase HPLC.
3.4.5 Abiotic phase measurements

3.4.5.1 Indene metabolites

Reverse phase HPLC was used to quantify all indene metabolites and normal-phase HPLC was used to determine the chiralities of both trans-indandiol and cis-indandiol. The HPLC systems used were (1) a Hewlett Packard 1050 series system, (2) a Waters 2690 separations module equipped with a Waters 996 photodiode array detector, or (3) a Hewlett Packard 1100 series system (at Merck Research Laboratories, Rahway, NJ).

Indene metabolites were quantified according to two separate reverse-phase assays. Samples were diluted in 1:1 acetonitrile:2-propanol, centrifuged, and pre-filtered using 0.2 μm pore PVDF 13 mm syringe filters (Alltech, Deerfield, IL). The primary assay used was on a Zorbax RX-C8 column (4.6 mm ID x 250 mm, Agilent Technologies) equipped with an RX-C8 guard column (Agilent Technologies). Metabolite separation was achieved using an acetonitrile (in deionized water) gradient run at 1.0 mL/min according to the following specifications: 10% to 40% (15 min) to 90% acetonitrile (25 min), followed by 90% acetonitrile run isocratically to 30 min, to 10% acetonitrile (35 min), run isocratically to 40 min. The column was equilibrated according to this gradient prior to sample injections. Indene metabolites were identified according to their characteristic ultraviolet spectra (Figure 3.6), and eluted at the following times (Figure 3.7): trans-indandiol (9.2 min), cis-indandiol (9.8 min), 1-keto-2-hydroxy- indan (11.1 min), 1-indenol (15.1 min), 1-indanone, (17.8 min), indan oxide (19.1 min), and indene (24.0 min). The column was periodically washed with 1.0 mL/min of 80% methanol for 60 minutes.

An alternative assay used a YMC CombiScreen C8 column (4.6 mm ID x 50 mm, YMC, Wilmington, NC). 10 μL of sample was injected into the column and an acetonitrile (in deionized water) gradient was run at 3.0 mL/min according to the following specifications: 10% acetonitrile run isocratically for 2.0 min, to 30% (3.0 min), to 90% (4.0 min) run isocratically until 6.0 min, to 10% at 8.0 min. The column was
equilibrated according to this gradient prior to sample injections. The indene metabolites eluted at the following times (Figure 3.8): trans-indandiol (1.1 min), cis-indandiol (1.2 min), 1-keto-2-hydroxy- indan (1.5 min), 1-indenol (2.5 min), 1-indanone, (2.8 min), indan oxide (3.6 min) and indene (4.2 min). The column was periodically washed with 3.0 mL/min of 80% methanol for 30 minutes.

The chiralities of trans-indandiol and cis-indandiol were determined using a Chiracel OJ normal-phase HPLC column (4.6 mm ID x 250 mm, Daicel Chemical Industries). Separation of the indandiol enantiomers was achieved by injecting 10 μL of sample and flowing a 9:1 mixture of hexanes and 2-propanol isocratically at 0.5 mL/min. The indene metabolites of interest eluted at the following times: trans-(1R,2R)-indandiol, 19.4 min; cis-(1S,2R)-indandiol, 22.2 min; cis-(1R,2S)-indandiol, 30.0 min (Figure 3.9).

Samples were prepared for analysis by first extracting the metabolites in an equal volume of ethyl acetate. Following vortexing and centrifugation at 15 000 g to rapidly pellet the cells and separate the organic and aqueous phases, the organic (top) layer was removed and evaporated under an air stream. The residual was dissolved in an appropriate volume of 2-propanol, and was filtered as described above prior to injection.

3.4.5.2 [14C]-Indene metabolites

Reverse phase HPLC was used to separate the indene-derived metabolites and measure the radioactive counts of labeled metabolites and the metabolite concentrations during the [14C]-transient experiments. Samples were taken from the shake flask and placed into 1.0 volumes of a 1:1 mixture of acetonitrile and 2-propanol at 4°C. Following centrifugation the supernatant was filtered as described above. 10 μL of the filtrate was injected into an HPLC system (Hewlett Packard Series 1050, San Fernando, CA) equipped with a Zorbax RX-C8 column (4.6mm x 25 cm) (Hewlett Packard). The HPLC protocol was followed as described above. The UV detector was used to quantify the concentrations of the unlabeled metabolite for each time point in the 14C-transient experiments.

A flow scintillation analyzer (Radiomatic 150TR, Packard Instrument Co., Meriden, CT) was connected downstream of the HPLC system for radioactivity
measurements. 3 mL/min of liquid scintillation cocktail (Ultima-Flo M, Packard Instrument Co.) was mixed in the analyzer with the 1.0 mL/min of HPLC column effluent. The peak areas (total counts) corresponding to each radiolabeled indene metabolite were reported.

3.4.5.3 Glucose

Glucose concentrations were measured using a commercial assay (Sigma) based upon the hexokinase-catalyzed phosphorylation of glucose and subsequent reduction of NAD\(^+\) to NADH according to the following reactions:

\[
\text{Glucose} + \text{ATP} \xrightarrow{\text{hexokinase}} \text{Glucose - 6 - Phosphate} + \text{ADP} \\
\text{Glucose - 6 - Phosphate} + \text{NAD}^+ \xrightarrow{\text{glucose-6-phosphate dehydrogenase}} 6 - \text{phosphogluconate} + \text{NADH}
\]

Samples were diluted such that the glucose concentration was less than 7.5 g/L. Glucose concentrations were determined based upon endpoint absorbance at 340 nm, measured 10 minutes following addition of enzyme to the sample.

3.4.5.4 Ammonia

Ammonia concentrations were measured using a commercial assay (Sigma) based upon the reductive amination of 2-oxoglutarate and concomitant oxidation of NADPH:

\[
2 - \text{oxoglutarate} + \text{NH}_3 + \text{NADPH} \xrightarrow{\text{glutamate dehydrogenase}} \text{Glutamate} + \text{NADP}
\]

Samples were diluted such that the ammonia concentration was less than 15 mg/L. Ammonia concentrations were determined based upon an endpoint decrease in absorbance at 340 nm proportional to NADPH oxidation, measured 10 minutes following addition of enzyme to the sample.
Figure 3.6 UV spectra for indene metabolites accumulated in *Rhodococcus* sps. 124 and KY1 cultures. (a) trans-indandiol, (b) cis-indandiol, (c) 1-keto-2-hydroxy-indan, (d) 1-indenol, (e) 1-indanone, (f) indan oxide, (g) indene.
Figure 3.7  Zorbax RX-C8 reverse-phase HPLC chromatogram at 220 nm. Indene metabolites elute as follows: *trans*-indandiol (9.2 min), *cis*-indandiol (9.8 min), 1-keto-2-hydroxy-indan (11.1 min), 1-indenol (15.1 min), 1-indanone, (17.8 min), indan oxide (19.1 min), and indene (24.0 min).
Figure 3.8  Elution of (a) trans-indandiol (1.1 min), cis-indandiol (1.2 min), 1-keto-2-hydroxy-indan (1.5 min), 1-indenol (2.5 min), 1-indanone, (2.8 min), and (b) indan oxide (3.6 min) and indene (4.2 min) on the YMC CombiScreen C8 HPLC column. Chromatograms shown are at 220 nm.
Figure 3.9  Elution of (a) cis-(1S,2R)-indandiol (22.2 min) and cis-(1R,2S)-indandiol (30.0 min) and (b) trans-(1R,2R)-indandiol (19.4 min) on Chiralcel OJ normal-phase HPLC column with 0.5 mL/min 9:1 hexanes:2-propanol. Chromatograms shown are at 220 nm.
3.4.5.5 Phosphate

Phosphate concentrations were measured using a commercial assay (Sigma) based upon the following reaction:

\[
\text{Inorganic phosphorus} + \text{H}_2\text{SO}_4 + \text{Ammonium Molybdate} \rightarrow \text{Unreduced phosphomolybdate complex}
\]

Samples were diluted such that the phosphate concentration was less than 150 mg/L. Phosphate concentrations were determined based upon an endpoint decrease in absorbance at 340 nm proportional to the inorganic phosphorus concentration, measured 10 minutes following addition of enzyme to the sample.

3.4.6 Enzymatic assays

3.4.6.1 Preparation of cell lysates

KY1 culture was centrifuged at 27,000 g at 4°C. The supernatant was removed and the cells were washed in lysis buffer at pH 7.0 containing 100 mM HEPES, 10% glycerol, 0.2 mM 2-mercaptoethanol, and 2 tablets/50 mL of protease inhibitor cocktail (Complete, EDTA-free, Boehringer Mannheim, Indianapolis, IN). The cells were centrifuged and resuspended in lysis buffer. The cells were passed through a pre-chilled French press cell (Spectronic Instruments, New York) at 16,000 psi three times. The mixture was then centrifuged at 27,000 g for 30 min at 4°C, and the top fraction was removed and stored at -80°C until use.

*E. coli* JM109 (and transformed strains) cell lysates were made by collecting three, 100 mL overnight cultures by centrifugation at 10,000 g at 4°C and washing with 4.0 mL of 50 mM potassium phosphate buffer (pH 7.0) per g wet cells, and repelleting in the same centrifugation conditions. The cells were then resuspended in the same buffer (2.0 mL buffer/g wet cells) and disrupted by sonication at 4°C for 2 min using a Heat Systems Sonicator® Ultrasonic Processor XL system. Cell debris was removed by centrifugation at 20,000 g at 4°C for 10 min. The cleared lysate was stored at 4°C until use (no more than 2 hours after lysate preparation).
3.4.6.2 Indene monooxygenase assay

A 2.5 mL culture sample was removed from a culture during indene bioconversion. The sample was placed in a 125 mL screw-cap shake flask. 2 µL aqueous ethanol containing 435 µCi/mL [14C]-indene was then added to the flasks. The flask was resealed and reincubated. Samples were taken at 0, 5, 10, 15, 30, and 60 min after addition of [14C]-indene. Cells were mixed with an appropriate volume of 1:1 acetonitrile:2-propanol so that the indene metabolite concentrations measured by HPLC were between 10 and 100 mg/L.

For inhibition experiments, appropriate amounts of trans-(1R,2R)-indandiol were immediately added to the shake flasks following removal from the primary culture to achieve the final metabolite concentration desired. The flask was sealed and placed in a shaker at 300 rpm, 30°C for 20 minutes prior to addition of [14C]-indene.

3.4.6.3 Limonene-1,2-epoxide hydrolase

1-3 mL of cell lysate was placed in a 125 mL screw-cap shake flask and (1S,2R)-indan oxide was added to the final concentration as indicated. Indan oxide stock was prepared by mixing 15 µL of 93% pure (1S,2R)-indan oxide (84% e.e., 1075 µg/mL) in 60 µL of chilled ethanol for a final stock concentration of 215 µg/mL. 5 µL of this stock was added per mL of lysate to give the final concentration desired. The lysate was incubated at 30°C and agitated at 300 rpm. Samples were taken periodically as described above for 60 min after indan oxide addition for HPLC analysis.

For inhibition experiments, appropriate amounts of trans-(1R,2R)-indandiol were added to the lysates to achieve the final metabolite concentration desired. The flask was sealed and placed in a shaker at 300 rpm, 30°C for 20 minutes prior to addition of (1S,2R)-indan oxide.
3.4.6.4 Indan oxide cell-free hydrolysis

10 mL of defined medium, adjusted from pH 5-10 using 2 M NaOH and 2 M HCl, was placed in a 125 mL screw-cap shake flask and (1S,2R)-indan oxide (83% ee) was added to a final concentration of 1.0 mM by the procedure used for the limonene-1,2-epoxide hydrolase assay. The flasks was incubated at 30°C and agitated at 300 rpm. Samples were taken as described above periodically for up to 12 h after indan oxide addition.
CHAPTER 4 - RESULTS AND DISCUSSION: DIRECTING THE STEREOSELECTIVE SYNTHESIS OF (2R)-INDANDIOL IN RHODOCOCCUS

In terms of a general framework for metabolic engineering of uncharacterized bioconversion networks, the work described in Chapter 2 satisfied the first two points of that proposed in the introduction. A steady state chemostat with a gas-phase indene delivery system was used as an experimental setup to apply selective pressure to the Rhodococcus sp. I24 strain and ultimately isolate the KY1 strain with improved bioconversion properties. Furthermore, the chemostat was used in conjunction with radiolabeled tracers to (1) identify the non-enzymatic hydrolysis of indan oxide, the consequent formation of both cis-(1S,2R)-indandiol and trans-(1R,2R)-indandiol enantiomers, and the lack of a trans-(1R,2R)-indandiol dehydrogenase, and (2) perform a rigorous flux analysis of the KY1 indene bioconversion network that identified the monooxygenase pathway as the key route of indene oxidation.

This chapter describes the genetic modification, phenotype confirmation, and system analysis of Rhodococcus sp. KY1 to facilitate improved production of (2R)-indandiol. These results are intended to validate the later three points of the proposed metabolic engineering framework: (iii) target identification, (iv) flux redistribution, and (v) analysis of the modified bioconversion phenotype to identify opportunities for further strain and/or process improvement.

4.1 Strategies for modification of indene bioconversion in Rhodococcus sp. KY1

4.1.1 Prioritization of targets in the indene bioconversion network

The identification of indene monooxygenase as the key oxidizing enzyme in Rhodococcus sp. KY1, and the non-enzymatic nature of indan oxide hydrolysis, provided several possible approaches to improving the yield of (2R)-indandiol in KY1 indene
bioconversions. Notwithstanding overall titer improvement that may in part entail monooxygenase overexpression, three distinct engineering paths were evident:

(i) Introduction of epoxide hydrolase-type activity catalyzing the selective hydrolysis of indan oxide to \textit{trans}-(1\textit{R},2\textit{R})-indandiol

(ii) Knockout of the \textit{cis}-(1\textit{S},2\textit{R})-indandiol dehydrogenase activity

(iii) Knockout of the naphthalene dioxygenase activity catalyzing the indene-to- \textit{cis}-(1\textit{R},2\textit{S})-indandiol reaction

A key aspect of metabolic engineering in uncharacterized microorganisms is the need to prioritize network targets in light of available biochemical and molecular biology tools. Each of the three aforementioned targets is discussed further below in the context of the previously uncharacterized nature of the I24 and KY1 strains.

4.1.1.1 Indan oxide node

A number of epoxide hydrolases that catalyze the asymmetric addition of water to an epoxide ring to give the corresponding \textit{trans}-dil have been characterized (see Chapter 2). Most relevant to \textit{Rhodococcus} indene bioconversion is the limonene-1,2-epoxide hydrolase cloned and purified from \textit{R. erythropolis} DCL14. This enzyme is a good candidate for expression in KY1 because of its small size and lack of cofactor requirements. Consequently, the metabolic burden (with respect to central carbon metabolism) and risk of introducing pleitropic effects as a result of epoxide hydrolase overexpression may be insignificant (Yaguchi, Mitsui et al. 2000; Hausler, Rademacher et al. 2001; Hicks, Lockington et al. 2001). Successful introduction of this activity would potentially increase the \textit{trans}-(1\textit{R},2\textit{R})-indandiol yield to approximately 97%, based upon steady state flux analysis results reported previously (Yanagimachi 2000). By reducing or eliminating the formation of \textit{cis}-(1\textit{S},2\textit{R})-indandiol from (1\textit{S},2\textit{R})-inden oxide, production of the byproduct 1-keto-2-hydroxy-indan would be drastically diminished by eliminating the primary substrate for that dehydrogenation reaction. Only a small amount of 1-keto-2-hydroxy-indan would presumably be formed, as a result of the naphthalene dioxygenase pathway giving \textit{cis}-(1\textit{R},2\textit{S})-indandiol that accounts for only
approximately 3% of the indene oxidation flux in steady state cultures. The theoretical result of epoxide hydrolase overexpression in terms of the KY1 indene bioconversion network is illustrated in Figure 4.1a.

4.1.1.2 cis-(1S,2R)-indandiol dehydrogenase knockout

A second approach to improvement in (2R)-indandiol yield would be to knockout the dehydrogenase enzyme(s) degrading cis-(1S,2R)-indandiol to the byproduct 1-keto-2-hydroxy-Indan. Similar to epoxide hydrolase expression, the theoretical yield of (2R)-indandiol would increase to 97%; however, this would be distributed between the cis- and trans-(2R)-indandiol enantiomers in a 4:3 ratio at the physiological pH used in flux analysis studies. The theoretical change in the KY1 indene bioconversion network as a result of cis-(1S,2R)-indandiol dehydrogenase knockout is illustrated in Figure 4.1b.

A number of substantial barriers make this option less desirable. First, gene knockout(s) require substantial development of molecular biology tools designed specifically for Rhodococcus sp. KY1 (and/or I24). One unique feature of the indene bioconversion network, that the enzymes involved in indene bioconversion presumably evolved for the metabolism of other compounds such as naphthalene and toluene, suggests that these tools would need to be applied not only to a single figurative “cis-(1S,2R)-indandiol dehydrogenase” enzyme, but instead to multiple enzymes: perhaps a naphthalene dihydrodiol dehydrogenase, a toluene dihydrodiol dehydrogenase, in addition to other enzymes that may be highly non-specific in nature. Although all of these genes have not yet been identified in I24 and KY1, the Rhodococcus sp. I24 genome has been sequenced and annotated, and DNA microarray studies are currently underway toward identifying these and other genes important in indene bioconversion (Jefferson Parker, personal communication).

A second barrier, as described above, is that a successful cis-(1S,2R)-indandiol dehydrogenase knockout would result in the primary products being a mixture of the cis- and trans-(2R)-indandiol enantiomers. While, as discussed earlier, either of these enantiomers are valid precursors in the chemical synthesis of (-)-CAI, in terms of
downstream processing it is preferable to design a biocatalyst giving an enantiomerically pure product.

4.1.1.3 Naphthalene dioxygenase knockout

A third option, albeit with minimal theoretical improvement in trans-(1R,2R)-indandiol yield (Figure 4.1c), is to knockout the naphthalene dioxygenase activity giving cis-(1R,2S)-indandiol. The genes encoding this activity have been identified in *Rhodococcus* sp. 124 (Treadway 1999; Treadway, Yanagimachi et al. 1999), where this enzymatic activity is more prominent and a knockout is a more substantial step toward improving (2R)-indandiol yield, by reducing the formation of both cis-(1R,2S)-indandiol and 1-indenol. However, with the naphthalene dioxygenase only supporting 3% of the indene oxidation flux in KY1, only after more promising approaches are exhausted should this knockout be pursued to potentially realize a small improvement in product yield.

4.1.1.4 Prioritization of targets

In light of the state of knowledge of *Rhodococcus* genetics, it was determined that emphasis should be placed on the identification of enzymes catalyzing the selective hydrolysis of indan oxide to trans-(1R,2R)-indandiol. In terms of genetic modification, this task is more efficient than focusing on the development of multiple enzyme knockouts. Such secondary targets to improve (2R)-indandiol yield, while valid, are likely to require substantially more investment relative to the improvement in yield realized. Plasmids that can replicate in *Rhodococcus* have been developed (Treadway, Yanagimachi et al. 1999) that served as the foundation of a vector for the expression of this epoxide hydrolase in KY1. This section next describes the expression of limonene-1,2-epoxide hydrolase in KY1 and the characterization of a non-enzymatic route to modifying the flux split ratio at the indan oxide node in *Rhodococcus* sp. KY1.
Figure 4.1  Theoretical (2R)-indandiol selectivities in KY1 derivatives as a result of (a) epoxide hydrolase expression, (b) cis- (1S,2R)-indandiol dehydrogenase knockout, and (c) naphthalene-inducible dioxygenase knockout, based upon metabolic flux analysis results reported previously (Figure 2.8).
4.1.2 Heterologous expression of limonene-1,2-epoxide hydrolase

4.1.2.1 *Escherichia coli* JM109

4.1.2.1.1 Genetic confirmation

The *limA* gene, which encodes limonene-1,2-epoxide hydrolase (Barbirato, Verdoes et al. 1998; van der Werf, Overkamp et al. 1998; van der Werf, Orru et al. 1999), was constructed by template-free PCR (as described in Chapter 3) and placed in the expression plasmid pDS3 (Figure 3.3) using the cloning strain *Escherichia coli* JM109. Plasmid integrity was validated through a battery of restriction digests on maxi-prepped DNA from the transformed strain (Figure 4.2).


4.1.2.1.2 Enzymatic activity

Cleared lysates of JM109(pDS3) and JM109(pAL282) were created by sonication and subsequent centrifugation, and used for a preliminary activity assay for limonene-1,2-epoxide hydrolase on (1S,2R)-inden oxide. The results of this assay are shown in Figure 4.3, with the accompanying SDS-PAGE results in Figure 4.4. The JM109(pDS3) lysate showed increased formation of trans-indandiol from the indan oxide substrate relative to the blank plasmid (pAL282) and buffer (cell-free) controls. Additional controls used were boiled lysates (both JM109(pDS3) and JM109(pAL282)) to account for non-enzymatic hydrolysis of indan oxide, which showed nearly the same hydrolysis profiles as did the buffer control. trans-(1R,2R)-indandiol was formed in 47% e.e. in the pDS3 lysate relative to the cis-(1S,2R)-indandiol enantiomer, while the trans-(1R,2R)-indandiol e.e. was 26% for both pAL282 and buffer. Accounting for non-enzymatic indan oxide hydrolysis, the epoxide hydrolase activity in the JM109(pDS3) lysate was approximately 0.05 nmol (mg protein)$^{-1}$ min$^{-1}$. While the data from this experiment confirmed epoxide hydrolase activity on indan oxide, the low activities indicated that E. coli likely would not be an optimal host for recombinant indene bioconversion.

![Graph showing conversion of limonene-1,2-epoxide hydrolase activity](image)

**Figure 4.3** Conversion of 2.0 mM (1S,2R)-inden oxide (83% e.e.) to trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol by lysates prepared from *E. coli* JM109 transformants. Residual indan oxide concentration not shown.
Figure 4.4  SDS-PAGE analysis of crude lysates prepared from JM109 transformants. Limonene-1,2-epoxide hydrolase corresponds to band at 17 kDa in JM109(pDS3) lane.

4.1.2.2 *Rhodococcus* sp. KY1

4.1.2.2.1 Genetic confirmation

The expression vector pDS3, harboring the *limA* gene, was transformed into competent KY1 cells by electroporation, as well as *Rhodococcus* sp. I24 (Appendix B). Following transformation, approximately 10 colonies were selected from LB-Gent10 agar plates, for which plasmid integrity was validated through a battery of restriction digests on genomic DNA retransformed into *E. coli* JM109 (Figure 4.5). Only candidate “E” demonstrated the correct digestion patterns expected for pDS3, while candidates “B”, “D”, and “H”, as well as the other candidates screened, appear to have insertions near or in the NG2 replication origin of pDS3. Thus, candidate “E” was subsequently denoted as KY1(pDS3). It is also noted that when creating the control strain harboring a blank plasmid, KY1(pAL282), using the same transformation protocol, over 40 candidates were screened using restriction digests to identify a transformant with an intact plasmid, with most of the other candidates also having an insertion in the NG2 replication region of pAL282.
Figure 4.5  Restriction digests of KY1(pDS3) candidates. Only candidate “E” has correct digestion patterns with respect to intact pDS3, while candidates “B”, “D”, and “H” appear to have insertions near or in the NG2 replication origin.

4.1.2.2.2  Indene bioconversion screening of transformation candidates

The KY1 transformation candidates were also screened for epoxide hydrolase activity in shake flask indene bioconversions. Indandiol concentrations were measured by reverse-phase HPLC 18 h following the addition of indene to the shake flask cultures. Several candidates did not show detectable indene bioconversion activity in the time-scale studied; those that did are shown in Figure 4.6. While all four transformation candidates that oxidized indene to a detectable extent showed an improved trans/cis-(2R)-indandiol ratio relative to the KY1 control, only candidate “E” showed product titers comparable to KY1. One possibility to explain the decreased titers in other transformation candidates is that the insertion event that appears to occur in the NG2 replication origin of pDS3 and pAL282 may not be confined to that plasmid or chromosomal location alone. If this event were to occur near the gene(s) responsible for indene monooxygenation, decreased oxidation activity may result. Nonetheless, the results of this screening assay, together with the results of the restriction digests indicate that the limA gene was successfully transformed into KY1 in a single candidate with an intact expression vector and improved formation of trans-indandiol at early culture times.
Figure 4.6  Indene bioconversions by KY1(pDS3) transformation candidates, as denoted in restriction digests of Figure 4.5

4.1.2.2.3 Enzymatic activity

Cleared lysates of KY1(pDS3) and KY1(pAL282) were created using a French press cell and subsequent centrifugation, and used to assay for limonene-1,2-epoxide hydrolase activity on (1S,2R)-indan oxide. Upon addition of (1S,2R)-indan oxide to KY1(pDS3) lysate, a significant increase in the rate of formation of trans-indandiol was observed relative to KY1 and controls (Figure 4.7). Chiral HPLC analysis confirmed that trans-(1R,2R)-indandiol was formed in approximately 85% e.e., corresponding to the 83% e.e. (1S,2R)-indan oxide substrate used for the assay. The specific epoxide hydrolase activity observed on indan oxide in KY1(pDS3) lysate was $77 \pm 12 \text{ nmol} \cdot (\text{mg protein} \cdot \text{min})^{-1}$. 
Figure 4.7  Conversion of 8.0 mM (1S,2R)-incene oxide (83% e.e.) to trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol by lysates prepared from *Rhodococcus* sp. KY1 transformants.

SDS-PAGE of the lysates showed a clear band at 17 kDa, which is the predicted size of this epoxide hydrolase (Figure 4.8) (van der Werf, Orru et al. 1999). The signal observed at 17 kDa for the KY1(pAL282) control is likely due to the similarly sized *aacC1* product. Taken together, these results demonstrate that the expression of limonene-1,2-epoxide hydrolase into KY1 results in the directed, enzymatic hydrolysis of (1S,2R)-indan oxide to trans-(1R,2R)-indandiol in cell lysates and *in vivo* by virtue of shake flask results.
Figure 4.8  SDS-PAGE analysis of cleared lysates prepared from KY1 transformants.

4.1.3 Directed, non-enzymatic hydrolysis of indan oxide by changes in medium pH

To determine the effect of pH on the kinetics of indan oxide hydrolysis, (1S,2R)-indan oxide (83% e.e.) was added to fresh cultivation medium adjusted from pH 5-10 at 30°C. Substantial differences in both the hydrolysis products formed and the rate of indan oxide hydrolysis, described by first order kinetics (Figure 4.9), were observed at different pH values (Figure 4.10). Of particular interest is the dramatic increase in the relative amount of trans-(1R,2R)-indandiol formed at pH 8 and higher. From just over 50% trans-indandiol formed from non-enzymatic hydrolysis of indan oxide at pH 7, the percentage of trans-indandiol formed relative to the total indandiol concentration, which for first order kinetics is equivalent to the ratio of the rate constants \( k_T/k_{C2R} \), increased to 71% at pH 8, 92% at pH 9, and 96% at pH 10. Conversely, at sub-physiological pH values, the formation of cis-(1S,2R)-indandiol is favored over trans-(1R,2R)-indandiol. At pH 6, only 34% of the (2R)-indandiol formed is in the trans configuration, and at pH 3, 4, and 5, 28% is in the trans configuration (\( k_T/k_{C2R} = 0.38 \)). At pH ≤ 5, the total indan oxide hydrolysis rate constant was extremely large (\( k_T + k_{C2R} ≥ 70 \)) and could not be measured more accurately due to sampling limitations. The total indan oxide hydrolysis rate constant (\( k_T + k_{C2R} \)) did not change as dramatically from pH 7 to pH 10. In all samples,
both the \textit{trans}-(1R,2R)-indandiol and \textit{cis}-(1S,2R)-indandiol were formed in approximately 80\% e.e., indicating that chirality at the C-2 carbon of the substrate is conserved at all pH values tested.

The drastically different hydrolysis results observed at different pH conditions suggest that multiple mechanisms are important in the degradation of indan oxide. It has been suggested previously that both \textit{S}$_{\text{N}}$1 and \textit{S}$_{\text{N}}$2 mechanisms may be important (Yanagimachi 2000), but the data presented herein clarify the relative importance of these hydrolysis routes under different conditions. At high pH, a nucleophilic \textit{S}$_{\text{N}}$2 addition of a hydroxyl anion likely accounts for the high enantiomeric excess of \textit{trans}-(1R,2R)-indandiol formed with respect to \textit{cis}-(1S,2R)-indandiol, while at low pH a \textit{S}$_{\text{N}}$1 mechanism likely predominates (Figure 4.11). In contrast to the previously proposed nature of \textit{S}$_{\text{N}}$1 hydrolysis, which suggests that \textit{trans}-(1R,2R)-indandiol is the favored product due to sterics, the \textit{cis}-(1S,2R)-indandiol is more likely favored via this mechanism as indicated by the results obtained at pH 3-5. Both mechanisms presumably are important at intermediate (physiological) pH levels. This suggests that \(k_T\) is comprised of two rate constants: \(k_T = k_{T(SN1)} + k_{T(SN2)}\), where \(k_{T(SN1)} = 0.38\ k_{C2R}\).
Figure 4.9  First order kinetics of \((1S,2R)\)-inden oxide non-enzymatic hydrolysis in cell-free defined medium at 30°C at (a) pH 10, (b) pH 9, (c) pH 8, (d) pH 7, (e) pH 6.
Figure 4.10  Effect of medium pH on (1S,2R)-indan oxide hydrolysis. 10 mM indan oxide was added to 10 mL culture medium adjusted to pH 5-10. Rate constants were derived from linear regression of semilog representations of indan oxide depletion data.

Figure 4.11  Proposed competing mechanisms for (1S,2R)-indan oxide hydrolysis.
4.1.4 Strategies for indene flux redistribution: Summary

Two independent approaches for the directed production of trans-(1R,2R)-indandiol have been developed. First, the limonene-1,2-epoxide hydrolase from Rhodococcus erythropolis DCL14 was expressed in Rhodococcus sp. KY1 and shown to hydrolyze (1S,2R)-inden oxide preferentially to trans-(1R,2R)-indandiol. Also, it was determined that the nature of indan oxide hydrolysis to trans- and cis-indandiol is strongly pH-dependant, with primarily trans-indandiol formed in basic conditions. Next, experiments are described to demonstrate improved indene bioconversion phenotypes as a result of both introduction of epoxide hydrolase activity and the implementation of high-pH conditions in KY1 and recombinant indene bioconversions.

4.2 Indene bioconversion in Rhodococcus sp. KY1 and recombinant derivatives

4.2.1 Effect of epoxide hydrolase expression on indene bioconversion

Following confirmation of limA expression, the ability of recombinant strain KY1(pDS3) to preferentially convert indene to trans-indandiol in batch-mode indene bioconversions was tested. In the following presentation, the results are described with respect to trans-indandiol selectivity as opposed to yield, since selectivity better reflects changes in the flux split ratio from indan oxide to both (2R)-indandiol enantiomers. As data shown in Table 4.1 indicate, selectivity changes are indicative of similar changes in yield. (2R)-indandiol selectivity (S) and yield (Y) are defined in Eqs. (4.1) and (4.2), where the selectivity calculation excludes the concentration of indan oxide due to its nature as an unstable intermediate:

\[ S = \frac{[T]}{[T] + [C2R] + [C2S] + [K] + [1 - \text{indenol}] + [1 - \text{indanone}]} \] (4.1)

\[ Y = \frac{[T]}{[\text{indene uptake}]} \] (4.2)
Table 4.1  (2R)-indandiol yields and selectivities in Rhodococcus sps. KY1 and I24 chemostat cultures. Steady state metabolite data from (Yanagimachi 2000).

| Steady State Values | Strain I24 | | | Strain KY1 | | | |
|---------------------|-----------|-----------|-----------|-----------|-----------|-----------|
|                     | D=0.10 h⁻¹ | D=0.10 h⁻¹ | D=0.065 h⁻¹ |
| (2R)-dial yield (μmol diol/μmol indene)⁺ | 0.32 0.30 | 0.66 0.61 | 0.62 0.55 |
| (2R)-dial selectivity | 0.32 0.30 | 0.61 0.54 | 0.53 0.55 |
| trans-(1R,2R)-indandiol (mg/L) | 0 0 | 86 181 | 151 262 |
| cis-(1R,2S)-indandiol (mg/L) | 18 23 | 6 8 | 5 8 |
| cis-(1S,2R)-indandiol (mg/L) | 106 129 | 24 52 | 35 55 |
| 1-keto-2-hydroxy indan (mg/L) | 32 51 | 25 93 | 96 154 |
| 1-indenol (mg/L) | 69 86 | 0 0 | 0 0 |
| 1-indanone (mg/L) | 82 118 | 0 0 | 0 0 |
| indan oxide (mg/L) | 0 0 | 21 42 | 34 55 |
| indene (mg/L) | 3 6 | 10 14 | 5 6 |
| Biomass (g DCW/L) | 2.7 2.4 | 3.2 3.6 | 4.9 3.7 |
| (2R)-dial productivity (μmol/g DCW/h)⁺ | 26 36 | 23 43 | 18 35 |

⁺(2R) diol productivity and yield determined for sum of trans and cis (2R) enantiomers.
First, to ensure that the blank expression vector did not affect the indene bioconversion profiles in KY1, wild-type KY1 and the control recombinant strain KY1(pAL282) were grown in separate fermentors using defined medium at pH 6.9 ± 0.1 and placed under similar indene feed conditions of 75 ppm indene in 1.0 vvm (1.0 L/min) of air. Under the indene feed conditions and cell densities studied, the dissolved oxygen concentration was at least 85% of air saturation at all times and thus the cultures were not oxygen-limited. As shown in Figure 4.12 (KY1) and Figure 4.13 (KY1(pAL282)), the indene metabolite profiles observed in the control KY1(pAL282) bioconversion were consistent with those for wild-type KY1. The first products formed upon initiation of the indene feed (denoted as t = 0 h) are indan oxide, and both trans-indandiol and cis-indandiol (with trans slightly favored). While the trans-indandiol concentrations continue to increase to approximately 800 mg/L in KY1 and 350 mg/L in KY1(pAL282), cis-indandiol no longer accumulates at the expense of the formation of its dehydrogenation product, 1-keto-2-hydroxy-indan. The higher product titers observed in KY1 are likely in part due to the higher biomass concentrations achieved in the defined medium by the wild-type strain; both KY1(pAL282) and KY1(pDS3) achieved lower final biomass concentrations than KY1 on the same defined medium.

When the recombinant KY1(pDS3) strain, harboring the limA gene encoding limonene-1,2-epoxide hydrolase, was tested for indene bioconversion phenotype under similar conditions, the new strain showed an improved overall indene conversion profile over that observed for KY1 and KY1(pAL282) batch cultures (Figure 4.14). The improvement exhibited by KY1(pDS3) was particularly pronounced during the initial period of the fermentation. For approximately 25 hours following the addition of indene, trans-indandiol was the primary terminal product formed in the bioconversion with an average selectivity of approximately 95%. This product presumably resulted from the activity of the newly introduced epoxide hydrolase on the (1S,2R)-indan oxide formed by the indene monooxygenase enzyme. Following this initial period, indan oxide began to accumulate, and then cis-(1S,2R)-indandiol and 1-keto-2-hydroxy-indan appeared in the fermentation medium after the indan oxide concentration reached a concentration of approximately 150 mg/L. This unusual indene bioconversion profile is
likely the result of an initial sole epoxide hydrolase activity followed with the competing non-enzymatic hydrolysis of indan oxide at later times as the concentration of indan oxide in the extracellular medium increases.

To confirm that epoxide hydrolase activity was not lost in the bioconversion period where cis-(1S,2R)-indandiol and 1-keto-2-hydroxy-indan begin to accumulate, lysates were created from cells removed from the KY1(pDS3) bioconversion at 90.5 h, 110.5 h, and 133.5 h. The cells were washed to remove residual indene and indene metabolites during the course of the lysing procedure. As shown in Figure 4.15, significant epoxide hydrolase activity was measured at 90.5 h, well after the undesired metabolite profile was observed (less robust but detectable epoxide hydrolase activity was also measured at the later two time points). Furthermore, it is noted that the initial detection of cis-(1S,2R)-indandiol and 1-keto-2-hydroxy-indan occurred after the cells had reached the stationary phase, indicating that the plasmid likely was not lost from actively growing cells. Nonetheless, cells were removed from the bioconversions of both KY1(pDS3) and KY1(pAL282) and successfully cultured on LB-Gent_{10} agar plates, indicating that both recombinant strains maintained resistance to gentamicin during the indene bioconversion and thus likely retained the respective plasmids.

4.2.2 Effect of high-pH conditions on indene bioconversion

Based on the results of in vitro, cell-free (1S,2R)-indan oxide hydrolysis, Rhodococcus sp. KY1(pDS3) was used for indene bioconversion under high-pH conditions in a single-phase batch fermentor. After indene feed initiation and substrate uptake occurred in the late-exponential phase of cell growth, the culture pH was adjusted to pH 8.6. With product titers comparable to those observed at physiological pH (6.9), the KY1(pDS3) strain at high pH showed dramatically improved indene bioconversion profiles over previous cultures of either KY1 or KY1(pDS3) at physiological pH (Figure 4.16). No more than 25 mg/L of the byproduct 1-keto-2-hydroxy-indan accumulated at any point in the bioconversion medium compared to approximately 300 mg/L for the KY1(pDS3) (pH 7.0) bioconversion.
Figure 4.12  (a) Indene bioconversion metabolite profiles for single-phase batch KY1 run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1, and biomass concentration of 7.0 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. The two products of non-enzymatic indan oxide hydrolysis, trans-(1R,2R)-indandiol and cis-(1S,2R)-indandiol, appear immediately upon the initiation of indene oxidation. (b) Gas and aqueous-phase indene concentrations. The gas inlet indene concentration was manually controlled.
Figure 4.13  (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pAL282) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1 and biomass concentration of 4.2 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. Indene bioconversion profiles are similar to those observed for KY1 (Figure 4.12). (b) Gas and aqueous-phase indene concentrations. The gas inlet indene concentration was manually controlled.
Figure 4.14  (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pDS3) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm), pH 6.9 ± 0.1, and biomass concentration of 4.9 ± 0.6 g DCW/L. Time is indicated as post-addition of indene to the cells. At early culture times (up to approximately 25 h), the only products formed are indan oxide and trans-(1R,2R)-indandiol. (b) Gas and aqueous-phase indene concentrations.
Figure 4.15  Epoxide hydrolase activity in cell lysates made from cells removed from KY1(pDS3) bioconversion (Figure 4.14) at 90.5 h. Cells were washed during the lysing procedure to remove residual indene and metabolites. 8.0 mM indan oxide was added to the lysates at the beginning of the assaying procedure.

4.2.3 Biphasic indene bioconversions with KY1(pDS3) and high pH: Product selectivities

A significant benefit of sparging indene into Rhodococcus indene bioconversions is that substrate concentrations can be maintained at low aqueous concentrations relative to the alternative of adding liquid indene directly to cultures. As Figures 4.12-4.14 and Figure 4.16 show, the aqueous-phase indene concentration is nearly zero during the most active phase of indene bioconversion. The indene uptake rate then steadily decreases for the remainder of the bioconversion phase of the culture. At low residual indene concentrations, substrate toxicity (discussed in Chapter 2) is presumably minimized. Unfortunately, this method of substrate addition is not practical in large-scale industrial applications where concerns about solvent containment are paramount. Particularly in the case of the system described here, indene is readily stripped into the exiting gas due to its low aqueous solubility (Yanagimachi 2000). Consequently, earlier work with
Rhodococcus sp. I24 and Pseudomonas putida F1 utilized a biphasic system where silicon oil is used as the organic phase to solubilize indene and reduce substrate stripping (Chartrain, Jackey et al. 1998; Buckland, Drew et al. 1999). Some data have indicated that the oxygenated indene derivatives partition primarily into the aqueous phase with only indene residing primarily in the organic phase (Kodzo Gbewonyo, personal communication), although no partition coefficient data is available. To test the potential of KY1 and/or the epoxide hydrolase bearing recombinant strain as industrially viable microorganisms, preliminary biphasic bioconversions were performed at the 23 L scale.

Four parallel biphasic bioconversions were performed in 23 L vessels containing a final volume of 13 L aqueous media and 2.0 L of silicon oil. 60 mL of indene was added to the fermentors together with the silicon oil when the biomass concentrations reached approximately 2.0 g/L. Both KY1 and KY1(pDS3) were cultivated separately under two conditions: one reactor at pH 7.0 throughout the run, and a second reactor for each strain was adjusted to pH 8.5 after the first indene bioconversion products were detected. The complete indene bioconversion profiles for each of these runs are shown in Appendix C. The total indene oxidized in the four runs ranged from 300 mg/L to 1100 mg/L for the KY1(pDS3), pH 7.0 run. While the biphasic system requires additional process development to improve trans-indandiol titers, the trans-indandiol selectivities in each of the cultures were consistent with those observed in the corresponding single-phase bioconversions (Figure 4.17), and indicate that these strains are suitable for developing an industrially viable process.
Figure 4.16  (a) Indene bioconversion metabolite profiles for single-phase batch KY1(pDS3) run in defined medium with an indene feed concentration of 75 ± 10 ppm (1.0 vvm) and pH 8.6 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. (b) Gas and aqueous-phase indene concentrations.
Figure 4.17 shows the effect of epoxide hydrolase expression and high-pH conditions on the indene bioconversion phenotype of KY1 and recombinant strains in both single-phase and biphasic experiments, described by the metric of trans-indandiol selectivity. As expected, the qualitative indene bioconversion profiles (hence selectivities) were consistent between the two types of experiments. Figure 4.17a shows that limonene-1,2-epoxide hydrolase affects trans-indandiol selectivities at early culture times (KY1(pDS3), pH 6.9), while high pH conditions result in improved selectivity late in the bioconversion period (KY1(pDS3), pH 8.6). trans-indandiol was produced by KY1 pDS3 (pH 8.6) at greater than 92% selectivity throughout the culture, and was resolved to 100% purity by the end of the bioconversion due to the slow degradation of 1-keto-2-hydroxy-indan. The data of Figure 4.17b confirms these results in a biphasic system, and also shows that the wild-type KY1 strain can be used under high-pH conditions to obtain trans-indandiol selectivities comparable to those seen with the epoxide hydrolase-expressing strain.

4.2.4 Recombinant indene bioconversions: Summary

The results of enzymatic assays with KY1(pDS3) cell lysate show that the in vitro activity of limonene-1,2-epoxide hydrolase on (1S,2R)-indan oxide compares favorably with the indene oxidation flux measured in continuous cultures of KY1 (Yanagimachi 2000). Indeed, the initial formation of only trans-indandiol from indan oxide in fermentations of KY1(pDS3) at pH 7.0 indicates that the limonene-1,2-epoxide hydrolase is fully functional in vivo. However, the level of activity apparent early in the bioconversion is not maintained for the duration of the process as evidenced by the presence of indan oxide in the medium and the later accumulation of cis-indandiol and 1-keto-2-hydroxy-indan. The inability of the expressed limonene-1,2-epoxide hydrolase activity to support the indene monooxygenase flux in batch culture required that alternative approaches be pursued to further reduce or eliminate the hydrolysis of indan oxide to cis-indandiol. The strong dependence of the selectivity of indan oxide hydrolysis on pH provides an excellent tool for further modulation of the flux distribution in the indene bioconversion network.
Figure 4.17  
*trans-*indandiol selectivity profiles for batch indene bioconversions performed in (a) single-phase (aqueous) and (b) biphasic (aqueous/silicon oil) modes.
From the bioconversion data alone, the reason(s) for the decline of the epoxide hydrolase-catalyzed indan oxide hydrolysis are not clear. It is possible that the epoxide hydrolase deactivates with time or is inhibited by the rising concentrations of some indene metabolites. Yet another possibility is that indan oxide formation by a membrane-bound monooxygenase is accompanied by significant indan oxide efflux to the medium and subsequent non-enzymatic hydrolysis yielding cis-indandiol. A third possibility is that the turnover of limonene-1,2-epoxide hydrolase expressed in KY1 is insufficient for the indan oxide concentrations achieved in these bioconversions. To investigate some of these possibilities, a series of in vitro inhibition experiments were performed, as described in the next section. However, it is noted that whatever the explanation for the appearance of extracellular indan oxide, a high pH medium would ensure its full conversion to the desired product.

4.3 Regulation of indene bioconversion in *Rhodococcus* sp. KY1 and recombinant derivatives

4.3.1 Theory

Previous work to analyze steady state flux distributions in *Rhodococcus* sp. KY1 (Yanagimachi 2000; see Chapter 2 for further details) showed that the kinetics of $^{14}$C-indene oxidation can be modeled by first order kinetics using Eqs. (2.8)-(2.13):

\[
\frac{d[I^*]}{dt} = -(k_{MO} + k_{DO})[I^*]
\]  
(2.8)

\[
\frac{d[IO^*]}{dt} = k_{MO}[I^*] - (k_{C2R} + k_T)[IO^*]
\]  
(2.9)

\[
\frac{dT^*}{dt} = k_T[IO^*] - k_{TDH}[T^*]
\]  
(2.10)

\[
\frac{d[C2R^*]}{dt} = k_{C2R}[IO^*] - k_{RDH}[C2R^*]
\]  
(2.11)

\[
\frac{d[C2S^*]}{dt} = k_{DO}[I^*] - k_{SDH}[C2S^*]
\]  
(2.12)
\[
\frac{d[K^*]}{dt} = k_{RDH}[C2R^*] + k_{SDH}[C2S^*] + k_{TDH}[T^*]
\]  
(2.13)

As a result of the first order behavior of \(^{[14]C}\)-indene oxidation, an effective whole-cell monooxygenase rate constant \(k_{MO}\) was used to quantify differences in the enzyme activity, as described below. Assuming the oxygenase enzymes are described by a Michaelis-Menten kinetic model and that the residual indene concentration is small relative to the Michaelis constant, Eq. (2.7) can be rewritten as shown in Eq. (4.3), where \(k_{MO}^{cat}\) and \(k_{DO}^{cat}\) are the first order reaction rate constants (turnover numbers) for the oxidations of indene to (1S,2R)-indan oxide and \(cis\)-(1R,2S)-indandiol respectively, and \([E_{MO}^{ACTIVE}]\) and \([E_{DO}^{ACTIVE}]\) are the total active enzyme concentrations (both free and substrate-bound) for the two oxygenase systems.

\[
\frac{d[I^*]}{dt} \approx \left( k_{MO}^{cat} \frac{[E_{MO}^{ACTIVE}]}{K_m^{MO}} \right) + \left( k_{DO}^{cat} \frac{[E_{DO}^{ACTIVE}]}{K_m^{DO}} \right) [I^*] = -(k_{MO} + k_{DO}) [I^*]
\]  
(4.3)

By virtue of the KY1 flux analysis that demonstrated that the naphthalene-inducible dioxygenase activity, which produces \(cis\)-(1R,2S)-indandiol, is almost negligible, the depletion of \(^{[14]C}\)-indene can be assumed to be due solely to the indene monooxygenase activity, which allows Eq. (4.3) to be written as follows in an integrated form by neglecting \(k_{DO}\):

\[
[I^* (t)] = [I^* (0)] \exp(-k_{MO}t)
\]  
(4.4)

According to Eq. (4.4), semilog plots of the normalized residual labeled indene counts following introduction of labeled indene allow determination of the rate constant \(k_{MO}\). Comparison of such constants determined at different culture conditions allows a direct determination of the relative concentrations of active enzyme at the corresponding physiological states. This is shown in Eq. (4.5) for two arbitrary states “A” and “B”:

\[
\frac{k_B^{MO}}{k_A^{MO}} = \frac{[E_{MO}^{ACTIVE}]^B}{[E_{MO}^{ACTIVE}]^A}
\]  
(4.5)
Note that Eq. (4.4) is analogous to the equations used for the calculation of steady state fluxes in the indene bioconversion network (Chapter 2), recalling Eq. (2.5):

$$k_i = \frac{J_i}{[M_{tot}]} X$$  \hspace{1cm} (2.5)

Thus, the use of radiolabeled tracers is not an enzymatic assay in the conventional sense, but indeed an in vivo flux determination adapted for the analysis of bioconversion regulation in a biocatalytic network. Eqs. (4.3-4.5) allow for the quantitative description of enzyme regulatory activity in a non-steady state system.

4.3.2 Monooxygenase activity in KY1 indene bioconversion

In single-phase batch cultures of *Rhodococcus* sp. KY1 and recombinant derivatives, trans-indandiol accumulates to 600-800 mg/L with fair consistency (Figures 4.12-4.14; Figure 4.16), but titers range from 400 mg/L to in excess of 1.0 g/L. Attempts to improve overall product titer could focus on a number of issues, among them:

(i) Product inhibition

(ii) Low biocatalyst concentration

(iii) Growth dependence of enzyme expression/activity

(iv) Substrate and/or product toxicity

To investigate issues (i) and (ii), a series of enzymatic studies were performed to probe the monooxygenase and epoxide hydrolase enzymes. First, the activity of the key transformation in wild-type KY1, the monooxygenation of indene to indan oxide, was probed with the use of $^{14}$C-indene during a fed-batch culture. During the bioconversion phase of the fed-batch run, small samples of culture were removed from the reactor vessel, placed into a shake flask, and spiked with $^{14}$C-indene to a final concentration of approximately 20 µM. The magnitude of $k_{MO}$ at different points in the culture was compared and analyzed together with the unlabeled indene metabolite profiles to identify possible systemic phenomena that may play a role in indene
bioconversion. The aforementioned issues (iii) and (iv) are being addressed in parallel studies performed elsewhere, as discussed in Chapter 5.

Figure 4.18 shows the time profiles of terminal secreted unlabeled indene metabolites (i.e. all metabolites except indan oxide) in a KY1 fed-batch culture performed at pH 6.9. \textit{trans}-indandiol accumulates to approximately 1.2 g/L and 1-keto-2-hydroxy-indan byproduct to 1.6 g/L. The biomass concentration during the bioconversion phase of the culture was 7.5 ± 0.5 g DCW/L. Although it is not clear what causes the cessation of indene oxidation, one possibility is that the indene monooxygenase enzyme is feedback inhibited at higher concentrations of one or more indandiols and/or 1-keto-2-hydroxy-indan. To investigate this possibility, cells were removed at several time points during the fed-batch culture and their monooxygenase activity was probed by adding [\textsuperscript{14}C]-indene.

The first order kinetics of [\textsuperscript{14}C]-indene oxidation to all detectable radiolabeled metabolites in cells removed from the fed-batch bioconversion at 22 h is shown in Figure 4.19. Consistent with earlier metabolic flux analysis work in KY1, the first product formed is [\textsuperscript{14}C]-inden oxide, followed by the appearance of both [\textsuperscript{14}C]-\textit{trans}-indandiol and [\textsuperscript{14}C]-\textit{cis}-indandiol. The monooxygenase reaction rate constant, \(k_{MO}\), determined in the KY1 fed-batch culture at 55 h post-indene addition was reduced by approximately 80% relative to that calculated at 22 h (Figure 4.20). This reduction in \(k_{MO}\) coincided with the attenuation of \textit{trans}-(1R,2R)-indandiol and 1-keto-2-hydroxy-indan accumulation in the culture, suggesting a possible regulatory effect of \textit{trans}-(1R,2R)-indandiol and/or 1-keto-2-hydroxy-indan on indene monooxygenase. Further experimentation was performed to explore this possibility.

4.3.3 Effect of \textit{trans}-indandiol concentration on \(k_{MO}\)

To investigate a possible inhibitory effect of \textit{trans}-indandiol on indene monooxygenase in KY1, cells were removed from indene-induced steady state chemostat cultures at low indene feed concentrations (0.055 h\(^{-1}\), 70 ppm indene, 2.0 g DCW/L). \(k_{MO}\) was determined by measuring the depletion kinetics of [\textsuperscript{14}C]-indene with time as
suggested by Eq. (4.4), and this measurement was repeated in the presence of varying concentrations of \( \text{trans-}(1R,2R)\)-indandiol added to cells removed from the chemostat. The steady state concentrations of terminal indene metabolites (± 10 mg/L) in the chemostat were 50 mg/L \( \text{cis} \)-indandiol, 100 mg/L \( \text{trans} \)-indandiol, and 100 mg/L 1-keto-2-hydroxy-indan.

For the control condition (run in triplicate) where no supplementary compound was added to the culture, an observed monooxygenase rate constant \( k_{MO} = 3.1 \pm 0.5 \text{ h}^{-1} \) was determined using Eq. (4.4). As shown in Figure 4.21, the tracer kinetics are well-described by first order kinetics. Based upon the value of \( k_{MO} \) calculated for this continuous culture, the addition of 1.7 g/L \( \text{trans-}(1R,2R)\)-indandiol to indene-induced KY1 cells showed significant inhibition of the indene monooxygenase activity (Figure 4.21b). The inhibition of the indene monooxygenase demonstrated by \( \text{trans-}(1R,2R)\)-indandiol appear to be consistent with the reduction in \( k_{MO} \) values calculated during the KY1 fed-batch culture.

The addition of several concentrations of \( \text{trans} \)-indandiol to indene-induced KY1 cells resulted in a significant reduction in the effective rate constant fit for the monooxygenase enzyme (Figure 4.22). As shown in Eq. (4.5), this can be interpreted as a reduction in concentration of active monooxygenase enzyme with the addition of \( \text{trans} \)-indandiol to induced KY1 cells. Thus, the addition of 1.4 g/L of \( \text{trans} \)-indandiol resulted in an approximately 55% concentration reduction of active monooxygenase. Higher concentrations of \( \text{trans} \)-indandiol resulted in a further increase in monooxygenase inhibition up to 70% at 3.8 g/L. Further experimentation with 1-keto-2-hydroxy-indan in analogous experiments also indicate that this degradation product could play a regulatory role in indene bioconversion (Appendix E). However, for purposes of indene bioconversion network analysis in the context of focusing on the indan oxide node, \( \text{trans} \)-indandiol is the compound of primary concern.
Figure 4.18  (a) Indene metabolite bioconversion profiles for KY1 fed-batch culture. *trans*-indandiol accumulates to approximately 1.2 g/L and 1-keto-2-hydroxy-indan accumulates to 1.6 g/L. (b) Gas-phase indene concentrations.
Figure 4.19  [¹⁴C]-indene oxidation kinetics to all detectable radiolabeled metabolites in cells removed from KY1 fed-batch culture (Figure 4.18) 22 h after [¹⁴C]-indene addition. [¹⁴C]-indene depletion data was used to calculate $k_{MO} = 10.1$ h⁻¹ from Eq. (4.4).
Figure 4.20  Observed monooxygenase rate constant values in KY1 fed-batch culture. The concentration of active enzyme declines rapidly as trans-(1R,2R)-indandiol and 1-keto-2-hydroxy-indan accumulate in the culture.
Figure 4.21  $[^{14}C]$-indene oxidation kinetics for trans-indandiol inhibition experiment in cells removed from indene-induced chemostat culture. Curves were fit to tracer data by non-linear regression using Eqs. (2.8-2.13). (a) Control condition with no trans-indandiol added ($k_{MO} = 3.3$ h$^{-1}$). (b) 1.7 g/L trans-(1R,2R)-indandiol (run in triplicate) incubated with cells for 30 min prior to tracer addition ($k_{MO} = 0.82$ h$^{-1}$). Legend: □ indene, ▲ indan oxide, ♦ trans-indandiol, ■ cis-indandiol, * 1-keto-2-hydroxy-indan; kinetic fits are $I(t)$ for indene, $IO(t)$ for indan oxide, $T(t)$ for trans-indandiol, $C(t)$ for cis-indandiol, and $K(t)$ for 1-keto-2-hydroxy-indan.
4.3.4 Limonene-1,2-epoxide hydrolase kinetics

4.3.4.1 Effect of (1S,2R)-indan oxide

To test for the possibility that limonene-1,2-epoxide hydrolase was subject to substrate inhibition by (1S,2R)-indan oxide, the epoxide hydrolase activity assay was performed in KY1(pDS3) lysates at multiple concentrations of indan oxide. It is noted that indan oxide hydrolysis kinetics are comprised of two components in KY1(pDS3) - both enzymatic and non-enzymatic contributions as in Eq. (4.6), assuming Michaelis-Menten kinetics:

$$\frac{d[IO]}{dt} = -(k_{C2R} + k_T)[IO] - \frac{\nu_{max}[IO]}{K_m + [IO]}$$  \hspace{1cm} (4.6)
where the magnitudes of $k_{C2R}$ and $k_r$ are dependent upon, among other parameters, the pH of the lysate. Assuming that $(v_{max}/K_m) >> (k_{C2R} + k_r)$, the non-enzymatic hydrolysis terms can be neglected. This assumption is borne out from the data for the determination of the $K_m$ for limonene-1,2-epoxide hydrolase on (1S,2R)-indan oxide (Figure 4.23). The experimentally determined $K_m$ value was $77 \pm 21$ mM.

![Figure 4.23](image)

**Figure 4.23** Determination of kinetic parameters for limonene-1,2-epoxide hydrolase *in vitro* assay on (1S,2R)-indan oxide.

The data show that limonene-1,2-epoxide hydrolase is not subject to substrate inhibition by (1S,2R)-indan oxide at concentrations up to 4.5 mM (590 mg/L), well above concentrations encountered in KY1(pDS3) single-phase bioconversions. The possibility of substrate inhibition at higher concentrations could not be conclusively determined due to the apparent toxicity of the heptane solvent used for the indan oxide stock at elevated levels. Nevertheless, the results of this experiment exclude the possibility that either a low $K_m$ for indan oxide or substrate inhibition are responsible for the decline in epoxide hydrolase activity observed at late culture times in KY1(pDS3) indene bioconversions. Thus, additional experimentation was performed to determine if the other primary product in the early phase of KY1(pDS3) bioconversion, *trans*- (1R,2R)-indandiol, plays a role in the reduction of limonene-1,2-epoxide hydrolase activity.
4.3.4.2 Effect of trans-(1R,2R)-indandiol

To test for product inhibition of limonene-1,2-epoxide hydrolase by trans-(1R,2R)-indandiol, the epoxide hydrolase activity assay was performed in KY1(pDS3) lysates at concentrations of indan oxide up to 5 mM (650 mg/L) and trans-indandiol concentrations between 1-25 mM (up to 3.6 g/L). As determined from $v_o$ values that were similar within experimental error, again assuming non-enzymatic hydrolysis of indan oxide to be insignificant relative to the epoxide hydrolase activity, there was no evidence of product inhibition of the epoxide hydrolase by trans-(1R,2R)-indandiol.

4.3.5 Regulation of indene bioconversion: Summary

The first order kinetics of [$^{14}$C]-indene depletion by Rhodococcus sp. KY1 indicates that the key monooxygenase enzyme is not saturated under the conditions examined, i.e. $[I] \ll K_M$. Consequently, the enzyme appears to be capable of supporting higher concentrations of substrate. Although some evidence gathered from a KY1 chemostat at a dilution rate of 0.10 h$^{-1}$ suggests that [$^{14}$C]-indene oxidation may be better described by zero-order kinetics at aqueous indene concentrations of 10-15 mg/L (Yanagimachi 2000), most [$^{14}$C]-indene depletion data for the KY1 strain conclusively follows first order kinetics.

The primary results of the inhibition experiments were that (1) the presence of trans-indandiol at concentrations as low as 1.0 g/L was found to result in a decrease in indene monooxygenation in vivo as determined by radiolabeled tracer experiments, and (2) limonene-1,2-epoxide hydrolase activity is not adversely affected by accumulation of either (1S,2R)-indan oxide or trans-(1R,2R)-indandiol. The latter result suggests that the attenuation of epoxide hydrolase activity observed in vivo in KY1(pDS3) indene bioconversion is the result of an inherent low turnover number of the epoxide hydrolase enzyme.
4.4 Discussion

The results described in this chapter encompass the final three steps (defined in Chapter 1) of the metabolic engineering framework for application to biocatalytic systems. More specifically, this work satisfies the primary objective to design a robust (2R)-indandiol production strain, and to perform subsequent phenotype evaluation of the recombinant strain. Analysis of the metabolic fate of both (2R)-indandiol enantiomers formed from the hydrolysis of indan oxide led to a prioritized set of indene bioconversion network targets, primary among them being the direction of indan oxide hydrolysis toward trans-(1R,2R)-indandiol to improve overall (2R)-indandiol selectivity. To modulate the indene oxidation flux in KY1, two parallel approaches were formulated. First, an epoxide hydrolase capable of selectively hydrolyzing indan oxide was constitutively expressed in Rhodococcus sp. KY1. The recombinant KY1(pDS3) strain showed enantioselective hydrolysis of the intermediate (1S,2R)-indan oxide to trans-(1R,2R)-indandiol at early culture times with subsequent attenuation of epoxide hydrolase activity. It was also determined that the direction of the first order hydrolysis of indan oxide to trans- and cis-indandiol can be manipulated by changing culture pH, with primarily trans-(1R,2R)-indandiol being formed at pH > 8. Indene bioconversions with KY1(pDS3) at pH 8.6 resulted in >95% selectivity of trans-(1R,2R)-indandiol compared to approximately 60% in KY1 at pH 7.0. These approaches to directing indan oxide hydrolysis toward trans-(1R,2R)-indandiol were validated in biphasic indene bioconversions using silicon oil to solubilize the aqueously immiscible indene substrate.

Subsequent analysis of the regulatory structure of the KY1 indene bioconversion network using radiolabeled tracers revealed that the enzymatic activity catalyzing the oxidation of indene to (1S,2R)-indan oxide is inhibited in vivo by concentrations of trans-(1R,2R)-indandiol at concentrations as low as 1.0 g/L. Further experiments exploring the loss of limonene-1,2-epoxide hydrolase activity in the recombinant KY1(pDS3) strain determined that epoxide hydrolase activity in KY1(pDS3) is not affected by accumulation of indan oxide or trans-(1R,2R)-indandiol. These results suggest several courses of action for industrial implementation of a bioconversion process for (2R)-
indandiol production, including: Directed evolution studies to remove inhibitory effects on the indene monooxygenase and/or improve epoxide hydrolase turnover, product removal using hydrophobic resins, the use of high-pH bioconversion conditions, and the integration of large-scale gene expression analysis using DNA microarrays. The utility of each of these approaches is evaluated in the next chapter.
CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

An analytical and experimental framework for systems analysis and metabolic engineering of uncharacterized biocatalytic systems, in the context of a systematic evaluation of the indene bioconversion networks of *Rhodococcus* strains facilitating the improvement of (2R)-indandiol production suitable for the manufacturing of Crixivan®, was developed as follows:

(i) Establishment of an experimental system for strain selection and metabolic network analysis. A gas-phase indene delivery chemostat system was developed and used to apply selective pressure to *Rhodococcus* sp. I24 (Yanagimachi 2000). This resulted in the emergence of the KY1 strain that lacked toluene-inducible activities, specifically the ability to produce the side products 1-indenol and 1-indanone. The chemostat system facilitated indene bioconversion network analysis as described in (ii).

(ii) Definition of the bioconversion network and quantification of network fluxes. Radiolabeled tracers were used to identify the metabolic differences between the I24 and KY1 strains. In addition to lacking the ability to produce 1-indenol and its isomerization product, it was found that KY1 lacks a toluene-inducible dioxygenase present in I24 and capable of producing *cis*-(1S,2R)-indandiol directly from indene. Also, tracer analysis revealed that KY1 lacks a *trans*- (1R,2R)-indandiol dehydrogenase shown to be present at low levels in I24, and that the hydrolysis of (1S,2R)-indan oxide is non-enzymatic and results in the formation of both *cis*-(1S,2R)-indandiol and *trans*- (1R,2R)-indandiol. Subsequently, the chemostat system was used in conjunction with radiolabeled tracers to determine the metabolic flux distribution in the KY1 indene bioconversion network. This analysis revealed that the primary route of indene bioconversion is through the monooxygenase pathway, which supports at least 94% of the indene oxidation in all steady states analyzed (Yanagimachi 2000).
(iii) **Target identification.** Analysis of the metabolic fate of indan oxide in the KY1 bioconversion network revealed the non-enzymatic hydrolysis of this intermediate to both (2R)-indandiol enantiomers as a key node in the network. Consideration of the metabolic fates of each (2R)-indandiol led to the hypothesis that directing the hydrolysis of indan oxide toward the formation of trans-(1R,2R)-indandiol would most efficiently improve (2R)-indandiol selectivity, as trans-(1R,2R)-indandiol is not subject to detectable dehydrogenase activity in KY1. Furthermore, the effort required to knockout multiple dehydrogenase(s) would at best result in an enantiomeric mixture of (2R)-indandiol. This analysis identified the indan oxide hydrolysis reaction as the prime candidate for modulation with the objective of improving (2R)-indandiol selectivity.

(iv) **Flux redistribution.** Two independent approaches were conceived to direct the hydrolysis of indan oxide toward trans-(1R,2R)-indandiol. First, the limA gene from *Rhodococcus erythropolis* DCL14 encoding a limonene-1,2-epoxide hydrolase, previously shown to enzymatically hydrolyze indan oxide, was constructed via template-free PCR and placed in the expression plasmid pDS3. Importantly, parallel to this work was the development of molecular biology tools (selectable markers, transformation systems) necessary to construct recombinant strains from *Rhodococcus* sp. KY1 (Treadway 1999). Expression of limonene-1,2-epoxide hydrolase in the recombinant KY1(pDS3) strain resulted in improved trans-(1R,2R)-indandiol formation at early culture times, with attenuation of epoxide hydrolase activity approximately 25 h after indene addition in batch cultures. Independently, (1S,2R)-indan oxide hydrolysis was determined to be pH-dependent, with primarily trans-(1R,2R)-indandiol formed at high pH, and cis-(1S,2R)-indandiol favored at low pH. Implementation of high-pH conditions in both KY1 and the recombinant strain resulted in substantial improvements in trans-(1R,2R)-indandiol selectivity to in excess of 95%, in both single-phase and biphasic bioconversions.
(v) Analysis of the modified bioconversion phenotype to identify areas for further development. Radiolabeled tracer experiments were performed to determine if the monooxygenase is subject to product inhibition that may limit (2R)-indandiol titers. By adapting a method using radiolabeled tracers that was developed for in vivo metabolic flux determination, it was found that the desired product, trans-(1R,2R)-indandiol, inhibits indan oxide formation at concentrations as low as 1.0 g/L. Conventional enzymatic studies on limonene-1,2-epoxide hydrolase expressed in KY1 showed that this enzyme is not inhibited either by its substrate, indan oxide, or its product. This indicated that the attenuation of epoxide hydrolase activity observed in vivo in recombinant cultures is likely due to inherent low turnover of the enzyme. Taken together, these results suggest a number of possible courses of action for development of KY1 or KY1(pDS3) as industrial production strains for trans-(1R,2R)-indandiol, as discussed below.

5.2 Conclusions

This work has applied the concepts and tools of metabolic engineering in creating an effective combination of a new recombinant Rhodococcus strain and fermentation environment capable of complete conversion of indene to trans-(1R,2R)-indandiol product. Through two approaches, the (2R)-indandiol selectivity in Rhodococcus sp. KY1 was increased to at least 95%, as originally suggested by the KY1 flux distribution. This work has demonstrated that KY1(pDS3) is a potential production strain capable of producing trans-(1R,2R)-indandiol from indene at high selectivity. Further optimization of bioconversion conditions is likely to improve the product titer obtained in this study. In particular, the use of a two-phase system similar to that employed previously may be useful to reduce exposure of the cells to the possibly toxic and/or inhibitory indene and indene metabolites (Connors, Chartrain et al. 1997; Chartrain, Jackey et al. 1998; Buckland, Drew et al. 1999). Alternatively, the use of a gas-phase indene delivery system similar to that described here may be preferable, as it was possible to maintain residual aqueous indene concentrations near zero under certain feeding conditions. In conjunction with fed-batch operation, concerted process development work (i.e.
parameter optimization) has significant potential to increase product titers beyond those obtained in this and previous studies.

Though not the focus of this thesis, it is important to recognize the central role of the chemostat in biocatalytic strain development. A mutant strain, *Rhodococcus* sp. KY1, with a drastically improved chiral indandiol product profile was generated by applying proper selective pressure to chemostat cultures of *Rhodococcus* sp. The work of (Yanagimachi 2000) represents a novel application of the chemostat system that differs from the traditionally exclusive use of the chemostat for the study of cell physiology. The improvement observed with respect to indene bioconversion is most likely due to a stereospecific monoxygenase of high activity catalyzing the conversion of indene to indan oxide of (2R) chirality, and the elimination of a toluene-induced dioxygenase pathway. This resulted in doubling the product yield with no significant change in productivity at lower indene concentrations. Further productivity increases are obtained by increasing indene feed concentration presumably due to greater induction of the monoxygenase activity.

The framework presented here is also applicable to other bioconversion systems where maximal production of chiral intermediates is sought through bioconversion network optimization (Panke, Wubbolts et al. 2000). This study also demonstrates the importance of rigorously defining the bioconversion network and determining the magnitude of the network fluxes to rationally identify targets for selectivity improvement. Flux distributions were instrumental in identifying targets in the bioconversion network for further genetic work in the KY1 strain to eliminate the remaining side products. Furthermore, a key determining factor in prioritization of network targets is the availability of molecular biology tools and genes. Only those targets that can be modified genetically or otherwise within a reasonable timeframe, as determined by manufacturing requirements, can be given full consideration.
5.3 Recommendations for future work

Analysis of the regulation of the indene monooxygenase and limonene-1,2-epoxide hydrolase enzymes in KY1 and KY1(pDS3), the first order nature of [¹⁴C]-indene oxidation, as well as other results and scientific developments indicate a number of opportunities for process development of indene bioconversion in Rhodococcus:

(a) Directed evolution of monooxygenase and/or epoxide hydrolase enzymes. The emergence of directed evolution and gene shuffling technologies has provided the ability to rapidly improve enzyme properties through a combinatorial approach, including an improvement in the indene oxidation profile of toluene dioxygenase from Pseudomonas putida F1 (Zhang, Stewart et al. 2000). In principle, this technology could be used to remove inhibitory effects of trans-indandiol on the monooxygenase enzyme (provided the gene(s) encoding this enzyme are cloned) and to improve the turnover of the limonene-1,2-epoxide hydrolase of the recombinant KY1(pDS3) strain.

(b) Active removal of trans-indandiol (or indan oxide) product. Removal of trans-indandiol from actively metabolizing cultures would reduce inhibition of the key monooxygenase enzyme and likely improve overall product titers. Polymeric absorbent resins have frequently been used in stereoselective bioconversions (for example, the enzymatic reduction of 3,4-methylene-dioxyphenyl acetone (Vicenzi, 1997)). Of particular interest are boronate resins that preferably bind cis-diol moieties. Preliminary experiments suggested that commercial phenylaminoboronate resins have the ability to bind trans-(1R,2R)-indandiol, as well as other indene metabolites (Appendix E). Alternatively, (1S,2R)-indan oxide removal would allow ex vivo application of high-pH conditions to obtain the final product. Use of an appropriate organic phase in biphasic growth conditions may allow for partitioning of indan oxide into an environment where it is stable, while still allowing for biocatalytic oxidation of indene simultaneously.
(c) **Integration of DNA microarray technologies.** The rapid advances in this field are making gene identification and genome-wide expression analysis relatively routine. The *Rhodococcus* sp. I24 genome has been sequenced (Jefferson Parker, personal communication), which will facilitate the rapid identification of the oxygenase genes, in addition to an unknown number of others, that play an important role in indene bioconversion. In addition, microarray experiments can be used to explore the growth dependence of indene bioconversion gene expression, to provide further context to the bioconversion results obtained in this study. This will also facilitate either directed evolution improvement of the monooxygenase, and/or other genetic modifications to improve (2R)-indandiol production.

(d) **Resolve indene toxicity issues.** As suggested by both single-phase and biphasic bioconversion results, indene concentrations play a critical role in biocatalyst viability and consequently (2R)-indandiol formation. The best titers apparently were obtained when aqueous concentrations of indene were maintained near zero, although this measurement could not be made accurately in the biphasic systems. Unfortunately, gas-phase delivery of indene is unlikely to be industrially feasible due to hazards associated with the containment of the volatile substrate. Consequently, rigorous work to determine the tolerance of indene by *Rhodococcus*, or perhaps other hosts, should be undertaken in the context of the biphasic bioconversion system. This can be potentially be done using multi-channel flow cytometry and/or using DNA microarrays. In addition, transport studies to determine the partitioning behavior of indene between aqueous and selected organic phases will be needed.
The methods and results of this work have broad implications for the design of metabolic engineering approaches to biocatalyst design in so-called "uncharacterized" biological systems with poorly characterized biochemistry and genetics, particularly with respect to the role of chemostats in strain improvement programs. A framework for optimizing such systems has been presented, and will benefit from large-scale gene expression analysis studies that will facilitate rapid gene identification. DNA microarray and directed evolution technologies will integrate into the framework by presenting more options for target identification and ultimately the engineering of robust industrial biocatalysts.
REFERENCES


Boyd, D., N. Sharma and A. Smith (1982) Chemical synthesis and optical purity determination of optically active 1,2-epoxyindan and alcohol products which are also derived from mammalian or microbial metabolism of indene or indanones. *J. Chem. Soc. Perkin Trans. 1*: 2767-2770.


TIGR Microbial Database (http://www.tigr.org/tdbmdb/mdbcomplete.html).


APPENDICES

Appendix A - Nomenclature

[C2R] \( \text{cis-(1S,2R)-indandiol concentration} \)
[C2R\(^*\)] \( [^{14}\text{C}]\text{-cis-(1S,2R)-indandiol concentration} \)
[C2S\(^*\)] \( [^{14}\text{C}]\text{-cis-(1R,2S)-indandiol concentration} \)
[E\(_{\text{ACTIVE}}\)] Active enzyme concentration
[I\(_{\text{tot}}\)] Indene concentration (labeled plus unlabeled)
[I\(^*\)] \( [^{14}\text{C}]\text{-indene concentration} \)
[I\(^{0*}\)] \( [^{14}\text{C}]\text{-indan oxide concentration} \)
[K\(^*\)] \( [^{14}\text{C}]\text{-1-keto-2-hydroxy-indan concentration} \)
\( K_m^i \) Michaelis-Menten constant for enzyme \( i \)

\( k_{\text{CAT}} \) Turnover number for enzymes following Michaelis-Menten kinetics
\( k_{\text{C2R}} \) Indan oxide hydrolysis rate constant to \( \text{cis-(1S,2R)-indandiol} \)
\( k_{\text{DO}} \) Dioxygenase rate constant
\( k_i \) Rate constant for enzyme \( i \)
\( k_{\text{MO}} \) Monoxygenase rate constant
\( k_{\text{RDH}} \) \( \text{cis-(1S,2R)-indandiol dehydrogenase rate constant} \)
\( k_{\text{SDH}} \) \( \text{cis-(1R,2S)-indandiol dehydrogenase rate constant} \)

\( k_T \) Indan oxide hydrolysis rate constant to \( \text{trans-(1R,2R)-indandiol} \)
\( k_{\text{TDH}} \) \( \text{trans-(1R,2R)-indandiol dehydrogenase rate constant} \)

[I\(_{\text{tot}}\)] Total metabolite concentration (labeled plus unlabeled)
[I\(^*\)] \( [^{14}\text{C}]\text{-metabolite concentration} \)
\( r_{\text{C2R}} \) \( \text{cis-(1S,2R)-indandiol excretion rate} \)
\( r_{\text{C2S}} \) \( \text{cis-(1R,2S)-indandiol excretion rate} \)
\( r_{\text{IND}} \) Indene uptake rate
\( r_{\text{IO}} \) Indan oxide excretion rate
\( r_{\text{K}} \) 1-keto-2-hydroxy-indan excretion rate
\( r_{\text{T}} \) \( \text{trans-(1R,2R)-indandiol excretion rate} \)
[I\(^*\)] \( [^{14}\text{C}]\text{-trans-(1R,2R)-indandiol concentration} \)
$J_{C2R}$  Indan oxide hydrolysis flux to $cis$-$(1S,2R)$-indandiol

$J_{DO}$  Dioxygenase flux

$J_{i}$  Flux for enzyme $i$

$J_{max}$  Maximum specific rate for an enzyme catalyzed reaction

$J_{MO}$  Monooxygenase flux

$J_{RDH}$  $cis$-$(1S,2R)$-indandiol dehydrogenase flux

$J_{SDH}$  $cis$-$(1R,2S)$-indandiol dehydrogenase flux

$J_{T}$  Indan oxide hydrolysis flux to $trans$-$(1R,2R)$-indandiol

$X$  Biomass concentration
Appendix B – Construction of *Rhodococcus* sp. I24(pDS3)

*Rhodococcus* sp. I24 has multiple dioxygenase activities in addition to the monooxygenase activity that is the primary indene oxidative enzyme in *Rhodococcus* sp. KY1. Because of the competing activities, expression of limonene-1,2-epoxide hydrolase is less likely to result in a high *trans-(1R,2R)*-indandiol selectivity in I24 as was shown to occur in KY1 in this study, although *trans-(1R,2R)*-indandiol has been shown to accumulate as high as 1.4 g/L in I24, and 1-indenol formation is less prevalent in batch culture relative to continuous cultures (Chartrain, Jackey et al. 1998, Yanagimachi 2000). In developing the transformation protocols for *limA* expression in *Rhodococcus*, I24 was used as the test strain due to the fact that molecular biology protocols for this strain were better developed than for KY1 (Treadway 1999).

Preparation of competent *Rhodococcus* sp. I24 and subsequent transformation of I24 with pDS3 by electroporation was performed as described in Chapter 3. Figure B.1 shows Southern blot (with pAL282 used as the probe) and agarose gel analysis of restriction digests on plasmid DNA prepared from I24(pDS3) transformants. In Figure B.1a, the transformants produced bands of the same size as the plasmid controls, providing confirmation that pDS3 was transformed intact into I24. Further confirmation was provided by the Hind III digests shown in Figure B.1b.

Based upon the study of indene bioconversion in KY1 and previous results (Chartrain, Jackey et al. 1998), some improvement in *trans*-indandiol formation would be expected in I24(pDS3) relative to the wild-type I24 strain. However, due to the production of 1-indenol and 1-indanone by this strain, it is not a good candidate for *(2R)*-indandiol production via epoxide hydrolase expression. This strain would require multiple knockouts to enable it to produce an enantiomerically pure *(2R)*-indandiol product.
Figure B.1  (a) Southern blot of recombinant I24(pDS3). Lanes are as follows: (1) plasmid control, (2) HinD III digest of I24(pDS3) DNA, (3) Sac II digest of I24(pDS3) DNA. (b) Restriction digest batteries of I24(pDS3): (4) BstE III digest of λ, (5-9) Hind III and four replicates, (10) Hae III digest of φX174.
Appendix C – Biphasic indene bioconversion profiles

As discussed in Chapter 4, a series of four, 15 L biphasic bioconversions using silicon-oil as a carrier for the indene substrate were performed with KY1 and the recombinant KY1(pDS3) strain. Shown in Figures C.1-C.4 are the indene metabolite profiles for each of the runs. In each of the runs, ammonium sulfate was added periodically to facilitate cell growth, and 30 mL of indene (2 mL/L) was added in 2.0 L of silicon oil at a culture OD₆₀₀ of approximately 10. Indene and metabolite concentrations were monitored during the bioconversions by reverse-phase HPLC. Cell growth could not be accurately monitored due to the presence of silicon oil, which likely adheres to cell membranes and complicates biomass analysis. Note that at 65 h post-indene addition, an additional 60 mL indene (4 mL/L) was added to replenish substrate lost both to indene oxidation and to stripping into the off-gas.

![Graph showing metabolite profiles](image)

**Figure C.1** Indene bioconversion metabolite profiles for biphasic batch KY1 run in defined medium at pH 7.0 ± 0.1. Time is indicated as post-addition of indene to the cells. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.
Figure C.2  Indene bioconversion metabolite profiles for biphasic batch KY1 run in defined medium at pH 8.5 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.
Figure C.3  Indene bioconversion metabolite profiles for biphasic batch KY1(pDS3) run in defined medium at pH 7.0 ± 0.1. Time is indicated as post-addition of indene to the cells. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.
Figure C.4  Indene bioconversion metabolite profiles for biphasic batch KY1(pDS3) run in defined medium at pH 8.5 ± 0.1. Time is indicated as post-addition of indene to the cells. Cells were grown at pH 6.9, and the pH was increased to the final value upon initiation of indene oxidation. Pulse of 60 mL indene was added at 60 h due to substrate being stripped into the gas phase.
Appendix D – KY1 monooxygenase inhibition by 1-keto-2-hydroxy-inden

In addition to examining the effect of the desired product, trans-(1R,2R)-indandiol, on indene monooxygenation, similar experiments using radiolabeled indene were performed to explore the effect of the byproduct 1-keto-2-hydroxy-inden on the same activity. KY1 cells were removed from a chemostat culture at 0.050 h⁻¹ and 70 ppm indene feed and were challenged with multiple concentrations of 1-keto-2-hydroxy-inden, followed by addition of [¹⁴C]-indene. The steady state concentrations of indene metabolites (± 10 mg/L) in the chemostat were 50 mg/L cis-indandiol, 100 mg/L trans-indandiol, and 100 mg/L 1-keto-2-hydroxy-inden.

For the control condition where no 1-keto-2-hydroxy-inden was added to the culture, an observed monooxygenase rate constant $k_{MO} = 2.2 ± 0.1$ h⁻¹ was determined. The addition of 1-keto-2-hydroxy-inden to indene-induced KY1 cells resulted in a significant reduction in the observed rate constant fit for the monooxygenase enzyme (Figure D.1). This can be interpreted as a reduction in concentration of active monooxygenase enzyme with the addition of 1-keto-2-hydroxy-inden to induced KY1 cells. Thus, the addition of this degradation product resulted in an approximately 45% concentration reduction of active monooxygenase. Higher concentrations of 1-keto-2-hydroxy-inden resulted in a further increase in monooxygenase inhibition up to 65% at 4.0 g/l.

This data has significant implications for wild-type KY1 bioconversions, where 1-keto-2-hydroxy-inden accumulates to a significant level at pH 7.0. This data provides additional support for focusing on the indan oxide node to improve (2R)-indandiol yield and productivity in the KY1 strain. However, independently of the approaches taken in this thesis work, removal of the byproduct by use of polymeric resins or the application of directed evolution to the monooxygenase may be useful in increasing product titers in KY1 by reducing byproduct inhibition effects.
Figure D.1  Inhibition of indene monooxygenase activity by 1-keto-2-hydroxy-indan.
Appendix E – Resins for removal of indene metabolites

Removal of trans-indandiol product from KY1 or KY1(pDS3) indene bioconversions could potentially reduce or eliminate product inhibition of the monooxygenase enzyme. One possible approach is the use of boronate resins commonly used for separations of molecules containing cis-diol moieties. Applications of these resins have been reviewed previously (Mazzeo and Krull 1989). Commercially available boronate is in the form of phenylaminoboronate immobilized to one of a number of support structures, including silica, agarose, and polyacrylamide. Phenylaminoboronate binds cis-diols through formation of a cyclic ester under basic conditions, as shown in Figure E.1:

![Diagram showing proposed binding mechanism of cis-(1S,2R)-indandiol to phenylaminoboronate. The hydroxylation of boronate to form a tetrahedral boronate intermediate is favored under basic conditions. The cyclic ester is hydrolyzed using 25 mM HCl. X is the polyacrylamide bead.]

Figure E.1 Proposed binding mechanism of cis-(1S,2R)-indandiol to phenylaminoboronate. The hydroxylation of boronate to form a tetrahedral boronate intermediate is favored under basic conditions. The cyclic ester is hydrolyzed using 25 mM HCl. X is the polyacrylamide bead.

Secondary interactions important to esterification include hydrophobic effect and ionic interactions. The presence of NH₄⁺ favors the formation of the boronate tetrahedral intermediate.
To test the utility of phenylaminoborinate resins in indene bioconversions, supernatant from *Rhodococcus* sp. KY1 shake flask indene bioconversions was applied to a 6.2 cm³ phenylaminoborinate column (Pierce) equilibrated with 0.2 M ammonium acetate (per manufacturer recommendations). The supernatant was adjusted to pH 7.2, and was supplemented with stock *cis*(1S,2R)-indandiol prior to addition to the column. The concentration of indene metabolites in the broth applied to the culture was 230 mg/L *cis*-indandiol (chirality undetermined), 200 mg/L *trans*-indandiol, and 410 mg/L 1-keto-2-hydroxy indan (± 10 mg/L). Indene was present in very low concentrations, and was found in previous experiments not to bind the resin.

The boronate resin acted non-specifically to bind both *cis*- and *trans*-indandiol, as well as 1-keto-2-hydroxy indan. The total amount of indene metabolites bound to the resin in this experiment were as follows: 2.1 mg *cis*-indandiol, 1.3 mg *trans*-indandiol, and 5.9 mg 1-keto-2-hydroxy indan. All three compounds broke through the column nearly simultaneously. The lack of specificity that was exhibited could be due to hydrogen bonding interactions between the hydroxyl groups of the boronate and the indandiols and 1-keto-2-hydroxy-indan, though some esterification between the boronate and indandiols presumably occurs. Though the resin was not selective with respect to specific indene metabolites, it is a possible method of preventing degradation of product and/or reducing product inhibition as it appears to bind all indene metabolites that were studied.
Appendix F – Plasmid maps

Plasmid name: pDS1
Plasmid size: 4418 bp
Constructed by: D. Stafford, P. Lessard
Construction date: 4 February 2000
Comments/References: synthetic limA gene based on sequence of R. erythropolis
limA cloned by overlapping primers and PCR, in pCR2.1-TOPO; pDS1 created by ligating Clai-XbaI fragment from plimA8 into Clai-SpeI sites of plimA2 to correct errors in plimA2
### DNA Strider 1.0 ###

**pDS1 seq -> All Sites**

**DNA sequence**  |  **4415 b.p.** |  **AGGCCCCAACATGC**  |  **AGGAAGCGGGAAG**  |  **circular**
--- | --- | --- | --- | ---

<table>
<thead>
<tr>
<th>Restriction</th>
<th>Sequence</th>
<th>Start Position</th>
<th>End Position</th>
<th>Affinity</th>
<th>Other Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apa I</td>
<td>gggccc</td>
<td>1</td>
<td>859</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>ApaI</td>
<td>g/tgacc</td>
<td>1</td>
<td>2401</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Bal I</td>
<td>tgtgcca</td>
<td>1</td>
<td>2011</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>BamHI</td>
<td>g/gatcc</td>
<td>1</td>
<td>252</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Bcl I</td>
<td>tgtgctc</td>
<td>1</td>
<td>1772</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Bgl II</td>
<td>a/catgt</td>
<td>1</td>
<td>4290</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>BspE I</td>
<td>t/ccgga</td>
<td>1</td>
<td>7271</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>BssH II</td>
<td>g/ccggc</td>
<td>1</td>
<td>2328</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>BstB I</td>
<td>tt/cgaa</td>
<td>1</td>
<td>3684</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Cla I</td>
<td>at/cgtt</td>
<td>1</td>
<td>545</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Dra III</td>
<td>ccmnn/gtg</td>
<td>1</td>
<td>1283</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>EcoRI</td>
<td>ccttn/nmnaggg</td>
<td>1</td>
<td>7951</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>HinC II</td>
<td>gty/rac</td>
<td>1</td>
<td>3561</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>HinD III</td>
<td>a/agcctt</td>
<td>1</td>
<td>234</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Pvu II</td>
<td>gttac/c</td>
<td>1</td>
<td>240</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Nco I</td>
<td>c/catgg</td>
<td>1</td>
<td>2363</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Sac II</td>
<td>gacgt/c</td>
<td>1</td>
<td>2461</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Sac II</td>
<td>ccggc/gg</td>
<td>1</td>
<td>342</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Sal I</td>
<td>g/ccggc</td>
<td>1</td>
<td>356</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Xba I</td>
<td>t/ccgag</td>
<td>1</td>
<td>853</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>Xmr I</td>
<td>gaann/nnttc</td>
<td>1</td>
<td>2798</td>
<td>4415</td>
<td>1</td>
</tr>
<tr>
<td>AlwII</td>
<td>cagnn/nctg</td>
<td>2</td>
<td>1550</td>
<td>2326</td>
<td>1</td>
</tr>
<tr>
<td>AlwI</td>
<td>g/tgacc</td>
<td>2</td>
<td>2730</td>
<td>1246</td>
<td>2</td>
</tr>
<tr>
<td>BbvII</td>
<td>gaagac</td>
<td>2/6</td>
<td>428</td>
<td>210</td>
<td>2</td>
</tr>
<tr>
<td>Bgl II</td>
<td>gc/mnnn/nggc</td>
<td>2</td>
<td>1037</td>
<td>2241</td>
<td>1</td>
</tr>
<tr>
<td>BspHI</td>
<td>t/catgt</td>
<td>2</td>
<td>553</td>
<td>3017</td>
<td>1</td>
</tr>
<tr>
<td>BstXI</td>
<td>cc/mnnn/nttg</td>
<td>2</td>
<td>282</td>
<td>541</td>
<td>2</td>
</tr>
<tr>
<td>EcoRV</td>
<td>gat/ctc</td>
<td>2</td>
<td>293</td>
<td>525</td>
<td>2</td>
</tr>
<tr>
<td>Nar I</td>
<td>gg/ccggc</td>
<td>2</td>
<td>1682</td>
<td>248</td>
<td>2</td>
</tr>
<tr>
<td>Not I</td>
<td>gc/gccggc</td>
<td>2</td>
<td>275</td>
<td>559</td>
<td>2</td>
</tr>
<tr>
<td>Nsi I</td>
<td>atgca/t</td>
<td>2</td>
<td>263</td>
<td>585</td>
<td>2</td>
</tr>
<tr>
<td>Pvu II</td>
<td>cggac/c</td>
<td>2</td>
<td>101</td>
<td>2019</td>
<td>2</td>
</tr>
<tr>
<td>ScI</td>
<td>agt/act</td>
<td>2</td>
<td>649</td>
<td>2270</td>
<td>1</td>
</tr>
<tr>
<td>TcH11</td>
<td>gacnc/nnttc</td>
<td>2</td>
<td>587</td>
<td>1459</td>
<td>2</td>
</tr>
<tr>
<td>Ase I</td>
<td>at/ctat</td>
<td>3</td>
<td>46</td>
<td>59</td>
<td>3</td>
</tr>
<tr>
<td>BspHI</td>
<td>acgtgc</td>
<td>4/8</td>
<td>1598</td>
<td>2229</td>
<td>3</td>
</tr>
<tr>
<td>Dra I</td>
<td>tttaa</td>
<td>3</td>
<td>2822</td>
<td>692</td>
<td>2</td>
</tr>
<tr>
<td>EcoRI</td>
<td>g/aaccc</td>
<td>3</td>
<td>303</td>
<td>68</td>
<td>3</td>
</tr>
<tr>
<td>Fsp I</td>
<td>tgg/cga</td>
<td>3</td>
<td>1031</td>
<td>1000</td>
<td>3</td>
</tr>
<tr>
<td>Nae I</td>
<td>gcc/ggc</td>
<td>3</td>
<td>734</td>
<td>446</td>
<td>3</td>
</tr>
<tr>
<td>PaeK7</td>
<td>c/tgg</td>
<td>3</td>
<td>270</td>
<td>246</td>
<td>3</td>
</tr>
<tr>
<td>Pat I</td>
<td>ctggc</td>
<td>3</td>
<td>298</td>
<td>515</td>
<td>3</td>
</tr>
<tr>
<td>Sph I</td>
<td>gcagc/c</td>
<td>3</td>
<td>265</td>
<td>581</td>
<td>3</td>
</tr>
<tr>
<td>Xho I</td>
<td>c/tgcgg</td>
<td>3</td>
<td>270</td>
<td>246</td>
<td>3</td>
</tr>
<tr>
<td>BsaHI</td>
<td>gr/cgc</td>
<td>4</td>
<td>590</td>
<td>1292</td>
<td>2</td>
</tr>
<tr>
<td>Eag I</td>
<td>c/gccggc</td>
<td>4</td>
<td>276</td>
<td>321</td>
<td>3</td>
</tr>
<tr>
<td>Pvu II</td>
<td>cag/cgt</td>
<td>4</td>
<td>53</td>
<td>928</td>
<td>2</td>
</tr>
</tbody>
</table>

---

89 sites found

179
Plasmid name: pDS2
Plasmid size: 4418 bp
Constructed by: D. Stafford, P. Lessard
Construction date: 15 February 2000
Comments/References: synthetic limA gene based on sequence of R. erythropolis limA; pDS2 derived by substituting the Ncol-BspEI fragment from pDS1 with the corresponding fragment from plimA9 (see plimAb) to correct point mutations in pDS1
pDS2 seq --> All Sites

<table>
<thead>
<tr>
<th>DNA sequence</th>
<th>4418 b.p.</th>
<th>AGCGCCCAATAC ... AGGAAACGGAAG</th>
<th>circular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apa I</td>
<td>gggcc/c</td>
<td>1  862(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Asp718</td>
<td>g/tgacc</td>
<td>1  240(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Bal I</td>
<td>tgg/cca</td>
<td>1  2014(4418) 1</td>
<td></td>
</tr>
<tr>
<td>BamH I</td>
<td>g/gatcc</td>
<td>1  252(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Bcl I</td>
<td>t/gatca</td>
<td>1  1775(4418) 1</td>
<td></td>
</tr>
<tr>
<td>BspLU11 I</td>
<td>a/catgt</td>
<td>1  4293(4418) 1</td>
<td></td>
</tr>
<tr>
<td>BspE I</td>
<td>t/ccgga</td>
<td>1  717(4418) 1</td>
<td></td>
</tr>
<tr>
<td>BssH II</td>
<td>g/ccgcc</td>
<td>1  2331(4418) 1</td>
<td></td>
</tr>
<tr>
<td>BstB I</td>
<td>tt/cgaa</td>
<td>1  368(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Cla I</td>
<td>at/cgat</td>
<td>1  545(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Dra III</td>
<td>cacnnm/gtg</td>
<td>1  1286(4418) 1</td>
<td></td>
</tr>
<tr>
<td>EcoN I</td>
<td>cctnn/mnnagg</td>
<td>1  798(4418) 1</td>
<td></td>
</tr>
<tr>
<td>HinC II</td>
<td>gty/rac</td>
<td>1  356(4418) 1</td>
<td></td>
</tr>
<tr>
<td>HincD III</td>
<td>a/agctt</td>
<td>1  234(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Kpn I</td>
<td>ggtac/c</td>
<td>1  240(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Nco I</td>
<td>c/cattg</td>
<td>1  2366(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Sac I</td>
<td>gacct/c</td>
<td>1  246(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Sac II</td>
<td>ccgcc/gg</td>
<td>1  342(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Sal I</td>
<td>g/tcgac</td>
<td>1  356(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Xba I</td>
<td>t/cctga</td>
<td>1  856(4418) 1</td>
<td></td>
</tr>
<tr>
<td>Xmi I</td>
<td>gaann/mnttc</td>
<td>1  2801(4418) 1</td>
<td></td>
</tr>
<tr>
<td>AlwN I</td>
<td>cagnnn/ctg</td>
<td>2  1553(2326) 1</td>
<td></td>
</tr>
<tr>
<td>Apal I</td>
<td>g/ctcac</td>
<td>2  2733(1246) 2</td>
<td></td>
</tr>
<tr>
<td>Bbv I</td>
<td>gaagac 2/6</td>
<td>2  428(210) 2</td>
<td></td>
</tr>
<tr>
<td>Bgl I</td>
<td>gccnnnn/nnggc</td>
<td>2  1040(2241) 1</td>
<td></td>
</tr>
<tr>
<td>BspH I</td>
<td>t/catga</td>
<td>2  553(3020) 1</td>
<td></td>
</tr>
<tr>
<td>BetX I</td>
<td>ccannnn/ntgg</td>
<td>2  282(544) 2</td>
<td></td>
</tr>
<tr>
<td>EcoR V</td>
<td>gat/atcc</td>
<td>2  293(528) 2</td>
<td></td>
</tr>
<tr>
<td>Nar I</td>
<td>gg/cgcc</td>
<td>2  1685(248) 2</td>
<td></td>
</tr>
<tr>
<td>Not I</td>
<td>gc/ggcccc</td>
<td>2  275(562) 2</td>
<td></td>
</tr>
<tr>
<td>Nsi I</td>
<td>atgca/t</td>
<td>2  263(588) 2</td>
<td></td>
</tr>
<tr>
<td>Pvu I</td>
<td>cgg/cg</td>
<td>2  1014(2019) 2</td>
<td></td>
</tr>
<tr>
<td>Sca I</td>
<td>agt/act</td>
<td>2  649(2273) 1</td>
<td></td>
</tr>
<tr>
<td>Tthl11 I</td>
<td>gacnn/mnttc</td>
<td>2  587(1462) 2</td>
<td></td>
</tr>
<tr>
<td>Ase I</td>
<td>at/taat</td>
<td>3  46(59) 3</td>
<td></td>
</tr>
<tr>
<td>BspM I</td>
<td>acctgcc 4/8</td>
<td>3  1601(229) 3</td>
<td></td>
</tr>
<tr>
<td>Dra I</td>
<td>ttt/aaa</td>
<td>3  2825(692) 2</td>
<td></td>
</tr>
<tr>
<td>EcoR I</td>
<td>g/aattc</td>
<td>3  301(68) 3</td>
<td></td>
</tr>
<tr>
<td>Fsp I</td>
<td>tgt/gca</td>
<td>3  1034(1000) 3</td>
<td></td>
</tr>
<tr>
<td>Nae I</td>
<td>gcc/ggc</td>
<td>3  734(449) 3</td>
<td></td>
</tr>
<tr>
<td>NgOM IV</td>
<td>g/ccgcc</td>
<td>3  734(449) 3</td>
<td></td>
</tr>
<tr>
<td>PaeR7 I</td>
<td>c/tcgag</td>
<td>3  270(246) 3</td>
<td></td>
</tr>
<tr>
<td>Pst I</td>
<td>ctgca/g</td>
<td>3  258(518) 3</td>
<td></td>
</tr>
<tr>
<td>Sph I</td>
<td>gctgc/c</td>
<td>3  265(584) 3</td>
<td></td>
</tr>
<tr>
<td>Xho I</td>
<td>c/tcgag</td>
<td>3  270(246) 3</td>
<td></td>
</tr>
<tr>
<td>Eag I</td>
<td>c/ggcccc</td>
<td>4  276(321) 3</td>
<td></td>
</tr>
<tr>
<td>Pvu II</td>
<td>cag/ctg</td>
<td>4  53(931) 2</td>
<td></td>
</tr>
<tr>
<td>BsaH I</td>
<td>gr/cgcc</td>
<td>5  590(163) 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>4418 b.p.</th>
<th>AGCGCCCAATAC ... AGGAAACGGAAG</th>
<th>circular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1  862(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  240(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  2014(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  252(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  1775(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  4293(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  717(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  2331(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  368(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  545(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  1286(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  798(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  356(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  234(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  240(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  2366(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  246(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  342(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  356(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  856(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1  2801(4418) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  1553(2326) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  2733(1246) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  428(210) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  1040(2241) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  553(3020) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  282(544) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  293(528) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  1685(248) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  275(562) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  263(588) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  1014(2019) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  649(2273) 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2  587(1462) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  46(59) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  1601(229) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  2825(692) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  301(68) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  1034(1000) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  734(449) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  734(449) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  270(246) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  258(518) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  265(584) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3  270(246) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4  276(321) 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4  53(931) 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5  590(163) 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

93 sites found
Plasmid name: pDS3
Plasmid size: 4437 bp
Constructed by: P. Lessard
Construction date: 22 February 2000
Comments/References: limA ORF under control of lambda PL promoter; derived by ligating the EcoRV-SphI fragment (containing limA) from pDS2 into the Scal and SphI sites of pAL245
### DNA Strider 1.0

#### Wednesday, February 23, 2000 1:44:07 PM

**pB3 seq -> All Sites**

<table>
<thead>
<tr>
<th>DNA sequence</th>
<th>4437 b.p.</th>
<th>gaattccgga ... cataacctcacc</th>
<th>circular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afl II</td>
<td>c/ttaag</td>
<td>1 3736 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Apal I</td>
<td>g/tgcac</td>
<td>1 1301 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>AspT18</td>
<td>g/tgacc</td>
<td>1 3136 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Bgl II</td>
<td>a/gatct</td>
<td>1 2002 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>BspLU11 I</td>
<td>a/catgct</td>
<td>1 1773 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>BspH II</td>
<td>g/ccggc</td>
<td>1 2342 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>BstB II</td>
<td>tt/cgaa</td>
<td>1 1049 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>BstE II</td>
<td>g/tgacc</td>
<td>1 3148 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>BstX II</td>
<td>ccannnn/nntgg</td>
<td>1 1129 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Cla I</td>
<td>at/cgat</td>
<td>1 872 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Dra I</td>
<td>ttt/aaa</td>
<td>1 2989 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Dra II</td>
<td>ccannnn/gtg</td>
<td>1 1438 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Hpa I</td>
<td>gtt/sac</td>
<td>1 46 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Kpn I</td>
<td>ggtac/c</td>
<td>1 3136 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Mlu I</td>
<td>a/cgcgt</td>
<td>1 3334 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Nco I</td>
<td>c/catgg</td>
<td>1 3884 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>PflM I</td>
<td>ccannnn/nntgg</td>
<td>1 3879 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Pvu II</td>
<td>cag/cgtg</td>
<td>2 2618 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Sal I</td>
<td>g/tcggac</td>
<td>1 1061 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Sca I</td>
<td>agt/act</td>
<td>1 768 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Spe I</td>
<td>a/ctagt</td>
<td>1 337 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Sph I</td>
<td>gcata/c</td>
<td>1 1152 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Spl I</td>
<td>c/gtacg</td>
<td>1 3170 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Xba I</td>
<td>t/tcagag</td>
<td>1 331 (4437) 1</td>
<td></td>
</tr>
<tr>
<td>Xmn I</td>
<td>g/nnttc/mttc</td>
<td>2 1218 (4437) 1</td>
<td></td>
</tr>
</tbody>
</table>

| AlwII  | cagmn/mctg | 2 1168 (1744) 2               | 2912 (2693) 1 |
| Bal I   | tgg/cca    | 2 3744 (242) 2                | 3906 (4195) 1 |
| BamHI   | g/gatcc    | 2 10 (333) 2                 | 343 (4104) 1 |
| BspH I  | acc/gac    | 2 2670 (810) 2               | 3480 (3627) 1 |
| EcoI    | ccctn/mnnag | 2 499 (115) 2               | 614 (3422) 1 |
| EcoR I  | gat/atc    | 2 1124 (1072) 2              | 2196 (3365) 1 |
| Nde I   | ca/tatg    | 2 2647 (614) 2               | 3261 (3823) 1 |
| Not I   | gc/ggctcg | 2 1140 (1147) 2              | 2287 (3290) 1 |
| PaeK I  | c/tcggag   | 2 901 (246) 2                | 1147 (4191) 1 |
| Pst I   | ctgca/g    | 2 601 (518) 2                | 1119 (3919) 1 |
| SnaB I  | tac/gta    | 2 2633 (28) 2                | 2661 (4409) 1 |
| Stu I   | agg/cct    | 2 288 (2708) 1               | 2996 (1729) 2 |
| Xho I   | c/tcggag   | 2 901 (246) 2                | 1147 (4191) 1 |

| Bbv II  | gaagac 2/6 | 3 779 (210) 3               | 989 (280) 2 |
| BsaHI   | gr/ccgcc   | 3 664 (163) 3               | 827 (2969) 1 |
| BspH I  | t/catga    | 3 281 (583) 3               | 864 (1961) 1 |
| Bsp8 I  | t/ccgga    | 3 576 (124) 3               | 700 (3566) 1 |
| Hind III | a/agctt   | 3 122 (3194) 1             | 3136 (759) 2 |
| Nae I   | gcc/ggcc   | 3 296 (26) 3                | 322 (361) 2 |
| NcoI IV  | g/ccggc    | 3 296 (26) 3                | 322 (361) 2 |
| Sac II  | ccgg/cggc  | 3 1075 (1179) 2             | 2254 (31) 3 |
| SgrA I  | cr/ccgcgg  | 3 682 (2032) 1              | 2714 (916) 3 |
| TchII11 I | gcc/nngtcc | 3 827 (964) 3               | 1791 (1483) 2 |

| Bgl I   | gccmn/nnggc | 4 355 (899) 4           | 1254 (1137) 2 |
| Bsm I   | gaatgc 1/-1 | 4 513 (2040) 1          | 2553 (297) 4 |
| EcoR7 III | age/gctc | 4 2297 (66) 4          | 2361 (45) 3  |
| EcoR I  | g/aatcc   | 4 1 (605) 2            | 606 (440) 3  |
| HinC II | gry/rac    | 4 46 (1015) 3          | 1061 (410) 4 |
| Fsp I   | tgc/gca    | 5 2600 (244) 4         | 2844 (9) 5 |
| Eag I   | c/ccggc    | 6 324 (496) 3         | 820 (321) 4  |
| BseR I  | r/ccgyy    | 7 296 (26) 7          | 322 (361) 4  |

119 sites found

183