Metabolic Engineering to Develop a Pathway for the Selective Cleavage of Carbon-Nitrogen Bonds

FINAL TECHNICAL REPORT

(October 2002-March 2006)

Principal Investigator: John J. Kilbane II

847-768-0723, john.kilbane@gastechnology.org

Report Issue Date: April 2006

DOE Contract No.: DE-FC26-02NT15382

Submitted by

GAS TECHNOLOGY INSTITUTE

1700 South Mount Prospect Road

Des Plaines, Illinois 60018

GTI Project No. 15334

Submitted to

FETC AAD Document Control

U.S. Department of Energy

Federal Energy Technology Center

P.O. Box 10940, MS 921-143

Pittsburgh, PA 15236-0940

DOE Technical Project Manager: Kathy Stirling

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

ABSTRACT

The objective of the project was to develop a biochemical pathway for the selective cleavage of C-N bonds in molecules found in petroleum. Specifically, the development of a novel biochemical pathway for the selective cleavage of C-N bonds in carbazole was the focus of research in this project.

The cleavage of the first C-N bond in carbazole is accomplished by the enzyme carbazole dioxygenase, that catalyzes the conversion of carbazole to 2-aminobiphenyl-2,3-diol. The genes encoding carbazole dioxygenase were cloned from *Sphingomonas* sp. GTIN11 and from *Pseudomonas resinovorans* CA10. Obtaining an enzyme capable of selectively cleaving the C-N bond in 2-aminobiphenyl-2,3-diol was the focus of much of the research in this project, however; no suitable enzyme was found.

Project accomplishments included expressing the genes for carbazole dioxygenase in *Rhodococcus erythropolis* and *Escherichia coli*, development of gene expression vectors for *Rhodococcus*, and isolation of a *Pseudomonas* sp. strain GTIN-G4 that has the novel biochemical ability to replace one of the nitrogen-associated hydrogen atoms in 2-aminobiphenyl with formaldehyde. *Rhodococcus* cultures are capable of metabolizing a wide range of substrates, including hydrophobic substrates like petroleum, and may find widespread use in the development of biotechnology processes in the future. The results of this project will directly benefit the development of future biotechnology processes/projects employing *Rhodococcus* hosts.

Three manuscripts were prepared and submitted for publication based on the data obtained in this project:

"Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4"

"Screening and Analysis of DNA Fragments that Show Promoter Activities in *Rhodococcus erythropolis*"

"Microbial Biocatalyst Developments to Upgrade Fossil Fuels".

TABLE OF CONTENTS

EXECUTIVE SUMMARY **5**
INTRODUCTION **8**
MATERIALS AND METHODS **14**
Bacterial Cultures and Growth Conditions 14
Nitrogen Bioavailability Assay 16
CarA resting cell assay 18
Thin Layer Chromatography for Identification of Metabolites 19
Gas Chromatography-Mass Spectrometry 20
High Performance Liquid Chromatography 20
Gibb's Reagent solid phase screening 21
β-Galactosidase Assay 21
Auto-oxidation 22
Genetic Techniques 22
Electroporation of *Rhodococcus erythropolis* 22
Construction of a Genomic DNA Library and Screening 23
PCR Primers 23
Error prone PCR 24
Reverse Transcription-Quantitative Polymerase Chain Reaction (RT-qPCR) for the Quantification of Gene Expression 25
DNA Sequencing 27
Primer Extension Analysis 27
In vitro Runoff Transcription 28
RNase Protection Assay 28
RESULTS AND DISCUSSION **30**
Enrichment Culture Experiments 30
Isolation of 2-ABP Degrading Bacterial Culture 34
Gibbs Assay to Investigate Substrate Range for C-N Bond Cleaving Enzymes 44
Directed Evolution of a Deaminase/Amidase Gene 45
Substrate specificity of aniline dioxygenase 55
Homology modeling 57
Genetic Studies 58
Construction of a Derivative of *Sphingomonas* GTIN11 that Lacks CarB and CarC Enzymes 63
Genetic Transformation of *Rhodococcus, Sphingomonas,* and *Pseudomonas* 65
Expression of CarA and AtdA in *Rhodococcus* 65
Investigation of *Rhodococcus* Promoters 72
CONCLUSIONS AND PROJECT IMPLICATIONS **84**
APPENDIX: Publications Resulting from This Project **87**

FIGURE 1. CARBAZOLE DEGRADATION PATHWAYS. 11

FIGURE 2. STRUCTURE OF ORGANONITROGEN COMPOUNDS USED IN GROWTH TESTS WITH PSEUDOMONAS SP. GTIN-G4 35

FIGURE 3. MASS SPECTRA OF 2-AMINOBIPHENYL AND A METABOLITE PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4. 42

FIGURE 4. MASS SPECTRA OF O-TOLUIDINE AND DIPHENYLAMINE AND THE CORRESPONDING METABOLITES PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4. 43

FIGURE 5. OPTICAL DENSITY OF CELLS IN MINIMAL MEDIA WITH BENZAMIDE AS SOLE NITROGEN SOURCE..... 46

FIGURE 6. OPTICAL DENSITY OF CELLS WITH TIME IN 2-AMINOBIPHENYL MINIMAL GROWTH MEDIUM. 47

FIGURE 7. DEAMINATION AND OXYGENATION OF (A) ANILINE AND (B) O-TOLUIDINE BY ATDA. 48

FIGURE 8. MINIMAL MEDIA ASSAY OF PAS93 TRANSFORMED *E. COLI* DH5α WITH (A) 2-ABP, (B) 2-EA AND MA. (C) CHEMICAL STRUCTURES OF 2-EA, MA AND 2-ABP. 49

FIGURE 9. ALIGNMENT OF ATD A3 WITH NAPTHTALENE DIOXYGENASE (1NDO_E). 50

FIGURE 10. SDS-PAGE ANALYSIS OF ATDA EXPRESSION BY PTRCA-2 AND PAS93 IN BL21. (A) SOLUBLE FRACTION (B) INSOLUBLE FRACTION. 51

FIGURE 11. SDS-PAGE ANALYSIS OF ATDA EXPRESSION BY PTA1-1 AND PTRCA-2 IN *E. COLI* JM109 INDUCED WITH 1 MM IPTG. (A) SOLUBLE FRACTION (B) INSOLUBLE FRACTION. 52

FIGURE 12. PLASMID MAPS OF PTA1-1, PTA2-3, AND PTA3-1. 53

FIGURE 13. NYLON MEMBRANE WITH A) JM109/PTRC99A (NEGATIVE CONTROL) CELLS; B) JM109/PTRCA-2 AFTER INCUBATION ON GIBB'S REAGENT PLATE. ... 53

FIGURE 15. E.COLI JM109 WITH PTA3-1 OR PTA2-3 AFTER 5 HRS OF INCUBATION WITH (A) 2-METHYLANILINE, AND (B) 2-ETHYLANILINE. 55

FIGURE 16. E.COLI JM109/PTRCA-2 COLONIES AFTER INCUBATION WITH VARIOUS SUBSTRATES AND GIBB'S REAGENT: RANDOM MUTAGENESIS OF *ATDA3* 56

FIGURE 17. (A) HOMOLOGY MODEL OF ATDA3, WITH 2-EA DOCKED TO THE ACTIVE SITE. RESIDUES IN RED FORM THE BINDING POCKET; (B) CLOSE UP OF ATDA3 BINDING POCKET AND THE SUBSTRATE CHANNEL. 58

FIGURE 18. MAPS OF THE CARBAZOLE DEGRADATION OPERONS OF SPHINGOMONAS SP. GTIN11, AND PSEUDOMONAS RESINOVORANS CA10 60

FIGURE 19. CONSTRUCTION OF *E. COLI* EXPRESSION VECTORS CONTAINING THE *CARA* GENES FROM *SPHINGOMONAS* SP. GTIN11 AND FROM *PSEUDOMONAS RESINOVORANS* CA10, AND A PROTEIN GEL ILLUSTRATING THE PRODUCTION OF THE CARA PROTEINS..... 62

FIGURE 20. CONSTRUCTION OF PACYCDUET-1-CARDO-GTIN11 THAT CONTAINS THE *SPHINGOMONAS* SP. GTIN11 *CARA* GENES IN AN *E. COLI* EXPRESSION VECTOR. 63

FIGURE 21. STRATEGY FOR THE REPLACEMENT OF THE *CARBA*, *CARBB*, AND *CARC* GENES OF SPHINGOMONAS GTIN11 WITH A GENE ENCODING KANAMYCIN RESISTANCE..... 64

FIGURE 22. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR PEBC26. 67

FIGURE 23. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR PSRKGFP. 68

FIGURE 24. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR PRESQ..... 69

FIGURE 25. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR PREP45BOXBNSIKAN. .. 71

FIGURE 27. NUCLEOTIDE SEQUENCES OF THE *E. COLI LACZ* PROMOTER AND DERIVATIVE PROMOTERS..... 75

FIGURE 28. RT-QPCR DATA FOR LUCIFERASE EXTERNAL MRNA CONTROL IN E. COLI 79

FIGURE 29. RT-QPCR DATA FOR EXPRESSION OF KANAMYCIN IN E. COLI BY PROMOTER VARIATIONS 80

FIGURE 30. RT-QPCR DATA FOR EXPRESSION OF KANAMYCIN IN R. ERYTHROPOLIS BY PROMOTER VARIATIONS 82

EXECUTIVE SUMMARY

The objective of the project was to develop a biochemical pathway for the selective cleavage of C-N bonds in molecules found in petroleum. Specifically a novel biochemical pathway will be developed for the selective cleavage of C-N bonds in carbazole. This is the Final technical report in a 3-year research project and encompasses the period from October, 2002 to March, 2006.

The development of biocatalysts with improved ability to cleave carbon-nitrogen bonds was addressed by cloning the genes for carbazole dioxygenase, aniline dioxygenase, and other enzymes. The removal of nitrogen from aromatic compounds like carbazole requires the cleavage of two carbon-nitrogen bonds. The cleavage of the first carbon-nitrogen bond is accomplished by the enzyme carbazole dioxygenase while an enzyme capable of selectively cleaving the second carbon-nitrogen bond has not yet been identified. The enzyme carbazole dioxygenase is encoded for by three genes: *carAa, carAc*, and *carAd*. The *carAacd* genes have been cloned and sequenced from several different microbial cultures including *Sphingomonas* sp. GTIN11 and *Pseudomonas resinovorans* CA10. Enrichment culture experiments and directed evolution experiments were performed to obtain an enzyme that can selectively cleave the second carbon-nitrogen bond in carbazole.

Enrichment culture experiments designed to isolate a culture capable of cleaving the carbon-nitrogen bond in 2-aminobiphenyl, and thus being capable of providing an enzyme for the cleavage of the second carbon-nitrogen bond in carbazole, resulted in the isolation of a unique microbial culture *Pseudomonas* sp. GTIN-G4. *Pseudomonas* sp. GTIN-G4 is capable of metabolizing 2-aminobiphenyl and related compounds, but does not appear to be capable of cleaving the carbon-nitrogen bond. Instead it has the unprecedented ability to modify 2-aminobiphenyl by replacing a hydrogen bound to the nitrogen atom with a formaldehyde group. Other enrichment culture experiments resulted in the isolation of numerous bacterial cultures capable of metabolizing a variety of organonitrogen compounds such as 2-aminobiphenyl, aniline, 4,4'-azodianiline, benzamide, carbazole, methylbenzylamine, piperidine, pyridine, pyrolidine, quinoline, quinazoline, ortho-, meta-, and para-toluidine, and triazine. However, none of these cultures had the ability to selectively cleave the C-N bond in 2-aminobiphenyl or related molecules.

Directed evolution experiments were performed by scientists at the university of Illinois at Chicago. Genes that encode the benzamide amidase from *Rhodococcus* sp. MP50, the melamine deaminase from *Pseudomonas* sp. NRRL B12227, and the aniline dioxygenase from *Acinetobacter* strain YAA were all included in directed evolution experiments designed to obtain derivative enzymes with altered substrate ranges. No enzyme capable of selectively cleaving the C-N bond in 2-aminobiphenyl was obtained.

Research performed in this project also included gene expression studies in *Rhodococcus erythropolis*. The genes that encode carbazole dioxygenase from *Sphingomonas* sp. GTIN11 and *Pseudomonas resinovorans* CA10 (*carAacd*), aniline dioxygenase from *Acinetobacter* sp. YAA (*atdA1, atdA2, atdA3, atdA4*, and *atdA5*), and resistance to the antibiotic kanamycin were cloned and introduced into *Rhodococcus erythropolis*, a bacterial host that tolerates exposure to petroleum. It was hoped that by simultaneously expressing the genes for carbazole dioxygenase and aniline dioxygenase in the same bacterial host then enrichment culture experiments employing carbazole as a sole nitrogen source may result in the isolation of derivative cultures with altered abilities to cleave C-N bonds. However, no culture capable of selectively cleaving both C-N bonds in carbazole was obtained. Because no gene encoding an enzyme that could selectively cleave the C-N bond in 2-aminobiphenyl was available, it was not possible to construct a new metabolic pathway for the selective removal of nitrogen from carbazole.

In the course of gene expression studies an increased understanding of gene transcription in *Rhodococcus erythropolis* was developed. Chromosomal DNA fragments were screened for promoter activity, transcription initiation sites were mapped, and promoter regions were examined to detect conserved sequences. No DNA sequences that resemble the two conserved hexanucleotide sequences typically found in *E. coli* promoters were present in DNA fragments that functioned as promoters in *R. erythropolis*. However, possible conserved DNA sequence motifs were observed and synthetic DNA fragments were constructed to test the importance of these motifs for promoter activity in both *E. coli* and *R. erythropolis*. A DNA sequence about 10 bp upstream from the transcriptional initiation site was demonstrated to be important for promoter activity in *R. erythropolis*. This sequence differs significantly from typical *E. coli* promoters. This information will aid in our understanding of gene regulation/expression in the genus *Rhodococcus*.

Rhodococcus cultures are capable of metabolizing a wide range of substrates, including hydrophobic substrates like petroleum, and may find widespread use in the development of biotechnology processes in the future. The results of this project will directly benefit the development of future biotechnology processes/projects employing *Rhodococcus* hosts.

Three manuscripts were prepared and submitted for publication based on the data obtained in this project:

"Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4" submitted to Biochemical and Biophysical Research Communications,

"Screening and Analysis of DNA Fragments that Show Promoter Activities in *Rhodococcus erythropolis*" manuscript in preparation,

"Microbial Biocatalyst Developments to Upgrade Fossil Fuels" Current Opinion in Biotechnology (In Press).

INTRODUCTION

With the decline in the production of light and medium weight crude oils, refineries will increasingly be forced to process heavier and sour crudes [1]. These crude oils are high in sulfur, nitrogen and metals. Nitrogen and sulfur in petroleum can foul catalysts decreasing the efficiency of hydrotreating and catalytic cracking processes [2, 3]. Sulfur-containing compounds typically found in petroleum that are recalcitrant to removal by hydrotreating are dibenzothiophene and alkylated derivatives of dibenzothiophene. The heavier gas oils and residua contain both basic and non-basic nitrogen compounds. The basic nitrogen compounds include pyridine, quinoline, acridine, phenanthridine, and their derivatives. These are responsible for poisoning of FCC catalysts by the reaction of the basic compounds with the acidic sites of the catalyst. The non-basics are predominantly mixed alkyl derivatives of carbazole and account for 70-75% of the total nitrogen content of crude oil (0.3% N). The neutralization of the active acid sites results in deactivation of the catalyst. Nitrogen poisoning also affects the selectivity of the reaction. Carbazole, a major constituent of the non-basic portion (and hence of the total nitrogen present), gets converted into basic derivatives during the cracking process and adsorbs and poisons the catalyst as described above. Nitrogen compounds in petroleum foul catalysts and thus decrease the efficiency of the existing hydrotreating and catalytic cracking processes. In addition to catalyst fouling, nitrogen compounds also promote corrosion of the equipment. Also, the combustion of nitrogen compounds leads to formation of nitrogen oxides (NOx) which, in the presence of other hydrocarbons (VOCs:volatile organic compounds) and sunlight lead to ozone formation. Both ozone and NOx are hazardous to human health. Removal of these organonitrogen compounds will not only significantly improve the efficiency of the catalytic cracking process and result in cost savings for the refinery but also decrease atmospheric pollution. The selective removal of nitrogen from petroleum is a relatively neglected topic in comparison with sulfur removal. Moreover, most metals in oil are associated with nitrogen compounds, and nitrogen compounds contribute to the instability of petroleum byproducts[4-6]. The selective removal of nitrogen from oil would be highly desirable, but effective processes are not currently available.

There is hence a need to develop alternate cost-effective and energy-efficient technologies for the removal of sulfur, nitrogen and metals. Existing thermochemical processes, such as hydrodesulfurization, can efficiently remove much of the sulfur and nitrogen from

petroleum but the selective removal of all organically bound sulfur and nitrogen, and the removal of metals cannot be efficiently accomplished using currently available technologies. The specificity of biochemical reactions far exceeds that of chemical reactions. Moreover biorefining can be performed at comparatively low temperatures and pressures and does not require hydrogen thus avoiding a significant amount of operating costs associated with the conventional hydrodesulfurization process. The selective removal of sulfur from dibenzothiophene and from petroleum by biochemical reactions performed by microorganisms has been demonstrated, but improved cultures are needed before a viable biodesulfurization process can be developed. Biorefining can also potentially be used to remove nitrogen and metals from petroleum, but so far this area of research has received very little attention.

Biorefining can complement existing technologies by specifically addressing compounds/contaminants refractory to current petroleum refinery processes [7]. Heteroatoms such as nitrogen, metals, and sulfur can poison the catalysts used in catalytic cracking and hydrotreating processes. Existing refineries are not capable of operating efficiently with heavy crude oils and residuum that have high heteroatom content. Bioprocesses could be used to pre-treat oil reducing the heteroatom content allowing the use of heavy crude oils that could not otherwise be treated with existing refinery processes. Biorefining processes can also be used in conjunction with existing processes to meet the increasingly stringent environmental requirements for contaminant reduction. Additionally, most current technologies focus on the removal of sulfur while the development of processes to remove nitrogen, and its associated heavy metals, is a comparatively neglected research topic that will increase in importance as the quality of available petroleum declines.

There is currently no biochemical pathway, or thermochemical process, for the selective removal of nitrogen from compounds typically present in petroleum. Previous research by GTI characterized the biochemistry and the genetics of microbial enzymes capable of cleaving one of the two carbon-nitrogen bonds in carbazole. This project sought to extend that work by constructing a biochemical pathway enabling the selective and complete removal of nitrogen from carbazole and related compounds.

Carbazole is a good model compound that is representative of the nitrogen-containing compounds present in the greatest abundance in many petroleum samples. For developing a biological process for the removal of nitrogen from petroleum no known carbazole-degrading

culture is particularly appropriate because nitrogen is only removed in the course of overall degradation. A microorganism capable of selectively cleaving C-N bonds in quinoline and removing nitrogen from petroleum was isolated and characterized (by GTI). But quinoline is not the predominant form of organonitrogen compounds in petroleum. What is wanted is selective cleavage of both C-N bonds in carbazole, and related compounds, resulting in the selective removal of nitrogen while leaving the rest of the molecule intact.

A variety of carbazole-degrading microorganisms have been reported in the literature including *Sphingomonas, Pseudomonas, Mycobacterium, Ralstonia* and *Xanthamonas* species [7]. Insofar as biodegradation pathways have been investigated, these differing species of carbazole degraders follow a similar carbazole degradation pathway that begins with the oxidative cleavage of the heterocyclic nitrogen ring of carbazole to form 2'-aminobiphenyl-2,3-diol. This compound is then oxidized through meta cleavage yielding 2-hydroxy-6-oxo-6-hexa-2e,4z-dienoate. The next metabolic steps result in the degradation of one of the aromatic rings releasing carbon dioxide. In existing pathways nitrogen is released from carbazole only after substantial carbon degradation. Figure 1 illustrates the carbazole degradation pathway employed by all currently known carbazole utilizing cultures as well as the potential pathway for selective removal of nitrogen from carbazole that was a goal in this project.

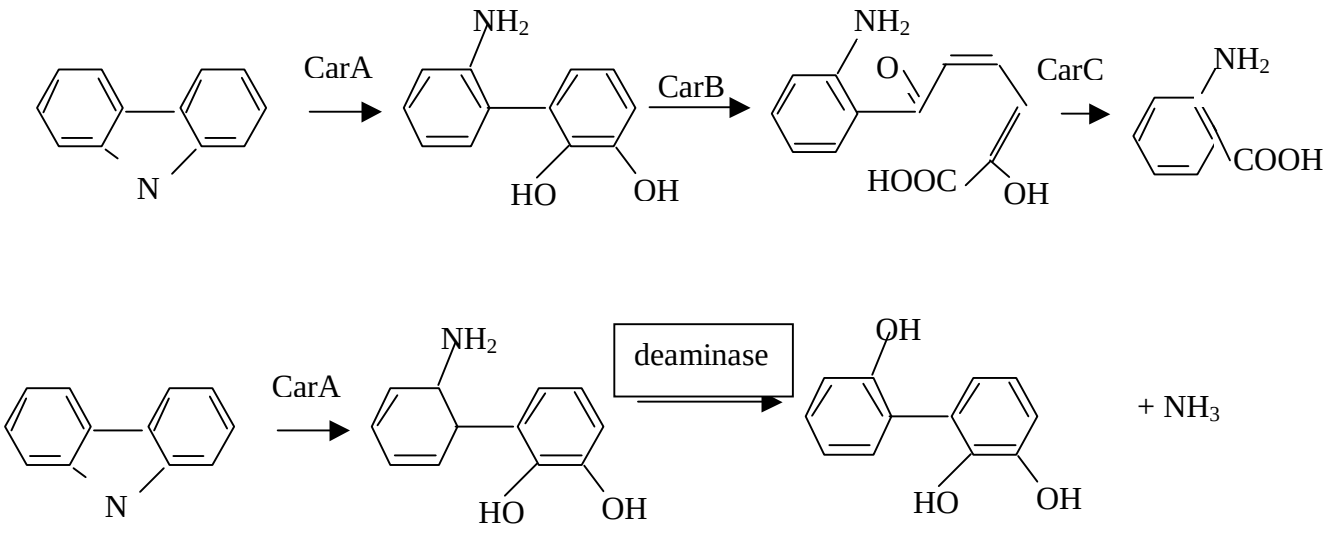


FIGURE 1. CARBAZOLE DEGRADATION PATHWAYS.

The top pathway illustrates the existing carbazole degradation pathway that results in overall degradation, whereas the bottom pathway illustrates the potential pathway for the selective removal of nitrogen from carbazole that was a goal in this project.

Sphingomonas sp GTIN11 [8] was isolated by GTI scientists and demonstrated to metabolize carbazole, and to a lesser extent C1 and C2 derivatives of carbazole, from petroleum. Moreover, the genes encoding a portion of the carbazole degradation pathway of *Sphingomonas* sp. GTIN11 have been cloned and sequenced. The reaction catalyzed by CarA converts carbazole to 2'-aminobiphenyl-2,3-diol accomplishing the cleavage of the first C-N bond in carbazole. There are no known deaminases that can metabolize 2'-aminobiphenyl-2,3-diol and accomplish the cleavage of the final C-N bond. This project employed enrichment culture, and directed evolution to isolate and/or create a deaminase that will recognize 2'-aminobiphenyl-2,3-diol as a substrate. If a gene encoding an appropriate deaminase was identified, it could be combined with the *carA* genes (*carAa*, *carAc*, and *carAd* encoding for the carbazole

dioxygenase, ferredoxin and ferredoxin reductase respectively) from *Sphingomonas* sp. GTIN11 and thereby construct a synthetic operon for the selective removal of nitrogen from carbazole, as shown in Figure 1. The *carA* genes from *Sphingomonas sp*. GTIN11 were used in this project because this is the only carbazole degrading culture demonstrated to remove nitrogen from petroleum. A preferred bacterial strain would lack the *carB* and *carC* genes so that complete biodegradation of carbazole would be avoided and the final product would be 2',2,3-trihydroxybiphenyl (or a similar compound).

Enrichment culture experiments were performed to isolate bacterial cultures capable of utilizing 2-aminobiphenyl as a sole nitrogen source. Cultures were then be tested to determine if they contain an enzyme that can deaminate 2-aminobiphenyl. Because 2'aminobiphenyl-2,3-diol is not commercially available enrichment culture experiments employed 2-aminobiphenyl. Known deaminases have limited ability to metabolize aromatic amides and their substrate range includes benzamide, toluamide, and anthranilamide. However, these enzymes can possibly by modified to recognize 2-aminobiphenyl as a substrate. Alternatively, the initial step in aniline degradation, catalyzed by aniline dioxygenase, is the conversion of aniline to catechol and ammonia. The chemical similarity between aniline and 2-aminobiphenyl make aniline dioxygenase a good candidate enzyme for potentially cleaving the C-N bond in 2-aminobiphenyl. The key goal of the project was to identify a bacterial culture that can deaminate 2-aminobiphenyl, then the deaminase gene could be cloned, sequenced, and combined with the *carA* genes of *Sphingomonas* sp. GTIN11 to create a novel metabolic pathway for the selective cleavage of C-N bonds. It was hoped that a novel metabolic pathway could be constructed that confers the ability to selectively cleave both C-N bonds in carbazole.

This project is relevant to the DOE's interests because it can contribute to greater efficiency in petroleum refining and can decrease environmental pollution due to the use of petroleum. Nitrogen in petroleum contributes to air pollution and decreases refinery efficiency by poisoning catalysts, but it is difficult to remove organically bound nitrogen without destroying the calorific value of the fuel. A key objective of the project was to develop biochemical pathways for the selective cleavage of C-N bonds in molecules found in petroleum. The reduction of nitrogen and metals in petroleum will allow refineries to operate more efficiently. This will decrease costs and will protect the environment.

In North America alone over 3 trillion barrels of known petroleum reserves are largely untapped or underutilized because of their high sulfur/nitrogen/metals content and attendant viscosity problems [9]. Energy statistics indicate that the U.S. imports 65% of its oil demand [10]. New technologies will allow a greater utilization of heavy oils and residuum while still maintaining refinery efficiency and environmental protection. The National Petroleum Refineries Association estimated the cost of meeting Clean Air Act regulations requiring a maximum sulfur content of 0.05% for diesel fuel by 1994 cost about $3.3 billion in capital expenditures and $1.2 billion in annual operating costs [10]. Similar estimates for the removal of nitrogen and metals from heavy oils and residuum are not available. However, diesel is far easier to treat than heavy oils so that one would predict that the costs associated with upgrading heavy oils and residuum would be correspondingly higher. The removal of nitrogen and metals prior to combustion of petroleum also protects the environment by eliminating contaminants that would otherwise contribute to air pollution.

MATERIALS AND METHODS

Bacterial Cultures and Growth Conditions

Environmental samples were obtained from petroleum and/or hydrocarbon contaminated soil. The environmental samples were used to inoculate nutristat and shake flask directed evolution/enrichment culture experiments to obtain cultures that may be suitable for the metabolism of organonitrogen compounds. A further description of the methodologies used in the isolation and characterization of bacterial cultures that can selectively cleave C-N bonds can be found in GTI publications: Kilbane II, J. J., A. Daram, J. Abbasian, and K. J. Kayser, 2002, "Isolation and characterization of *Sphingomonas* sp. GTIN11 capable of carbazole metabolism in petroleum" Biochemical & Biophysical Research Communications 297: 242-248, and Kilbane II, J. J., R. Ranganathan, L. Cleveland, K. J. Kayser, C. Ribiero, and M. M Linhares, 2000, "Selective removal of nitrogen from quinoline and petroleum by *Pseudomonas ayucida* IGTN9m", Applied & Environmental Microbiology 66: 688-693.

Multiple nutristats were set up employing a defined nitrogen-free mineral salts medium (Mod A):

KH_2PO_4	0.37 g/L
$MgSO_4.7H_2O$	0.25 g/L
$CaCl_2.2H_2O$	0.07 g/L
$FeCl_3$	0.02 g/L
Glucose/glycerol/succinate	20.0 g/L

This medium was adjusted to pH 6.5 to 7 and nitrogen was supplied in the form of an organonitrogen test compound in the 0.5-20 mM concentration range. For the positive nitrogen control, 10 mM NH_4Cl (0.535 g/L) was used.

Two improved recipes for a defined mineral salts media, A2 and MMN, that yielded better growth of some bacterial isolates was also used. A mixture of glucose, glycerol, and succinate was employed as a carbon source to encourage the growth of a wide range of microbial species.

Nitrogen-free medium A2 is composed of 5.0 g glucose, 6.3 g KH_2PO_4, 8.0 g K_2HPO_4, 0.2 g $MgSO_4$, 1.0 ml vitamin solution and 2.0 ml metal solution per liter. The vitamin solution contains 400 mg Ca-pantothenate, 400 mg niacin, 400 mg pyridoxine-HCL, 200 mg inositol, 200 mg p-aminobenzoate, 200 mg D-biotin, 50 mg riboflavin, 50 mg folic acid, 15 mg thiamine HCL and 50 mg cyanocobalamin per liter of distilled water. The metal solution contains 100 mg $MnSO_4$, 31.7 mg $CuCl_2$, 25 mg $(NH_4)_6Mo_7O_{24}$, 20 mg H_3BO_3, 25 mg $Co(NO_3)_2$, 25 mg $ZnCl_2$ and 10 mg NH_4VO_3 per liter of distilled water.

The recipe of the improved nitrogen-free minimal media MMN is:

Compound Name	1 X
EDTA	3.2 mg
MOO_3	0.1 mg
Na_2HPO_4	1,419.6 mg
$KH_2\ PO_4$	1,360.9 mg
$MgSO_4$	98.5 mg
$CaCl_2 \bullet 2H_2O$	5.88 mg
H_3BO_4	1.16 mg
$FeSO_4 \bullet 7H_2O$	2.78 mg
$ZnSO_4 \bullet 7H_2O$	1.15 mg
$MnSO_4 \bullet H_2O$	1.69 mg
$CuSO_4 \bullet 5H_2O$	0.38 mg
$CoCl_2 \bullet 6H_2O$	0.24 mg

*** In 1 L ddH_2O

Nutristats and shake flasks were operated at temperatures of 25 (room temperature), 37, and 45 °C. The working volume of nutristats is one liter and shake flask experiments generally utilize 25 to 100 mL of liquid medium. The organonitrogen test compound is routinely varied during the course of operation of the nutristats/shake flask experiments. Nutristats are operated in series so that the effluent of one nutristat serves as the influent for the next. Carbazole, 2-aminobiphenyl, benzamide, aniline, 4,4'-azodianiline, ortho-, meta-, and para-toluidine, quinoline, pyridine, quinazoline, quinoxaline, piperidine, pyrolidine, triazine or other test compounds are added to the fresh media influent at concentrations of 0.5-20 mM. Flow rates of the nutristats are adjusted to achieve hydraulic retention times ranging from 35 hours to 60 hours. The flow rates and the organonitrogen test compound are altered as needed to ensure that the nutristats create an environment suitable for the selection of cultures with improved abilities to selectively cleave C-N bonds. This means that the bacterial cell density in the nutristats/shake

flasks ranges from 10^2 to 10^8 cells/mL, but generally cell densities of 10^4 to 10^5 cells/mL are maintained. The bacteria isolated from the effluent of nutristats and/or from shake flasks or nitrogen bioavailability assays are subjected to short wave ultraviolet (UV) irradiation. Cell populations are mutagenized under conditions that result in the death of about 99% of the population. The mutagenized cells are then used to reinoculate nutristats, start additional shake flask experiments, and to streak onto agar plates containing organonitrogen test compounds. Care is taken to ensure that the amount of biomass that is added back to nutristats in the form of inocula is insufficient to provide a significant amount of nitrogen in the form of dead biomass. Hence significant bacterial growth in the nutristat experiments should be due to the utilization of nitrogen from the organonitrogen test compounds and not from readily available sources such as dead biomass. The effluent of nutristats and cells from shake flasks and from agar plates are routinely tested using the nitrogen bioavailability assay.

Nitrogen Bioavailability Assay

The nitrogen bioavailability assay utilizes defined mineral salts medium in growth tests in which organonitrogen model compounds such as 2-aminobiphenyl, benzamide, aniline, 4,4'-azodianiline, diphenylamine, maleimide, ortho-, meta-, and para-toluidine, quinoline, pyridine, carbazole, quinazoline, piperidine, pyrolidine, and triazine serve as sources of carbon and/or nitrogen. For selective cleavage of carbon-nitrogen bonds, a culture should be capable of utilizing an organonitrogen compound as a nitrogen source but not as a carbon source. Accordingly, growth tests are performed using the following eight conditions:

1. Test compound as sole source of carbon and nitrogen.
2. Test compound as sole source of carbon (alternative nitrogen source, ammonia, is available).
3. Test compound as sole source of nitrogen (alternative carbon source, glucose/glycerol/succinate, is available).
4. Test compound present as well as alternative sources of carbon and nitrogen.
5. Only alternative nitrogen (ammonia) and carbon (glucose/glycerol/succinate) sources are available. The test compound is not present.
6. No carbon or nitrogen compounds of any kind are present.

7. Only alternative nitrogen (ammonia) is present. No carbon or test compound is present.
8. Only carbon (glucose/glycerol/succinate) sources are available. No nitrogen compounds (ammonia or test compound) are present.

These eight growth conditions constitute a bioassay for the ability of a culture to metabolize organonitrogen compounds. The basis of the nitrogen bioavailability assay is that all microorganisms require nitrogen for growth. When carbon and nitrogen sources other than the test compounds were needed, they were supplied in the form of glucose/glycerol/succinate, and as ammonia respectively.

The nitrogen bioavailability assay described above can be performed with any organonitrogen test compound that is ordinarily used at a concentration of from 0.5 to 20 mM. The various cultures to be tested are inoculated into test tubes or shake flasks containing medium components appropriate for the eight test conditions. The cultures are then incubated aerobically for 2 to 28 days, at 25, 37, and 45°C. The growth of the cultures is monitored easily by measuring the turbidity/optical density of the cultures in the various test conditions, or by determining colony-forming units. The unamended sample (test condition No. 6) serves as a negative control while the samples amended with both a carbon and nitrogen source (test conditions No. 4 and 5) serves as positive controls and should produce healthy microbial growth unless the test compound is toxic to the culture being tested. If the test compound is toxic only condition No. 5 should result in healthy growth. The amount of bacterial growth observed in test conditions 1, 2, and 3 in comparison with the amount of growth observed in test conditions 4, 5 and 6, indicate the ability of cultures to use the organonitrogen test compound as a source of carbon and/or nitrogen. Those cultures which show better growth in test condition No. 3 than conditions Nos. 1 or 2 may be preferentially utilizing the organonitrogen compound as a nitrogen source only, and should be examined more thoroughly and included in further experiments. Conditions 7 and 8 serve as negative controls, and no growth should occur unless there is a problem with the batch of media used, or if the bacterial culture being tested has the ability to fix atmospheric nitrogen.

E. coli JM109 and BL21 (DE3) were purchased from Novagen (Madison, WI), while chemical competent *E. coli* DH5α was purchased from the Cell Media Facility at the University

of Illinois (Urbana, IL). The plasmids pACYC Duet and pTrc99A were from Novagen (Madison, WI) and Amersham Pharmacia respectively. The plasmid pUCARA containing the carbazole dioxygenase gene cluster (CarA1-A5) was kindly provided by Dr Nojiri from the Biotechnology Research Center, The University of Tokyo, Tokyo, Japan. The plasmids, pAS91 and pAS93, both containing the aniline dioxygenase gene cluster were kindly provided by Dr Takeo from the Dept of Applied Chemistry, Himeiji Institute of Technology, Hyogo, Japan. Ampicillin, chloramphenicol, β-D-thiogalactopyranoside (IPTG), were purchased from Sigma (St. Louis, MO).

Carbazole, aniline, 2-ethylaniline, 2-isopropylaniline, 2-sec-butylaniline, 2-tert-butylaniline, 2-methylanthranilate, 2-aminobiphenyl and Gibb's reagent were purchased from Sigma-Aldrich (St. Louis, MO).

Carbazole was dissolved in dimethyl sulfoxide (DMSO) to form a stock solution of 100 mg/ml and filtered once to remove insoluble impurities. Aniline, 2-ethylaniline, 2-isopropylaniline, 2-sec-butylaniline, 2-tert-butylaniline, 2-methylanthranilate and 2-aminobiphenyl were dissolved in ethanol to a final concentration of 0.5M.

CarA resting cell assay

The 2xYT media used in carbazole dioxygenase CarA (pUCARA) assay was prepared as described by Sambrook et al. [11]. The minimal media (MM) used in carbazole dioxygenase CarA (pUCARA) assay had the following composition: 800 ml Millipore water, 1.32 g $Na_2HPO_4.7H_2O$, 0.64 g KH_2PO_4, 0.012 g $FeSO_4.7H_2O$, 0.012 g $CaCl_2.2H_2O$, 0.0059 g $MgSO_4$, 0.02 g yeast extract. The initial pH was adjusted to 7.0. The media was then filter sterilized.

A colony of BL21 (DE3) harboring the pUCARA plasmid was grown overnight in 2ml 2xYT with 100 μg/l ampicillin at 37 ^oC under constant shaking. The overnight culture was then inoculated into 200 ml 2xYT with 100 μg/l ampicillin and incubated in a 37^oC shaker until the OD at 550 nm reaches about 0.4. At that point, protein expression is induced by adding IPTG to a final concentration of 1mM. The culture is then incubated in a 30^oC shaker for 6 to 8 hrs. After the induction period, the cells were spun down and washed with ice cold MM twice. The cells were resuspended in MM or 2xYT with a final volume of 20 ml. Carbazole was then added to desired concentrations. DMSO was added to a final concentration of 3% v/v.

At stipulated times, samples of the resting cell assay were taken and acetonitrile (ACN) was added at a 1.33 volume ratio to dissolve any insoluble carbazole. The mixture is vortexed for 1 min and spun down at 400 rpm for 10 min at 4 ^{o}C in an Eppendorf 5810R centrifuge. The supernatant was collected and analyzed by HPLC. To purify 2'-aminobiphenyl-2,3-diol (2'-ABPD), the 20 ml of resting cell culture was acidified with 50 µl of 1M HCl and extracted with ethyl acetate. The organic phase is then evaporated and the extracts were redissolved in 0.5 ml acetronitrile. The redissolved extracts were then run through a silica gel column using a mobile phase comprising of methanol (7.5%) and toluene (92.5%). The eluted fraction containing 2'-ABPD was vacuum evaporated and redissolved with DMSO.

Thin Layer Chromatography for Identification of Metabolites

Thin layer chromatography (TLC) was performed on Whatman Silica C-18 plates by the method described by Watson and Cain [12]. Running phase solvents used were chloroform-toluene (1:3), and hexane-acetic acid-xylene (5:1:2). Supernatants from bacterial cultures grown with an organonitrogen test compound as the sole source of nitrogen were obtained after centrifugation at 10,000 x g for 15 minutes. These supernatants were used at neutral or alkaline pH. Typically 10 ml of aqueous supernatant was acidified to pH 1 to 2 with HCl and extracted with ethyl acetate (1:1 or 1:0.5 v/v). The organic phase was separated from the aqueous phase by centrifugation or by using a separatory funnel. The ethyl acetate extract was then evaporated in a hood resulting in the concentration of the sample from 20 to 1000-fold prior to the analysis of the extracts by TLC. 10 to 50 µL of ethyl acetate sample that had been concentrated 100-fold relative to the volume of aqueous supernatant extracted was spotted onto TLC plates. Typical running times of the TLC plates were about 20 minutes. These plates were later observed under normal lighting, short (245 nm), and long wave (366 nm) UV light.

Some experiments also utilized resting cells that were prepared by centrifuging from 500 ml of log phase cultures grown with either an organonitrogen compound or ammonia as nitrogen sources. Then the washed cell pellets were resuspended in 5 to 50 ml of mineral salts medium achieving final cell densities of from 10^{10} to 10^{11} cells/ml. These cell suspensions were incubated with from 1 to 20 mM test compound (organonitrogen compound) for periods ranging from 15 minutes to 24 hours. The incubator was agitated at about 200 rpm and maintained at the

microorganism's optimum temperature. The ethyl acetate extract was stored in amber vials at 4°C until they were analyzed by TLC, HPLC and/or GC-MS.

Gas Chromatography-Mass Spectrometry

GC-MS analysis was performed on extracts derived from growing and resting cell cultures exposed to organonitrogen test compounds, and on compounds eluted from spots observed on thin layer chromatography plates.

Extraction of the supernatants from resting cells as well as growing cells were carried out either by ethyl acetate solvent extraction or with C-18 solid phase extraction cartridges as described above for the preparation of samples for TLC analysis. Additionally, TLC spots of possible metabolites were scraped from the TLC plates and eluted with ethyl acetate and concentrated for analysis by GC-MS.

For analysis of the extracts a Hewlett Packard 5971 mass selective detector and 5890 series II GC with HP 7673 auto sampler tower and a 30 meter Rezteck XTI-5 column was used. The final oven temperature was maintained at 300°C. The detection limit was 1 ng or 1 μg/ml with a 1 μl injection. Mass spectrographs were compared with various libraries of mass spectrograph data prepared from known standard compounds. Several chromatograph libraries were consulted to determine the identity of metabolites of organonitrogen compounds. The presence or absence of nitrogen in various compounds was also determined by GC-AED using the nitrogen-specific wavelength of 174.2 nm for detection.

Some samples were derivatized with trimethylsilane (TMS) prior to analysis by GC-MS. The TMS-silylating reagent was prepared using 9 parts pyridine, 3 parts hexamethyldisilizane (Sigma-Aldrich catalog # 440191), and 1 part chlorotrimethylsilane (Sigma-Aldrich catalog # C72854). Approximately 25 mg of a sample (3 drops of ethyl acetate extract) was added to 2mL of TMS-silylating reagent under anaerobic conditions. The mixture was agitated/mixed and allowed to incubate at room temperature for 30-60 minutes. The samples were centrifuged and the supernatants were used for GC-MS analysis.

High Performance Liquid Chromatography

The extracts derived from growing and resting cell experiments were analyzed by HPLC. Extraction was carried out with ethyl acetate as described in the TLC section. The ethyl acetate was then evaporated completely and the residue (nonvolatile organics) was suspended in

acetonitrile before injecting into the HPLC system. A Waters system equipped with a Symmetry C_{18} (3.5 μm, 4.6 × 100 mm) column and a 600 controller was used for this purpose. Detection of compounds was carried out using a 996 photodiode array detector coupled to the HPLC system. An isocratic mobile phase of acetonitrile:water at the flowrate of 1.5ml/min was used as the running solvent. For other assays the identification and quantitation of carbazole and 2'-aminobiphenyl-2,3-diol (2'-ABPD) was done using an Agilent 1100 series HPLC with the following solvents: H_2O with 1% acetic acid (solvent C) and acetonitrile with 1% acetic acid (solvent D). The column used is the SB-C8 reverse phase column. The method used is 10% to 70% solvent D in 20 min under a flow rate of 1.5 ml/min. Carbazole and 2'-ABPD were detected using a wavelength of 260 nm.

Gibb's Reagent solid phase screening

The solid phase screening assay was carried out as described elsewhere with slight modifications. *E. coli* JM109/pTrcA-2 was first grown overnight at 37°C on LB plates with 100 mg/l ampicillin. Using a nylon membrane, the colonies formed were lifted onto a fresh LB plate with 80 μl of 50 mM IPTG spread on top and incubated at 30°C for 8 hrs. After induction with IPTG, the nylon membrane with colonies was transferred onto Whatman #1 filter papers soaked with 0.5 mM of 2-ethylaniline or other surrogate substrates and incubated at 30°C for 3hrs. The nylon membranes were then transferred onto Whatman #1 filter papers soaked with 0.025% Gibb's reagent and incubated at room temperature for about 20 min.

β-Galactosidase Assay

β-Galactosidase activity was assayed in cultures of exponentially growing cells, and the activity was quantified essentially by the method of Miller [13], except that *Rhodococcus* cells were sonicated whereas *E. coli* cells were treated with chloroform and sodium dodecyl sulfate to obtain cell lysates that were used for the β-galactosidase activity determinations and protein determinations. Protein assays were performed using the Pierce BCA Protein Assay Reagents (Pierce, Rockford, IL) according to the manufacture's directions. Activities were expressed in

Miller units per milligram of protein rather than in units per $O.D._{600}$ unit, since cell size and therefore light-scattering measurements vary slightly at different growth rates.[6]

Auto-oxidation

JM109/pTA2-3 pr pTA3-1 colonies were picked and streaked onto LB plates containing 0.5 mM IPTG and incubated overnight at 37°C. Using nylon membranes, the colonies were then transferred to M9 minimal media agar plates containing 0.5 mM IPTG and 5mM of the desired substrates and incubated at 30°C.

Genetic Techniques

Methods used in genetic experiments are described in detail in the recent publications from GTI's biotechnology laboratory:

"New Host Vector System for *Thermus* spp. Based on the Malate Dehydrogenase Gene", K. J. Kayser and J. J. Kilbane II, Journal of Bacteriology 183: 1792-1795. (2001) [14].

"Inducible and Constitutive Expression Using New Plasmid and Integrative Expression Vectors for *Thermus* sp." K. J. Kayser, J.-H. Kwak, H.-S. Park, and J. J. Kilbane II. Letters in Applied Microbiology 32: 1-7 (2001) [15].

Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Pfu Turbo DNA polymerase was obtained from Stratagene (La Jolla, CA). Taq DNA polymerase was purchased from Promega (Madison, WI). Plasmid, PCR and gel purification kits were obtained from Qiagen (Valencia, CA). DNA primers were ordered from Integrated DNA Technologies (Coralville, IA).

<u>Electroporation of *Rhodococcus erythropolis*</u>

1. Grow cells in NZ for 24-48 hours at 30°C.
2. Dilute 1/5 in fresh NZ, (50 ml total volume).
3. Grow at 30°C until OD_{600} of 0.6.
4. Harvest by centrifugation at 4°C and wash 4 times with 1.0 ml ice cold 0.3M sucrose.
5. Resuspend cells in 1.0 ml 0.5M sucrose.

6. Add 100µl of cells to ice cold 0.2cm electroporation cuvette, and add 2 µl plasmid DNA.
7. Pulse at 25 µF, 2.4kV, 800Ω.
8. Dilute immediately in 5.0 ml NZ broth and incubate for 4 hours at 30°C.
9. Concentrate cells by centrifugation and resuspend in 1.0 ml NZ.
10. Plate 100µl of cells on appropriate antibiotic plates and incubate at 30°C.

Construction of a Genomic DNA Library and Screening

Genomic DNA from *Rhodococcus erythropolis* IGTS8 was digested with Sau3AI and fractionated by agarose gel electrophoresis, and fragments from 50 to 1000 bp were purified using the Qiaex gel extraction kit (Qiagen, Germany). Approximately 60 ng of purified genomic DNA fragments were ligated to 300 ng of dephosphorylated, BamHI-cleaved pEBC26 in a final volume of 20 µl. The ligation mixture (10 µl) was then electroporated into CPE648 and the transformed cells were plated on NZ agar containing tetracycline (12.5 µg/ml) and X-gal. Blue colonies were restreaked on the same plates and subsequently inoculated into tetracycline (12.5 µg/ml)-containing NZ medium for either plasmid isolation or for β-galactosidase assay.

PCR Primers

The PCR primers used for the construction of promoters derived from the *E. coli lacZ* promoter were:

5'Pro Xba1 (-130 of the *lacZ* promoter region)
TCATCTAGATGCAGCTGGCACGACAGGTTTC

5'Pro Xba2 (-50 of the *lacZ* promoter region)
TCATCTAGAGGCACCCCAGGCTTTACACTTT

5'Pro Xba3 (alters the -35 hexanucleotide region of the *lacZ* promoter)
TCATCTAGAGGTGGCCTTTATGCTTCCGGCTCGTATGT

3' Pro Nsi1 (+36 of the *lacZ* promoter)
TCAATGCATAGCTGTTTCCTGTGTGAAATTGTTAT

3' Pro Nsi2 (+1 of the *lacZ* promoter, and alters the RBS)
TCAATGCATTCCTTTTCCACACAACATACGAGCCGG

3' Pro Nsi3 (+1 of the *lacZ* promoter, alters the RBS and the –10 hexanucleotide region)

CAATGCATTCCTTTTCCACACTCCGCGCGAGCCGGAAGCATAAAGTGTAA

The primers used for amplification of carbazole dioxygenase genes were:

CarAa GTIN11 Forward Primer
5'-ACACACGCATGCATGGCTAACCAACCATCAATCGC-3'

CarAa GTIN11 Reverse Primer
5'-ACACACGGTACCTCAACGTGCTTCCTGAATGCCG-3'
Restriction sites were added into the PCR primers to give 5'SphI and 3'KpnI sites.

CarAc Forward Primer
5'-ACACACGGTACCATGACCGCAAAGGTCCGCGTGA-3'

CarAc Reverse Primer
5'-ACACACCCCGGGTCATGACTGTGCCTTCCGGCAGTT-3'
Restriction sites were added into the PCR primers to give 5'KpnI and 3'XmaI sites.

CarAd Forward Primer
5'-ACACACCCCGGGATGTACCAACTCAAAATTGAAGG-3'

CarAd Reverse Primer
5'-ACACACCTGCAGTTAGAAAAATGCGTCAAAATG
Restriction sites were added into the PCR primers to give 5'XmaI and 3'PstI sites.

CarAa CA10 Forward Primer
5'-ACACACGGATCCATGGCGAACGTTGATGAGGCAAT-3'

CarAa CA10 Reverse Primer
5'-ACACACGAGCTCTCAAGCGCCTCCTTTCAACATATCGG-3'
Restriction sites were added into the PCR primers to give 5'BamHI and 3'SacI sites.

CarAcd CA10 Forward Primer
5'-ACACACGAGCTCATGAACCAAATTTGGTTGAAAG-3'

CarAcd CA10 Reverse Primer
5'-ACACACGAGCTCTTAGAAAAATGCGTCAAAATGAAT-3'
Restriction sites were added into the PCR primers to give 5'SacI and 3'SacI sites.

Error prone PCR

Error prone PCR was used to create a library of mutant AtdA3 genes according to the protocol described elsewhere. The primers used for error-prone PCR (EP-PCR) of AtdA3 are 5'-GTGGGTGATGTATAACCTAGGAGG-3' and 5'-TGGCGTACCTCAACCGGTTTA-3'. Restriction sites are underlined in the primer sequences.

Reverse Transcription-Quantitative Polymerase Chain Reaction (RT-qPCR) for the Quantification of Gene Expression

Total RNA isolations from DH5 α cells harboring the promoter versions were achieved by a modification of the acid phenol method [16]. RNA isolations from Rhodococcus CW25 cells were carried out by using the Fast RNA Pro Blue kit (Q-BIO gene, OH). About 25 ml of an overnight culture in LB-broth, incubated with 20μg/ml of kanamycin (kan) was used for RNA isolation from DH5α. In the case of CW25 cells 100 ml of LB-broth incubated with 50μg/ml of kan at 30°C was employed for all RNA isolations. The 0.1mm Zirconia/silica bead and the 2ml sterile vials for bead beating were from Biospec products, Inc, OK. All the reagents required for the acid phenol method of RNA isolation as well as the Quantitative RT-PCR Ready Mix kit was purchased from Sigma, MO. After RNA isolations, the samples were subjected to DNase treatment in order to remove contaminating DNA. A DNA-free kit (Ambion, TX) was used for the DNase treatment. The samples were then stored at –80°C until further use.

Reverse transcriptase-q PCR was employed to quantify kanamycin mRNA's as a direct measure of the promoter strengths in *Rhodococcus erythropolis* CW25 and *Escherichia coli* DH5α cells. The attempt here was designed to include and external reference mRNA as a control for the efficiency of the different procedures involved in this protocol [16]. Luciferase (luc) mRNA (Promega) was used as the external reference mRNA. 2μl of a 1:100 dilution of the luc mRNA was added to the samples during RNA isolation. Primers and probes (MWG Biotech, NC) were designed for kanamycin (primary target), tetracycline (marker on the base plasmid pREP45boxBnsikan) and luciferase (external reference).

A typical RT-qPCR reaction consisted of 10μl of the 2x master mix (Sigma), 2μl of the forward primer (1000nM), 2μl of the reverse primer (1000nM), 1.6μl of the respective probe (200nM), 1μl of reverse trancriptase and 2μl of the appropriate RNA sample. The reaction volume was then made up to 20μl with DNase and RNase free sterile water. For reactions involving the kan primers and probe, a final concentration of 3mM of magnesium chloride ($MgCl_2$) was added to the reaction in addition to the $MgCl_2$ provided in the ready mix to increase sensitivity of the reaction. This was identified during the standardization of the RT-qPCR conditions. All the RT-qPCR runs included the kan, tet and luciferase standards also other than the experimentals. A six concentration (10^8-10^3 copies) standard curve was set up for each

standard; by diluting the standards in DEPC treated sterile water. The thermal profile for the RT-qPCR had an initial hold at 48°C for 30 min's, followed by a hold at 95°C for 2 min's and 40 cycles of 95°C for 15 sec's and 60°C for 60 sec's (data acquisition).

The actual DNA and RNA concentrations in the standards and experimentals were measured using commericial kits from Molecular Probes/Invitrogen, Inc, Eugene, OR. Quant-iT Pico Green dsDNA Assay kit was used for measuring the DNA concentrations and a Quant-iT RiboGreen RNA Reagent and Kit was used for measuring RNA concentrations. Both these kits were used as per the manufacturers instructions. A Fluoromax-3 spectrofluorometer (Jobin Yvon Inc, NJ) was used to measure the fluorescence.

For RT-qPCR the TaqMan amplification method was used. The DNA sequences of the primers and probes were designed using the Primer3 program (http://frodo.wi.mit/edu/cgi-bin/primer3). The primers were designed to have Tm values of about 62°C and the probes had Tm values of about 70°C. The probes were fluorescently labeled at the 5' end using FAM dye and the 3' end contained the BHQ1 fluorescence quenching dye. The size of RT-qPCR amplicons were 67 to 78 bp. Primer concentrations were 500 nM, and the probe concentration was 200 nM. RT-PCR Ready Mix (Sigma-Aldrich, St. Louis, MO) was used. A Corbett Research Rotor-Gene RG-3000 thermocycler(Mortlake NSW, Australia) was used for RT-qPCR. Cycling conditions consisted of an initial denaturing at 95°C for 1 minute followed by 35 cycles of 95°C for 15 seconds, and 60°C for 60 seconds. Data acquisition occurred at 60°C.

luc-forward
TACAACACCCCAACATCTTCGA

luc-reverse
GGAAGTTCACCGGCGTCAT

luc-probe
CGGGCGTGGCAGGTCTTCCC

Kan-forward
TACGCTTGATCCGGCTACCT

Kan-reverse
GCTTCCATCCGAGTACGTGAA

Kan-probe
GCCCATTCGACCACCAAGCG

DNA Sequencing

Plasmid DNA was prepared by using the Maxi-prep product (Quiagen) and sequencing was performed with the dideoxy chain termination method by using United States Biochemical (USB) Sequenase^{R} 7-deaza-dGTP Sequencing Kit and [^{35}S]dATP according to the supplier's recommendations. Appropriate primers were employed so that DNA sequences of both strands of targeted regions were determined. Sequencing reactions were run on 7 M urea, 6% Long Ranger Gel (AT Biochem, PA) in 0.6X TBE buffer for 2 hours at 1650 volts with the Sequigen (Bio-Rad Laboratories, CA) sequencing apparatus. The gels were dried and autoradiography was used to visualize DNA bands.

Primer Extension Analysis

The oligonucleotides (2 pmol) for the primer extension reaction were labeled with 50 µCi of [^{32}P] ATP (~5,000 Ci/mmol), and the free nucleotides were separated from labeled oligonucleotide by passage over a 2-ml Sephadex G-25 column equilibrated in TE buffer (pH 7.0). *Rhodococcus* sp. CPE648 total RNA and *E. coli* total RNA (2-5 ug) were mixed with 5 pmol of 5'-end-labeled oligonucleotide primer in 20 µl of 0.1 M Tris hydrochloride (pH 8.3)-0.14 M KCl-0.01 M $MgCl_2$-33 U of RNAsin. This solution was incubated at 75°C for 5 min and cooled at room temperature for 30 min. Dithiothreitol was added to a final concentration of 0.01 M, and the four deoxynucleotide triphosphates were added to a final concentration of 1 mM each. Avian myeloblastosis virus reverse transcriptase (18 U) was added, and the mixture was incubated at 42°C for 1 h. The reaction was stopped by heat inactivation at 75°C for 10 min. DNase-free RNase (2 U) was added, and incubation was continued for 30 min at 45°C. Ammonium acetate (4 M, 0.1 volume) was added with 3 volumes of ethanol, and the mixture was kept overnight at -20°C. The suspension was centrifuged at a maximum speed for 20 min in an eppendorf centrifuge, and the precipitate was washed in 70% ethanol, dried, suspended in

sequencing mixture (98% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 2 mM EDTA), and separated on an acrylamide sequencing gel.

In vitro Runoff Transcription

The template DNA was linearized by digestion to completion with a suitable restriction enzyme to obtain transcripts of a defined length. A reaction mixture containing 40% glycerol, 80 mM Tris (pH 8.0), 100 mM KCl, 20 mM $MgCl_2$, 80 mM 2-mercaptoethanol, 0.2 mM EDTA, 10 mM DTT, placental ribonuclease inhibitor (40 U) and 0.5 mM each ATP, CTP, GTP was prepared. To this mixture was added 5 µCi of [^{32}P]UTP (Amersham, Arlington Heights, IL), 10 U of SP6 RNA polymerase (Ambion, Austin, TX), and 1 µg linearized template DNA. The transcription reaction, in a final volume of 20 µl, was carried out at 37°C for 1 hour. RNase-free DNase I (2 U) was added to this reaction, mixed thoroughly, and incubated at 37°C for 15 minutes. An equal volume (20 µl) of a solution of 50 mM EDTA and 200 µg of tRNA per ml was then added to each sample. At the end of the DNase I incubation, an equal volume of gel loading buffer (80% formamide, 0.1% xylene cyanol, 0.1% bromophenol blue, 2 mM EDTA) was added to the reaction, and the tube was heated for 5 minutes at 80-90°C. All of the reaction mixture was loaded on a 0.75 mm thick 5% acrylamide gel/8 M urea and run for about 30 minutes at 200 volts. Autoradiography was followed and the exposed film was used to localize precisely the area of the gel that contained the full-length labeled transcript, the most intense band on the autoradiogram. The gel containing the full-length transcript was excised and eluted in 350 µl of elution buffer (0.5 M NH_4OAc, 1 mM EDTA, 0.1% SDS) at 37°C overnight. The amount of radioactive label in an aliquot of the eluted probe was determined by scintillation counting, and these *in vitro* runoff transcripts were used directly for the ribonuclease protection assays.

RNase Protection Assay

The mixture of ^{32}P-labeled probe (prepared by *in vitro* runoff transcription) and *Rhodococcus* IGTS8 total RNA (about 10 µg) was precipitated, and the pellet was resuspended

in 20 µl hybridization buffer (80% deionized formamide, 100 mM sodium citrate pH 6.4, 300 mM sodium acetate pH 6.4, 1 mM EDTA). After the denaturation of RNA, the tubes were incubated overnight at 42°C to allow hybridization of probe and complementary mRNA in the *Rhodococcus erythropolis* IGTS8 total RNA. After hybridization, to each tube containing sample RNA, 200 µl of RNase digestion buffer (10 mM Tris-Cl, pH 7.5, 300 mM NaCl, 5 mM EDTA) and 200 µl of the diluted RNase A-RNase T1 mixture (40 µg/ml ribonuclease A and 10 units/ml ribonuclease T1) were added, and the tubes were incubated at 37°C for 1 hour to digest unprotected single-stranded RNA. After phenol/chloroform extraction, the protected double-stranded RNA was precipitated, and then the pellets were resuspended in 8 µl of gel loading buffer (80% formamide, 0.1% xylene cyanol, 0.1% bromophenol blue, 2 mM EDTA). After complete denaturation of RNA, each sample was loaded on 5% polyacrylamide/8 M denaturing gels and run at about 200 volts in 0.5 X TBE buffer. The gel was transferred to Whatman paper No.1, covered with plastic wrap, and exposed to X-ray film overnight with an intensifying screen at -80°C.

RESULTS AND DISCUSSION

Enrichment Culture Experiments

Enrichment culture experiments were performed using soil from a variety of petroleum/hydrocarbon-contaminated sites. Of primary interest was the isolation of a culture capable of utilizing 2'aminobiphenyl-2,3-diol, which is the product of the conversion of carbazole by the CarA enzyme, carbazole 1,9a-dioxygenase. Initially the nitrogen-free recipe ModA described in the Materials and Methods section was used, but an improved media was eventually adopted that contained other essential trace elements and was thought to be more likely to result in the growth of a wider range of microbial species and increase our chances of isolating bacteria that possess useful C-N bond cleaving enzymes. The organonitrogen chemicals used in enrichment culture experiments included 2-aminobiphenyl, aniline, 4,4'-azodianiline, benzamide, carbazole, methylbenzylamine, piperidine, pyridine, pyrolidine, quinoline, quinazoline, ortho-, meta-, and para-toluidine, and triazine. A mixture of glucose, glycerol, and succinate was employed as a carbon source to encourage the growth of a wide range of microbial species. The recipe of the improved nitrogen-free minimal media MMN is:

Compound Name	1 X
EDTA	3.2 mg
MOO_3	0.1 mg
Na_2HPO_4	1,419.6 mg
$KH_2\ PO_4$	1,360.9 mg
$MgSO_4$	98.5 mg
$CaCl_2 \bullet 2H_2O$	5.88 mg
H_3BO_4	1.16 mg
$FeSO_4 \bullet 7H_2O$	2.78 mg
$ZnSO_4 \bullet 7H_2O$	1.15 mg
$MnSO_4 \bullet H_2O$	1.69 mg
$CuSO_4 \bullet 5H_2O$	0.38 mg
$CoCl_2 \bullet 6H_2O$	0.24 mg

*** In 1 L ddH_2O

Table 1. Composition of Nitrogen-Free Minimal Media

Eventually, we succeeded in obtaining the growth of mixed cultures on various organonitrogen compounds and then streaked these mixed cultures on MMN agar plates or nutrient agar plates to obtain isolated bacterial colonies. Pure cultures were subsequently tested to determine if they could grow with organonitrogen compounds as sole sources of nitrogen and the results obtained with the most promising cultures are summarized below along with the

identification of each culture (MIDI closest match) based on the analysis of the DNA sequence of its 16S rRNA gene.

Culture	MIDI Closest match
GTIN16 – *Azo degrader (azo 2-6)	*Rhodococcus globerulus*
GTIN17 – Azo degrader (from m-tol MGP 10)	*Sphingomonas yanoikuyae*
GTIN18 – Carbazole degrader (from m-tol MGP 10)	*Methylobacterium radiotolerans*
GTIN21 – Carbazole degrader (from azo 2-6)	*Gordonia bronchialis*
GTIN22 – Aniline degrader (An #3)	*Gordonia bronchialis*
GTIN23 – Carbazole degrader (from azo 2-6)	*Rhodopseudomonas palustris*
GTIN24 – Carbazole degrader (from m-tol MGP 10)	*Rhodopseudomonas palustris*
GTIN25 – Azo degrader (azo 2-6)	*Gordonia bronchialis*
GTIN26 – Azo degrader (azo 2-6)	*Rhodococcus globerulus*

* Azo = 4, 4'-azodianiline

Representative results obtained in nitrogen bioavailability growth tests were:

Nitrogen bioavailability test conditions (GGS = glucose/glycerol/succinate)

1. MMN + test compounds (C & N sources)
2. MMN + NH_4Cl + test compounds (C source)
3. MMN + GGS + test compounds (N soruce)
4. MMN + GGS + NH_4Cl + test compounds
5. MMN + GGS + NH_4Cl
6. MMN + GGS

<u>GTIN16</u>

- Growth on NH_4Cl, Benzamide, Azodianiline and Carbazole as nitrogen sources
- Azodianiline as test compound
- First growth on condition 4 and next on condition 2 and 4
- After 2 weeks condition 3 was grown slightly
- Azodianiline may be used as only carbon source.

<u>GTIN17</u>

- Full Growth (+++) on NH_4Cl, benzamide, Azodianiline, p-toluidine, aniline, and 2-aminobiphenyl
- Slight growth on m-toluidine (++), o-toluidine (+), carbazole (++)
- Negative control grow slightly (+)
- First growth on condition 5 and next on condition 3 and 4
- No growth on condition 2
- After 10 days, the growth was observed on condition 1
- Azodianiline may be first used as sole nitrogen source
- After the depletion of carbon source (GGS), test compound may be used as carbon source

GTIN18

- First growth on MMN + GGS + carbazole (Faster growth than NH_4Cl medium)
- Full growth on NH_4Cl (pink floating biomass)
- Slight growth on m-toluidine, p-toluidine, o-toluidine, aniline, and azodianiline
- First growth on condition on 3 & 4 and next on condition 5 &2
- They first attack on C-C bond (condition 2).
- Carbazole as sole nitrogen source (condition 3).

GTIN 20

- Same identification as GTIN17
- Growth inhibition effect with test compound (condition 4 grow faster than 5).
- Grwoth on condition 2 (No growth on GTIN17)

GTIN21

- Full Growth (+++) on m-toluidine, o-toluidine, and aniline
- Brown/pink color was developed on the growth of m-toluidine, and o-toluidine.
- Slight growth on carbazole medium (+), and NH_4Cl (+)
- Faster growth (first growth) on MMN + GGS + carbazole (condition 3) than condition 1 and 2
- Carbazole as sole nitrogen source (condition 3)

GTIN22

- Same identification as GTIN21
- Full grow on NH_4Cl, m-toluidine, p-toluidine, aniline and carbazole (+++)
- Slight growth on o-toluidine, and azodianiline (+)
- Brown/pink color was developed on the growth of m-toluidine, and p-toluidine
- Condition 1, 2, 4, and 5 grow at the same time.
- Condition 3 grow slightly and brown color development
- Carbazole may be used as carbon and nitrogen sources

GTIN23 and 24

- Same identification
- Condition 1, 2, 3, and 4 grow at the same time
- Carbazole may be used as carbon and nitrogen sources

GTIN26

- Similar growth pattern with GTIN16

The most promising cultures appeared to be GTIN17, GTIN18, and GTIN21. Each of these cultures exhibited some ability to utilize carbazole as a nitrogen source so PCR tests were performed with primers designed to detect *carA* genes having sequences related to the *carA* genes of *Sphingomonas* sp. GTIN11, *Pseudomonas resinovorans* CA10, and *Sphingomonas* sp.

CB3; however, none of these cultures yielded a PCR product of the expected size with any of these primer pairs. While amplifications with controls all yielded amplicons of the expected sizes, none of the three candidate cultures (GTIN17, GTIN18, or GTIN21) yielded any PCR amplification products suggesting that the carbazole degradation genes in these strains are unique. Another possible interpretation of the lack of products in PCR experiments with *car* gene primers is that these cultures may not actually metabolize carbazole.

Enrichment cultures were successful in obtaining several mixed cultures that were capable of utilizing organonitrogen compounds as sole nitrogen sources and these mixed cultures were then processed to obtain pure bacterial cultures capable of utilizing organonitrogen compounds as sole nitrogen sources. The potentially useful cultures that were isolated included *Rhodococcus globerulus, Sphingomonas yanoikuyae, Methylobacterium radiotolerans, Gordonia bronchialis,* and *Rhodopseudomonas palustris* that were capable of metabolizing carbazole, aniline, or 4,4'-azodianiline. Of all of the pure microbial cultures listed above, none proved to be worth detailed characterization because no unique enzyme capable of selectively cleaving C-N bonds could be demonstrated. Upon further characterization it was found that these cultures did not selectively cleave C-N bonds in organonitrogen compounds, but would also utilize organonitrogen compounds as sole carbon sources. Therefore none of these cultures offered advantages versus *Sphingomonas* sp. GTIN11: a previously characterized carbazole-degrading microorganism.

The carbazole degradation operon of *Sphingomonas* sp. GTIN11 was cloned and sequenced [8] and contains genes that specify the CarA enzyme (the *carA, carAc,* and *carAd* genes) that is responsible for the cleavage of the first C-N bond in carbazole such that carbazole is converted to 2-aminobiphenyl-2,3-diol. However, an enzyme that selectively cleaves the C-N bond in 2-aminobiphenyl-2,3-diol is lacking. Therefore, obtaining such an enzyme that selectively cleaves the C-N bond in 2-aminobiphenyl-2,3-diol became the key focus of enrichment culture experiments. If an appropriate deaminase enzyme was found that can selectively cleave the C-N bond in 2-aminobiphenyl-2,3-diol the gene(s) that encode the enzyme could be combined with the *carA, carAc,* and *carAd* genes of *Sphingomonas* sp. GTIN11 to form a new metabolic pathway enabling the selective cleavage of both C-N bonds in carbazole.

The most important goal of enrichment culture experiments was to obtain cultures capable of utilizing 2-aminobiphenyl (2-ABP), since 2-aminobiphenyl-2,3-diol is not

commercially available. Hydroxybiphenyl and aminobiphenyl compounds are toxic to many microorganisms and it was challenging to obtain cultures that tolerate exposure to such compounds, let alone are capable of degrading them.

Despite the toxicity of 2-ABP we eventually succeeded in obtaining a mixed culture, SL1-5, that is capable of utilizing 2-aminobiphenyl as a sole nitrogen source. The ability of SL1-5 to utilize 2-aminobiphenyl as a sole nitrogen source was confirmed by at least six independent experiments, all of which showed the expected results for nitrogen-free and nitrogen-added controls. The ability to utilize the nitrogen present in 2-aminobiphenyl implies that the C-N bond in 2-aminobiphenyl is being cleaved. SL1-5 does not utilize 2-aminobiphenyl or biphenyl as a carbon source.

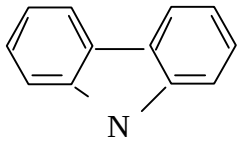
Isolation of 2-ABP Degrading Bacterial Culture

The selective removal of nitrogen from petroleum can complement hydrotreatment for upgrading petroleum and create petroleum products that have better stability. It is important to note that recent investigations reported in the literature show that the typical biochemical pathways for the metabolism of carbazole are not suitable for upgrading petroleum. 2-ABP is converted to 2-aminobiphenyl-2,3-diol by the first step in previously characterized carbazole degradation pathways. It has been speculated that the cleavage of one of the two C-N bonds in carbazole, and related compounds, could be sufficient to allow upgrading of petroleum in combination with hydrotreating. However, the study by Bressler et al. [17, 18] demonstrated that compounds such as 2-aminobiphenyl-2,3-diol are transformed by the hydrotreatment process resulting in ring closure, condensation, and polymerization. Deamination was not observed. The implication of this for this project was that biochemical cleavage of both C-N bonds in carbazole is required for upgrading petroleum and previously characterized biochemical pathways for carbazole biodegradation are not appropriate.

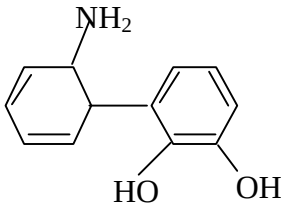
After repeated streaking of the 2-ABP-degrading mixed culture SL1-5 onto MMN agar plates containing a mixture of glucose/glycerol/succinate as a carbon source and 2-ABP as a nitrogen source, eventually a pure culture was obtained that appeared to grow in liquid culture with 2-ABP as the sole nitrogen source. DNA sequence analysis of a portion of the 16S rRNA gene of this culture identified it as a *Pseudomonas* species most closely related (4.6% difference) to *Pseudomonas fluorescens*. This culture was designated as *Pseudomonas* sp. GTIN-G4. The metabolites of 2-ABP, and related compounds, formed by *Pseudomonas* sp. GTIN-G4 were

investigated. The structures of some of the organonitrogen compounds used in these tests are shown in Figure 2.

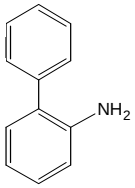
FIGURE 2. STRUCTURE OF ORGANONITROGEN COMPOUNDS USED IN GROWTH TESTS WITH PSEUDOMONAS SP. GTIN-G4



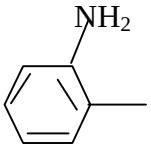
Carbazole



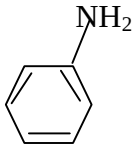
2-aminobiphenyl-2,3-diol



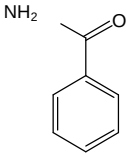
2-aminobiphenyl



ortho-toluidine



Aniline



Benzamide

Pseudomonas sp. GTIN-G4 was isolated based on its ability to grow in a defined mineral salts medium containing 2-aminibiphenyl (2-ABP) as the sole nitrogen source. Initial microbial growth tests utilized a stock solution of 100 millimolar 2-ABP dissolved in ethanol. Growth experiments employed a final concentration of 1 mM 2-ABP in MMN medium incubated aerobically at 30°C. *Pseudomonas* sp. GTIN-G4 grew under these conditions and appeared to be utilizing 2-ABP as a sole source of nitrogen and carbon. After repeated testing of the culture it became clear that ethanol was a good carbon source for the growth of the culture. Subsequent growth tests exclusively used crystals or neet liquids of test compounds rather than stock solutions dissolved in ethanol.

The results of a typical growth test are shown in Table 2 in which the ability of organic and inorganic nitrogen sources were tested to see if they support the aerobic growth of *Pseudomonas* sp. GTIN-G4. The results appear to indicate that carbazole, 2-ABP, o-toluidine, and diphenylamine can be used as nitrogen sources but not as carbon sources. However, unexpected growth was also observed when ethanol was used as a carbon source, but no nitrogen source was provided. *Pseudomonas* sp. GTIN-G4 is a fragile culture that frequently failed to yield growth upon subculturing so that growth tests were repeated multiple times until it could be concluded with confidence that a given condition allowed growth of this culture. This is why it was not initially observed that *Pseudomonas* sp. GTIN-G4 could grow in the absence of a specifically supplied organic or inorganic nitrogen compound. However, repeated testing yielded the results summarized in Table 2. The test substrates 2-ABP, o-toluidine, and diphenylamine were transformed to colored products in growth experiments and these colored products were not observed in sterile controls. It was clear then that *Pseudomonas* sp. GTIN-G4 was metabolizing 2-ABP, o-toluidine, and diphenylamine but it was not clear if these compounds were serving as nitrogen sources for the growth of the culture.

Table 2. Aerobic Growth of *Pseudomonas* sp. GTIN-G4 with Various Nitrogen Sources

Carbon Source	**Nitrogen Source**	**Growth**	**Color of Sample**
None	None	-	clear
ethanol	None	++	gray
ethanol	NH_4Cl	++	gray
ethanol	NH_4NO_3	++	gray
ethanol	carbazole	++	gray
ethanol	2-aminobiphenyl	++	copper/brown
ethanol	o-toluidine	++	light brown
ethanol	diphenylamine	++	slightly purple
None	NH_4NO_3	-	clear
None	carbazole	-	clear
None	2-aminobiphenyl	-	clear

The growth of the culture in the apparent absence of a nitrogen source suggested the possibility that the culture could fix atmospheric nitrogen. To investigate this possibility, growth tests were performed in sealed serum bottles using pure oxygen or air as the headspace gas. Multiple growth experiments were performed and the results are summarized in Table 3, which illustrated that no growth was obtained when pure oxygen was used while air did yield growth supporting the conclusion that *Pseudomonas* sp. GTIN-G4 can fix N_2 gas. However, the failure of the culture to grow in the presence of pure oxygen even when NH_4Cl or NH_4NO_3 were provided raised the possibility that perhaps pure oxygen was toxic to the culture.

Table 3. Growth of *Pseudomonas* sp. GTIN-G4 in the Absence of Nitrogen Gas

Carbon Source	**Nitrogen Source**	**Headspace Gas**	**Growth**
None	None	O_2	-
ethanol	None	O_2	-
ethanol	NH_4Cl	O_2	-
ethanol	NH_4NO_3	O_2	-
ethanol	2-aminobiphenyl	O_2	-
ethanol	o-toluidine	O_2	-
ethanol	diphenylamine	O_2	-
None	2-aminobiphenyl	O_2	-
None	None	Air	-
ethanol	None	Air	++
ethanol	2-aminobiphenyl	Air	++

This possibility was investigated in experiments summarized in Table 4 in which the growth of *E. coli* and *Pseudomonas* sp. GTIN-G4 were compared using various concentrations of oxygen. To achieve the desired concentration of oxygen the sealed serum bottles were first sparged for 2 minutes with pure oxygen then the required volume of gas was removed and replaced with helium. However, growth of *Pseudomonas* sp. GTIN-G4 was not obtained even with 25% oxygen, which is very near the oxygen concentration in air. Therefore some experiments were performed in which the sealed serum bottles were first sparged with helium and then the required volume of gas was removed and replaced by oxygen in order to achieve the

desired oxygen concentration. As shown in Table 4, *Pseudomonas* sp. GTIN-G4 is sensitive to exposure to pure oxygen while *E. coli* is not.

Table 4. Growth of *Pseudomonas* sp. GTIN-G4 and *Escherichia coli* with Various Concentrations of Oxygen.

Bacterial Culture	**Media composition**	**Headspace Gas**	**Growth**
GTIN-G4	MMN,ethanol, NH_4NO_3	100% O_2	-
GTIN-G4	MMN,ethanol, NH_4NO_3	50% O_2	-
GTIN-G4	MMN,ethanol, NH_4NO_3	25% O_2	-
GTIN-G4	MMN,ethanol, NH_4NO_3	20% O_2, He first	+
E. coli	L Broth	100% O_2	++
E. coli	L Broth	50% O_2	++
E. coli	L Broth	25% O_2	++
E. coli	L Broth	20% O_2, He first	++

To more fully define the growth capabilities of *Pseudomonas* sp. GTIN-G4 experiments were performed using cultures initially sparged with helium and subsequently adjusted so the gas composition was 20% oxygen. The results of such tests are shown in Table 5 where it is observed that when growth of *Pseudomonas* sp. GTIN-G4 occurs in the absence of N_2 gas colored metabolites of 2-ABP, o-toluidine, or diphenylamine are not observed. Growth of *Pseudomonas* sp. GTIN-G4 with air is required for the production of colored products derived from organonitrogen substrates.

Table 5. Growth of *Pseudomonas* sp. GTIN-G4 with Various Organonitrogen Compounds in the Absence of nitrogen Gas

Nitrogen Source	**Headspace Gas**	**Growth**
NH_4Cl	20% O_2/80% He	++, clear
NH_4NO_3	20% O_2/80% He	++, clear
NH_4NO_3, 2-aminobiphenyl	20% O_2/80% He	++, clear
NH_4NO_3, o-toluidine	20% O_2/80% He	++, clear
NH_4NO_3, diphenylamine	20% O_2/80% He	++, clear
2-aminobiphenyl	20% O_2/80% He	-, clear
o-toluidine	20% O_2/80% He	-, clear
diphenylamine	20% O_2/80% He	-, clear
2-aminobiphenyl	Air	++, copper/brown
o-toluidine	Air	++, light brown
diphenylamine	Air	++, slightly purple

Pseudomonas sp. GTIN-G4 metabolizes 2-aminobiphenyl, diphenylamine, o-toluidine, and aniline in liquid cultures to produce colored metabolites that are not observed in sterile controls or with cultures of other bacteria incubated with these substrates. GC-MS analyses identified some unique products that appear to result from the metabolism of *Pseudomonas* sp. GTIN-G4. Mass spectral analysis was performed on aqueous extract samples in an attempt to identify potential biodegradation byproducts formed by *Pseudomonas* sp. GTIN-G4 from 2-aminobiphenyl, diphenylamine, and o-toluidine. Compounds were found in all three sample types suggesting a substitution of a hydrogen on the nitrogen atom with a CHO group. The proposed structure was confirmed for the diphenylamine by comparing it to a standard of N,N-diphenylformamide. The structures related to the 2-aminobiphenyl and o-toluidine samples are

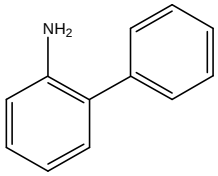
based on the similarities of their mass spectra to that of N,N-diphenylformamide. GC-MS data illustrating the formaldehyde-modified derivatives of 2-ABP, o-toluidine, and diphenylamine produced by *Pseudomonas* sp. GTIN-G4 are shown in Figures 3, and 4 respectively. Similar results were also obtained for the metabolism of aniline by *Pseudomonas* sp. GTIN-G4 (data not shown).

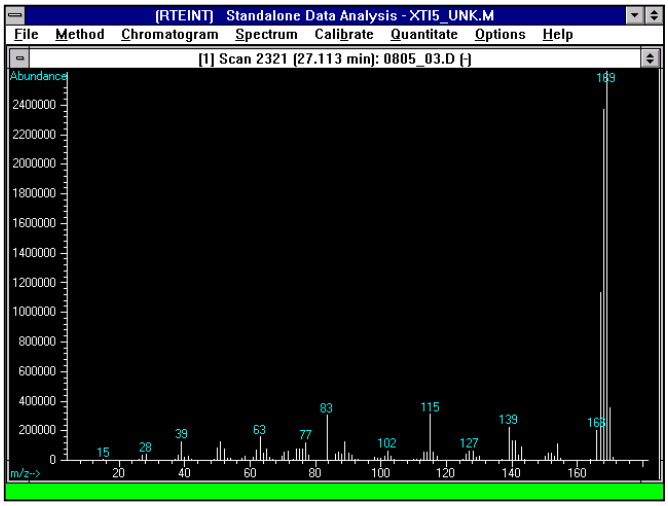
The microbial modification of 2-ABP by the addition of a formaldehyde group is without precedent, and is an interesting biochemical ability. Since there are no other publications describing the biodegradation of 2-ABP the most similar studies with which to compare these data are studies of the biodegradation of aniline. The predominant mechanism for the biodegradation of aniline is by aerobic oxidation whereby aniline dioxygenase oxidatively deaminates aniline yielding catechol [19, 20]. However, minor degradation products consisting of acylated or formylated derivatives of aniline were also observed derived from anaerobic metabolism [21]. The derivitization of compounds is commonly observed in the microbial degradation of aromatic compounds by anaerobic microorganisms. The anaerobic degradation of aniline involves the carboxylation of aniline to 4-aminobenzoate which is converted to 4-aminobenzoyl-CoenzymeA and subsequently deaminated and further degraded [22, 23].The metabolism of toluene under anaerobic conditions by the denitrifying bacterium *Thauera aromatica* is initiated by the formation of benzylsuccinate from the combination of toluene and fumarate catalyzed by benzyl succinate synthase [24]. Similarly, the biodegradation of phenol by *Thauera aromatica* involves phosphorylation and carboxylation [25].

No culture capable of utilizing 2-ABP as a sole source of nitrogen and/or carbon was obtained from enrichment culture experiments, nor are there reports of the biodegradation of 2-ABP in the literature. Aminobiphenyl compounds are toxic and carcinogenic [26]. 4-aminobiphenyl has been shown to react with DNA and cause mutagenesis. 2-aminobiphenyl and 3-aminobiphenyl are less mutagenic and their primary mechanism of toxicity is thought to be reactions with proteins and interference with the electron transport system in cells [26, 27]. This serves to highlight the toxicity of 2-ABP and its importance as an environmental contaminant. While the formylation of aniline is known this is the first report of the biodegradation of 2-ABP. The formation of N-formamide derivatives of 2-aminobiphenyl, diphenylamine, aniline, and o-toluidine is a unique capability of *Pseudomonas* sp. GTIN-G4.

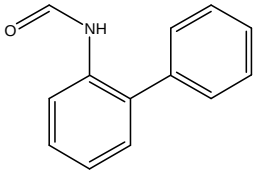
FIGURE 3. MASS SPECTRA OF 2-AMINOBIPHENYL AND A METABOLITE PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4.

2-aminobiphenyl





2-formamidylbiphenyl



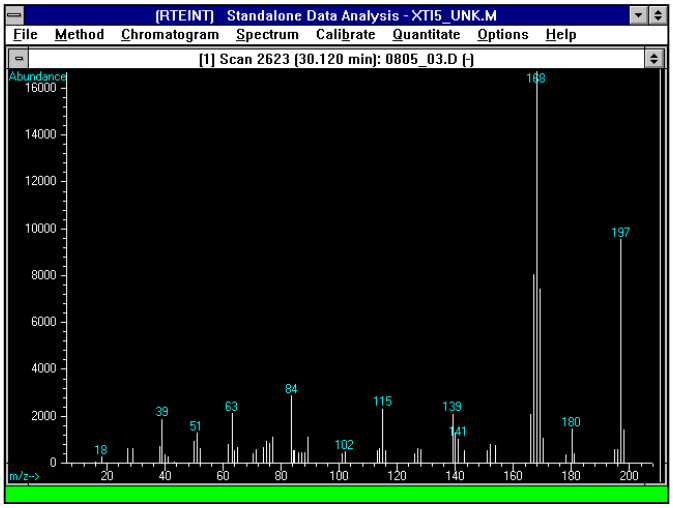
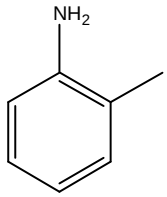
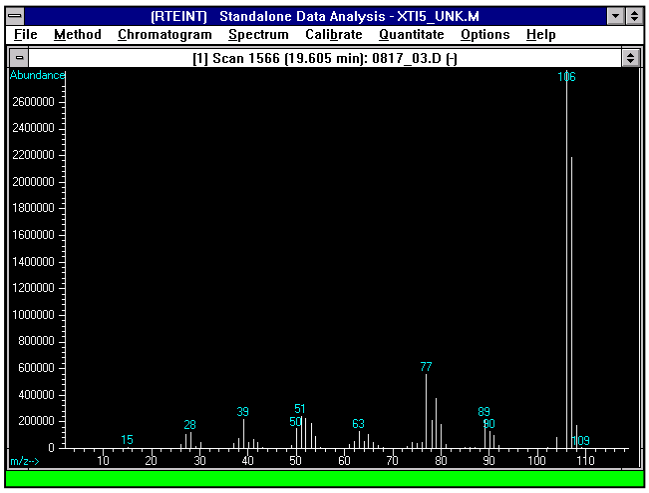


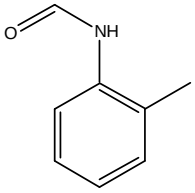
FIGURE 4. MASS SPECTRA OF O-TOLUIDINE AND DIPHENYLAMINE AND THE CORRESPONDING METABOLITES PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4.

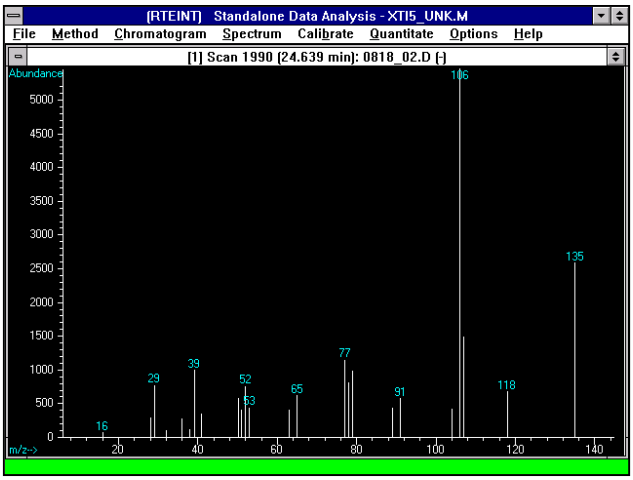
o-toluidine



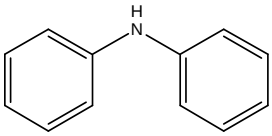


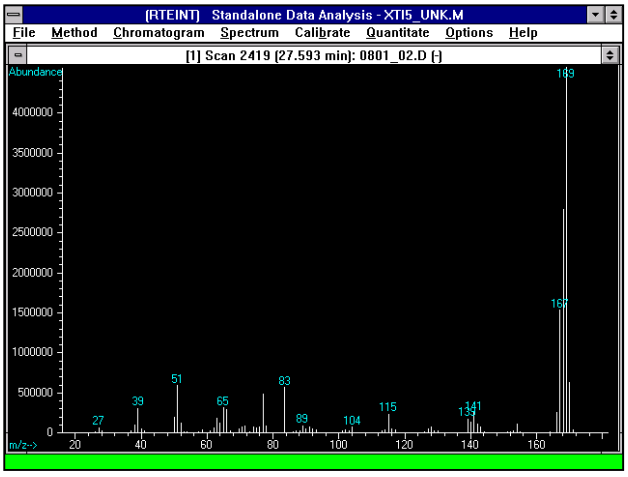
2-methylbenzylformamide



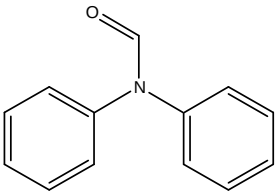


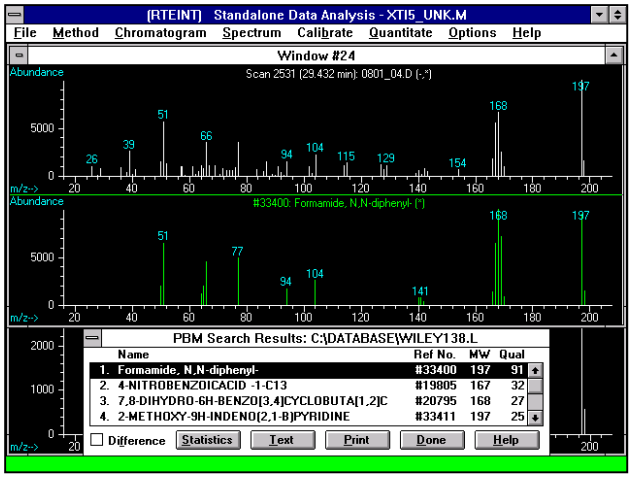
Diphenylamine





N,N-diphenyl formamide





Gibbs Assay to Investigate Substrate Range for C-N Bond Cleaving Enzymes

Investigations of biodesulfurization of the aromatic sulfur compound dibenzothiophene benefited from the use of the Gibbs assay. The Gibbs assay employs 2,6-dichloroquinone-4-chloroimide to react with aromatic hydroxyl groups and form colored products that can be measured spectrophotometrically. It is not yet known if the Gibbs assay can be used to measure reaction products resulting from the metabolism of organonitrogen compounds, so that was investigated. Various compounds related to the metabolism of organonitrogen compounds were tested to determine if they would react with the Gibbs reagent. Model compounds tested included: 4-hydroxybenzoic acid, 4-hydroxycoumarin, 3-hydroxy isoquinoline, 3,4-dihydroxybenzoic acid, dihydroxyfumaric acid hydrate, 2-hydroxybiphenyl, 2-aminobiphenyl, 4,4'-dihydroxybiphenyl, 2,2'-dihydroxybiphenyl, coumarin, catechol, carbazole, and quinoline. All of the compounds that contain an aromatic hydroxyl group produced a reaction, but the color of the reacted products varied and blue, purple, brown, gray, and orange reaction products were observed. The carbazole and quinoline substrates did not react.

To test the use of the Gibbs assay to investigate the substrate range of C-N bond cleavage enzymes, experiments were performed with growing and resting cells of *Sphingomonas sp.* GTIN11 and *Pseudomonas ayucida* IGTN9. *Sphingomonas sp.* GTIN11 grew with carbazole as the sole nitrogen source, but no products that reacted with the Gibbs reagent were observed. Similarly, resting cells of *Sphingomonas sp.* GTIN11 incubated with carbazole failed to yield products that reacted with the Gibbs reagent. This is probably because carbazole is metabolized intracellularly and intermediate degradation products containing aromatic hydroxyl groups are not released into the medium. Fresh, quinoline-grown cells of *Pseudomonas ayucida* IGTN9 were washed and concentrated, then resuspended for 2 hours in MMN media containing various substrates. After the 2-hour incubation the Gibbs assay was performed and spectrophotometric scans were performed to identify absorption wavelengths that were best suited for the detection/quantification of various reaction products. IGTN9 cells without any substrate yielded a light brown color, while in the presence of quinoline a metabolite with an absorption maximum at 330 nm (light blue) was observed. In the presence of carbazole no reaction was observed, probably indicating that carbazole does serve as a substrate for the quinoline-metabolizing enzymes. In the presence of 2-aminobiphenyl, a metabolite with an absorption maximum at 418 nm (blue/green) was observed while the 2-aminobiphenyl control yielded a blue reaction with a

different absorption maximum. This indicates that 2-aminobiphenyl may be metabolized by the enzymes in IGTN9 even though this compound does not serve as a sole source of nitrogen for this culture. Melamine and diphenylamine yielded no reaction products while aniline yielded a purple color with an absorption maximum at 380 nm.

Directed Evolution of a Deaminase/Amidase Gene

In addition to using enrichment culture experiments to obtain an enzyme capable of cleaving the C-N bond in 2-aminobiphenyl we also tried to modify known deaminases, amidases, and dioxygenases using directed evolution. The objective of this task was to obtain an enzyme capable of selectively cleaving the C-N bond in 2-aminobiphenyl-2,3-diol by modifying known enzymes that act on structurally similar substrates.

When this project started no suitable enzyme for the selective cleavage of C-N bonds in 2-aminobiphenyl or related compounds was known to be available. Accordingly, experiments were performed with amidase, deaminase, and dioxygenase enzymes for which genetic information was available and that acted on substrates somewhat structurally related to 2-aminobiphenyl. Initially the candidate enzymes employed were the benzamide amidase from *Rhodococcus* sp. MP50 [28] and the melamine deaminase (*triA*) from *Pseudomonas* sp. NRRL B12227 [29]. Both of these enzymes are encoded by single genes which simplifies the mutagenesis steps required in directed evolution experiments.

The amidase from *Rhodococcus* sp. MP50 was amplified by PCR using primers based on published DNA sequence data and cloned into the *E. coli* vector pGEMT-Easy. This DNA fragment was subsequently cloned into *E. coli* expression vector pQE80, and *Rhodococcus* expression vector pYgal-K2. Although benzamide is structurally not as close to the target substrate, 2'-aminobiphenyl-2,3-diol, as desired, it was chosen as the alternative nitrogen source in the minimal media because it is a known substrate of amidase MP50 and serves as an useful substrate to test for enzyme activity. The concentrations of benzamide used in the minimal media assay ranged from 1 to 32 mM. The cells were first induced with 0.1 mM of IPTG for 1 hr before they were inoculated into the minimal media to ensure that the innoculum possessed the amidase enzyme required for transforming benzamide. The minimal media contains 0.01 mM of IPTG. This low concentration ensures that the cells will not expend too much resource producing the enzymes and will still be able to replicate.

From the results of the assay (see Figure 5) it was found that the optimal benzamide concentration for cell growth was about 8 mM. At higher benzamide concentrations, the growth of the cells slowed down. At a benzamide concentration of 32 mM, there was no growth detected, indicating that the benzamide may be toxic to the cells at high concentrations.

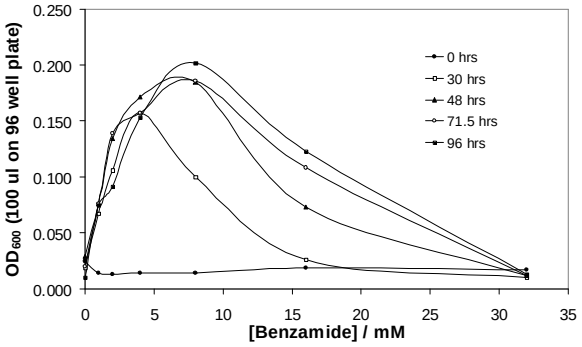


FIGURE 5. OPTICAL DENSITY OF CELLS IN MINIMAL MEDIA WITH BENZAMIDE AS SOLE NITROGEN SOURCE.

From the results of the SDS-PAGE and minimal media assay, it can be concluded that the amidase enzyme can be functionally expressed in *E. coli* DH5 alpha. By changing the nitrogen source to the desired substrate, the minimal media assay can be also used as a selection method for the directed evolution of the amidase enzyme.

The activity of amidase on 2-aminobiphenyl (2-ABP) was tested by using the minimal growth media containing 2-aminobiphenyl as the sole nitrogen source. In this set of experiments, the highest concentration of 2-ABP used was restricted to 1 mM due to the low solubility of the substrate in aqueous solution. The optical density of the cell culture is shown in Figure 6. From the figure, it can be seen that the sample containing ammonium and saturated with 2-ABP grew at a much slower rate of than that with ammonium only. This shows that 2-ABP slows the growth of the cells but is not lethal to the cells. Cells supplemented with 2-ABP as the sole nitrogen source did not show signs of growth. This might be due to two factors. Firstly, the amidase enzymes do not act on 2-ABP and hence the cells cannot catabolize 2-ABP. Secondly, the low solubility of 2-ABP (about 1.38 mM at 25°C) may have resulted in the lack of nitrogen source for cell growth.

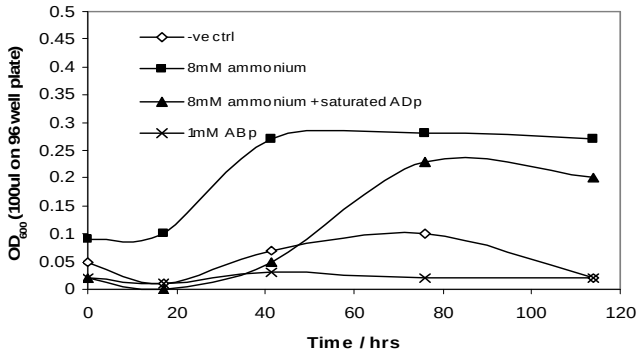


FIGURE 6. OPTICAL DENSITY OF CELLS WITH TIME IN 2-AMINOBIPHENYL MINIMAL GROWTH MEDIUM.

To engineer the amidase MP50 enzyme to hydrolyze 2'-aminobiphenyl-2,3-diol (2'-ABPD), random mutations were introduced into the amidase MP50 gene by error prone PCR (epPCR). The mutant genes were then cloned and transformed into *E. coli* DH5 alpha and put through selection on a minimal media plate with 2-aminobiphenyl (2-ABP) as the only source of nitrogen. As such, only cells producing mutant amidase enzymes capable of hydrolyzing 2-ABP to produce ammonia can obtain nitrogen necessary for growth. 2-ABP was used as a substitute for 2'-ABPD in these experiments as 2'-ABPD is not available commercially.

epPCR followed by minimal media selection was performed twice on the amidase gene but no positive clones were found. This is because 2-ABP does not possess the alpha-carbon double-bonded to an oxygen atom, as is the case of an amide. This C=O bond is essential in the amidase hydrolysis mechanism. As it is almost impossible to completely change the enzymatic mechanism of an enzyme, other candidate enzymes were investigated for use in directed evolution experiments.

The melamine deaminase encoded by the *triA* gene of *Pseudomonas* sp. NRRL B12227 is capable of selectively cleaving the C-N bonds in melamine, ammeline, and ammelide. Therefore, it may be possible to adapt this deaminase to metabolize 2-aminobiphenyl. We demonstrated that *Pseudomonas* sp. NRRL B12227 can grow in defined mineral salts media with melamine, ammeline, or ammelide serving as sole nitrogen sources. The *triA* gene was cloned into *Rhodococcus-E. coli* genetic vectors to see if the *triA* gene could confer the ability to cleave C-N bonds to an *E. coli* or *Rhodococcus* host. Most *E. coli* cultures won't grow in mineral salts media because of requirements for one or more vitamin or amino acid. However, *E. coli* strain BL21 is

a prototroph and can grow in a defined mineral salt medium, such as the MMN media used to test for the utilization of organonitrogen compounds. The *triA* clone was also included in directed evolution experiments, but no cultures containing the triA gene were obtained from directed evolution experiments that could utilize 2-ABP.

The deamination of 2'-aminobiphenyl-2,3-diol (2-ABPD) by the aniline dioxygenase (AtdA), from *Acinetobacter* sp. strain YAA [19] was also investigated. AtdA is a multicomponent enzyme isolated from *Acinetobacter* sp. Strain YAA involved in the simultaneous deamination and oxygenation of aniline. The gene encoding for AtdA was found to have 5 open reading frames (ORFs), atdA1-A5.

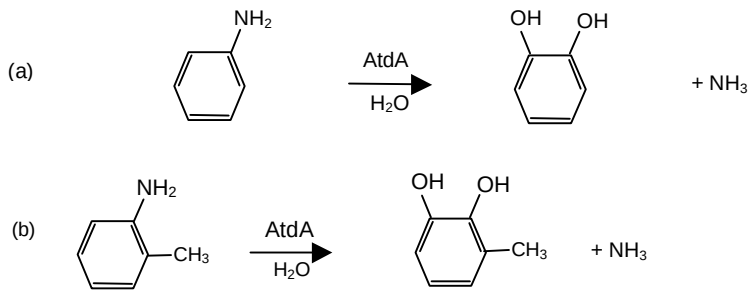


FIGURE 7. DEAMINATION AND OXYGENATION OF (A) ANILINE AND (B) O-TOLUIDINE BY ATDA.

The plasmid pAS93 [19], consisting of the five *atdA* open reading frames (ORFs) cloned into a pUC19 vector, was obtained from Dr Takeo in the Department of Applied Chemistry, Himeji Institute of Technology, Japan. *E. coli* JM109 was used as host for the pAS93 plasmid.

Besides aniline, AtdA was found to be able to oxidize o-toluidine, which has an additional methyl side chain at the ortho position (Figure 7). As 2'-ABPD can be viewed as an aniline molecule with a bihydroxylated phenyl attached at the ortho position, there is a potential that the AtdA enzyme can remove the amine group.

When the *E. coli*/pAS93 cells were grown in a minimal media with 2-ethyl aniline (2-EA), methyl anthrinilate (MA), or 2-ABP as the sole source of nitrogen, it was found that AtdA was only able to deaminate 2-EA and use the released ammonia as a source of nitrogen (Figure 8). The larger ortho-substituent groups on ME and 2-ABP may not be able to fit into the binding

pocket of the AtdA enzyme and hence no deamination activity was detected. Thus the key to engineering AtdA to accept these substrates may lie in increasing the size of the enzyme binding pocket.

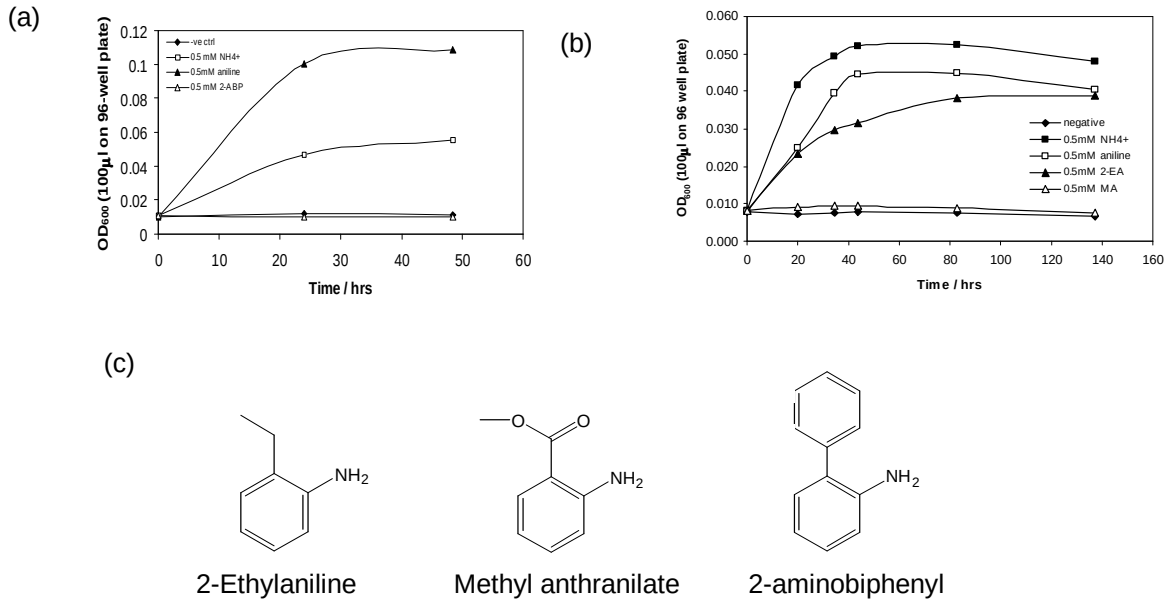


FIGURE 8. MINIMAL MEDIA ASSAY OF pAS93 TRANSFORMED *E. COLI* DH5α WITH (A) 2-ABP, (B) 2-EA AND MA. (C) CHEMICAL STRUCTURES OF 2-EA, MA AND 2-ABP.

The alpha subunits of Rieske non-heme dioxygenases often determine substrate specificity. Using the protein sequence alignment program available on the Biology Workbench (http://workbench.sdsc.edu), it was found that AtdA3 has 27.4 % homology with the alpha subunit of naphthalene 1,2-dioxygenase (NDO), a Rieske non-heme iron dioxygenase (RDO) enzyme that oxidizes naphthalene. Hence, it is highly possible that the substrate specificity of AtdA lies in the AtdA3. The crystal structure of naphthalene dioxygenase was found previously by Kauppi et al. [30]. The alignment of AtdA3 with NDO using ALIGN program on the Biology Workbench is shown in Figure 9. In the figure, the residues highlighted in blue represent the residues in NDO coordinating the Rieske center [2Fe-2S] of the enzyme. The residues highlighted in light gray are residues in NDO coordinating the non-heme mono-nuclear iron ion in the active site. Asp205 of NDO (in dark gray) provides a direct route of electron transfer between one Rieske center of an alpha-subunit and mononuclear iron in the adjacent alpha-subunit. It was found that AtdA3 possesses all of these key residues as well. This makes NDO a

suitable candidate to model AtdA3 against. Extensive modeling and simulation were performed on the AtdA3 subunit to identify key residues that can be mutated to alter the substrate specificity of aniline dioxygenase.

```
1NDO_E        MNYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWLFLTHDSLIPAPGDYVTAKMG
              :.  :...  .::  ... .. .  .:: :.  ::  ::.:: : : ::   :: :..::
Atd_A3        MKTINQLI--QSGRVHRK-VYTEASIFQAEMDKIFQANWVFLLHASQIPKLDDYQTVRMG
                        10         20        30        40        50

                     70        80        90       100       110       120
1NDO_E        IDEVIVSRQNDGSIRAFLNVCRHRGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPFE
                 .:: :..:  ..:.:: : ::  .   ..::.: :.: ::: : ..:.   .:
Atd_A3        GRPLIVVRKGDDEFQALLNRCPHRGAKVCRNDSGNSKTFTCPYHGWKFRNSGKAFVIP-G
               60        70        80        90       100       110

                    130       140       150       160       170
1NDO_E        KDLYGESLNKKCLGLKEVARVESFHGFIYGCFDQEAPPLMDYLGDAAWYLEPMFKHSGG-
               . ::::..:  ...  . ::::..::...  ...:  : ..::.:  :..  . :.::
Atd_A3        ANAYGEGFDKDNFSMTAIPRVESYRGFVFATSNENAVSLEEHLGSARQYIDEWLAHQGGE
               120       130       140       150       160       170

            180       190       200       210       220       230
1NDO_E        LELVGPPGKVVIKANWKAPAENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPEGAG
              ...     .  :: :::   .: .::.::: ..: : :.    . .   :.. .   : .
Atd_A3        IKVSKSVQRYEIKCNWKLVFDN-AGDGYHVPFSHQSLLQ----MTTLRYGGGDIQYFGNA
               180       190        200       210           220       230

            240       250       260       270       280         290
1NDO_E        LQMTSKYGSGMGVLWDGYSGVHSADLVPELMAFGGAKQERLN--KEIGDVRARIYRSH--
                 .. : :. .: .:.: .   :  ::.   .:  :.: .  .:  ....:   :.
Atd_A3        ----DETGMGLYALGNGHSVI---DQRPEMHKESGWDQQRPQPGRESYETHVRNNSSQPA
                        240          250       260       270       280

                          300       310       320        330       340
1NDO_E        ----------LNCTVFPNNSMLTCSGVFKVWNPIDANTTEV-WTYAIVEKDMPEDLKRRL
                        .: ..:::  .:  .. ..: .::..: : . :  ...  :  :    :.
Atd_A3        RDLERAVGAGMNLNIFPN--LLLIGNQIQVIDPISVNETVLHWHATLLAGDNEELNAIRM
                 290       300         310       320       330       340

                 350       360       370       380         390       400
1NDO_E        ADSVQRTFGPAGFWESDDNDNMETASQNGKKYQSRDS--DLLSNLGFGEDVYGDAVYPGV
                 .:. :   :  : ::  :.:.  :.: . . .    :.  ... ::.    : :  :
Atd_A3        --RTQEDFPIMG--EVDDVANFESC-QEGLETMPEIEWIDFSRHMNEGEN----ACYQDV
                     350         360        370       380           390

                   410       420       430       440
1NDO_E        VGKSAIGETSYRGFYRAYQAHVSSSNWAEFEHASSTWHTELTKTTDR
              . ..  .:   : ..          ..: ..  : .        . ..
Atd_A3        IQHKPTSEIHSRHYF---------DTWLQLMSAVNK-----ENQSEV
                 400                 410       420
```

FIGURE 9. ALIGNMENT OF ATD A3 WITH NAPTHTALENE DIOXYGENASE (1NDO_E).

A goal in this task was to relax the substrate specificity of AtdA by rational design so that the enzyme can act on larger substrate molecules such as 2-ABP and ultimately, 2'-ABPD. This can be accomplished by first using a homology modeling program, Insight II, to model the structure of AtdA against known enzyme structures. With the model, the key amino acid residues in the binding pocket of the enzyme can then be identified with other computer programs like

Molecular Operating Environment (MOE). The aim is to reduce the size of the residues in the binding pocket so that larger substrate molecules will be able to enter the binding pocket. The residues in the actual enzyme can then be changed by site directed mutagenesis.

Cloning of aniline dioxygenase operon into expression vectors

The SalI fragment from the pAS91 plasmid was cloned into the pTrc99A expression vector (pTrcA-2) and transformed into *E. coli* BL21 (DE3). The expression level of AtdA cloned in pTrc99A compared to that in pUC19 is shown in Figure 10. From the figure, it can be seen that the expression level of AtdA subunits is higher in pTrc99A than pAS93. With pTrcA-2, all the subunits of the AtdA operon can be seen except AtdA5 (~37.2 kDa), the last subunit in the operon.

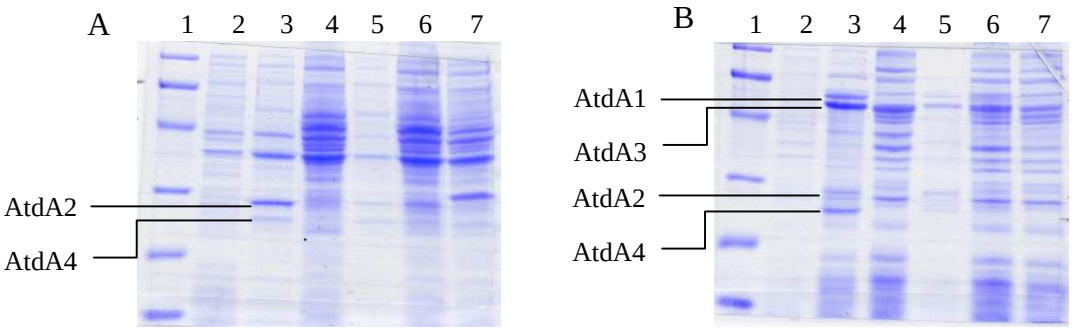


FIGURE 10. SDS-PAGE ANALYSIS OF ATDA EXPRESSION BY pTRCA-2 AND pAS93 IN BL21. (A) SOLUBLE FRACTION (B) INSOLUBLE FRACTION.

Lane: (1) Protein marker (from top to bottom: 97.4, 66.2, 45.0, 31.0, 21.5, and 14.4 kDa), (2) pTrcA-2 before induction, (3) pTrcA-2 after 4 hr induction with 1mM IPTG, (4) pTrc99A (negative control), (5) pAS93 before induction, (6) pAS93 after 4 hr induction with 1mM IPTG, (7) pUC19 (negative control).

Overlap extension PCR was then used to introduce AvrII and AgeI restriction sites before and after the atdA3 gene respectively, to form the plasmid pTA1-1. SDS-PAGE analysis of IPTG induced *E. coli* JM109 with pTA1-1 showed that the expression level of AtdA in pTA1-1 is similar to that of pTrcA-2 (Figure 11).

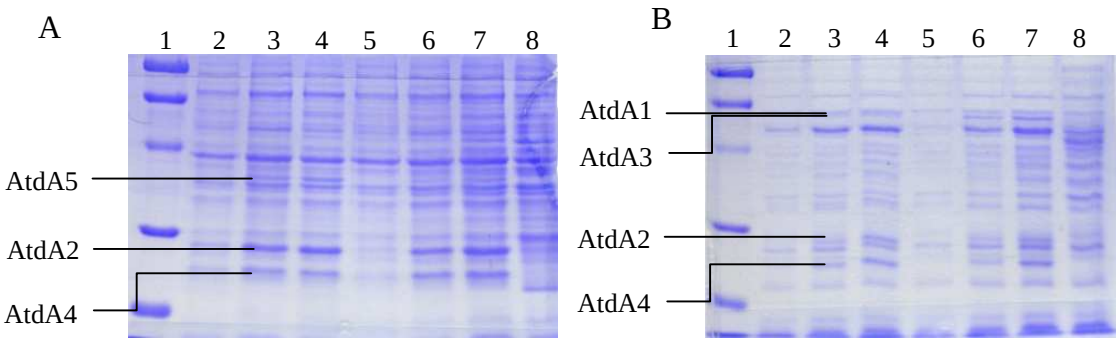


FIGURE 11. SDS-PAGE ANALYSIS OF ATDA EXPRESSION BY pTA1-1 AND pTRCA-2 IN *E. COLI* JM109 INDUCED WITH 1 MM IPTG. (A) SOLUBLE FRACTION (B) INSOLUBLE FRACTION.

Lane: (1) Protein marker (from top to bottom: 97.4, 66.2, 45.0, 31.0, 21.5, and 14.4 kDa), (2) pTA1 before induction, (3) pTA1 after 2 hr induction, (4) pTA1 after 4 hr induction, (5) pTrcA-2 before induction, (6) pTrcA-2 after 2 hr induction, (7) pTrcA-2 after 4 hr induction, (8) pTrc99A (negative control).

The AvrII restriction site in pTA1-1 was then mutated to EcoRI restriction site, while the other EcoRI restriction sites on pTA1-1 were removed by introduction of silent mutations. The resulting plasmid, pTA2-3, gave similar AtdA expression levels as pTA1-1 in *E. coli* JM109 (data not shown). To improve the ligation efficiency of *atdA3* into the vector during directed evolution work, the AgeI restriction site at the 3' end of *atdA3* in pTA2-3 was switched to SalI, which is a more commonly used restriction site, using the Quikchange XL mutagenesis kit. At the same time, the SalI restricition sites flanking the entire *atdA* operon were switched to AgeI using the same method (Figure 12). The resulting plasmid is pTA3-1.

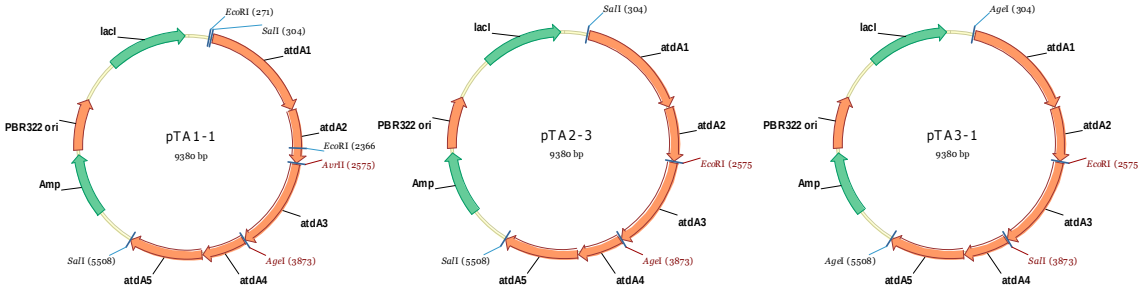


FIGURE 12. PLASMID MAPS OF pTA1-1, pTA2-3, AND pTA3-1.

Solid phase screening

Two methods of detecting the phenolic products of the aniline denitrogenation reaction were studied. The first method is the Gibb's reagent solid phase screen adapted from with slight modifications, as described in the Experimental section. Figure 13 shows the result after JM109 cells carrying pTrc99A (empty vector) and pTrcA2 were induced with IPTG, exposed to 2-ethylaniline (2-EA), and subsequently Gibb's reagent. The background on the membrane turns blue after incubation on the Gibb's reagent due to the presence of 2-EA. The area around JM109 cells carrying the empty vector remains blue whereas the area surrounding JM109/pTrcA-2 turns to a purplish pink color due to the presence of the hydroxylated product. The products formed did not mix with the substrate because of the diffusion limitation on the solid phase. As such, active mutants can easily be identified via a visual screen of the plate. This allows for a high throughput screen without the laborious step of colony picking.

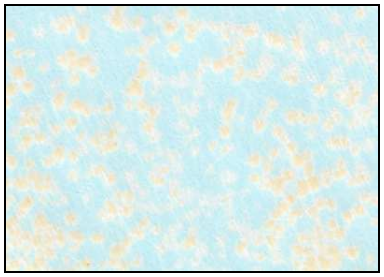
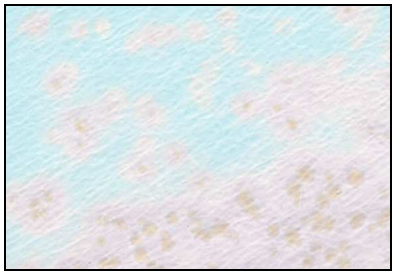
A  B 

FIGURE 13. NYLON MEMBRANE WITH A) JM109/pTRC99A (NEGATIVE CONTROL) CELLS; B) JM109/pTRCA-2 AFTER INCUBATION ON GIBB'S REAGENT PLATE.

The sensitivity of the Gibb's reagent screen was verified by its ability to identify *E.coli* JM109/pTrcA-2 from a mixture of JM109/pTrcA-2 and JM109/pTrc99A. Separate overnight cultures of JM109/pTrcA-2 and JM109/pTrc99A were first diluted 10^4 folds and then plated onto an LB plate at a ratio of 1:10. After overnight incubation in 37°C, the colonies were membrane transferred onto an LB plate with 1mM IPTG for 8 hrs, and then incubated with 2-ethylaniline for another 3 hrs. As shown in the Figure 14, the active clones can be identified from those with empty vectors by the pink halo formed around the colonies through the reaction of Gibb's reagent with the phenolic products from 2-EA dioxygenation. This assay is sensitive enough to identify active colonies among a population of approximately 500 colonies per plate.

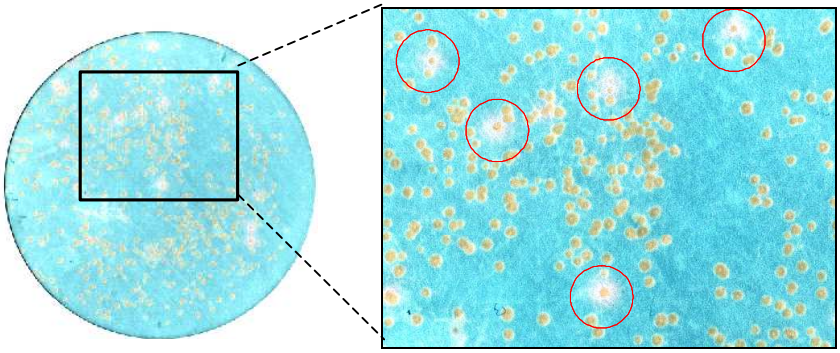


FIGURE 14. E.COLI JM109 WITH pTRCA-2 AND JM109 WITH pTRC99A (EMPTY VECTOR) ON A NYLON MEMBRANE AT A RATIO OF 1:10 AFTER INCUBATION WITH 2-EA AND GIBB'S REAGENT SCREEN. Colonies containing the active AtdA enzyme can be identified by the pink halos around them.

The second method is of detecting the phenolic compound is via the auto-oxidation of these compounds. Hydroxylated aromatic compounds are unstable in the presence of light and spontaneously polymerize to form colored compounds, thus allowing the identification of phenolic products around active clones. In this method, colonies were first picked and streaked onto LB plates containing 0.5 mM IPTG and incubated overnight at 37°C so that when the colonies are formed they will already be expressing the AtdA protein. This method reduces the overall assay time, as the colonies do not need to be further incubated with IPTG before the substrates are introduced. The membranes were then transferred to M9 minimal media agar plates containing 0.5 mM IPTG and 5mM of the desired substrates and incubated at 30°C. Using this method, the characteristic colored ring around the colonies appeared after only 5 hrs of

incubation with 2-methylaniline and 2-ethylaniline (Figure 15). Previously, incubation of uninduced colonies with IPTG and the substrate took at least 12 hr before the colored rings are formed.

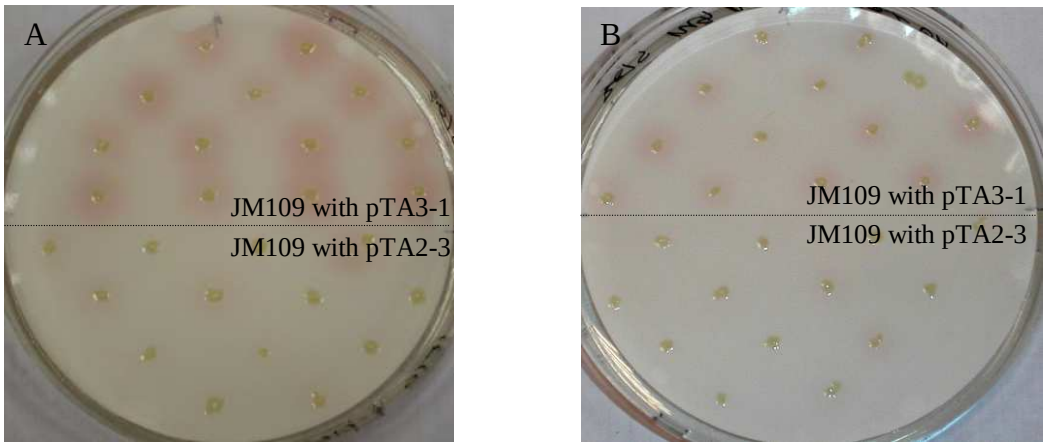


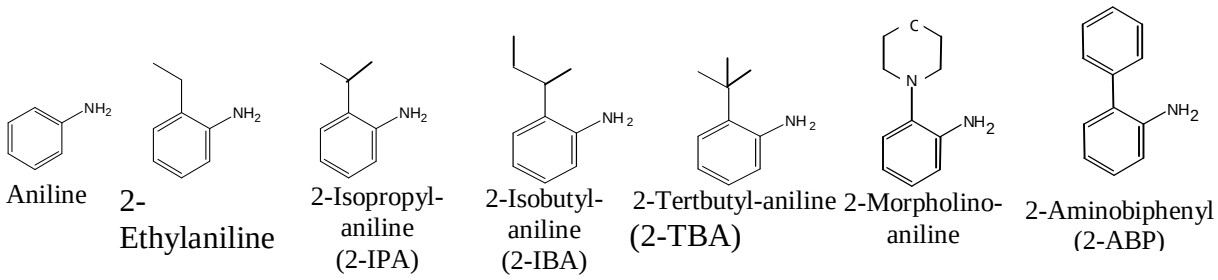
FIGURE 15. E.COLI JM109 WITH pTA3-1 OR pTA2-3 AFTER 5 HRS OF INCUBATION WITH (A) 2-METHYLANILINE, AND (B) 2-ETHYLANILINE.

Activity of AtdA from plasmid pTA3-1

An interesting observation that can be made from Figure 15 is that the colored rings around the JM109 colonies harboring pTA3-1 were larger and had greater intensity than those of pTA2-3, even though the only difference between these plasmids is the restriction sites used. The change in restriction sites may have affected the ribosomal binding site of the plasmids, resulting in increased functional expression of the protein.

Substrate specificity of aniline dioxygenase

E. coli JM109 with pAS93 was grown in liquid M63 minimal media with the following substrates as sole N source. Only substrates that can be denitrogenated by AtdA will be able to supply N for cell growth.



It was found that only aniline and 2-EA were able to support growth of the cells. Cells grown in complete M63 liquid media with ammonium sulfate did not convert 2-IPA as well. However, AtdA may still be active on 2-IPA or larger substrates but its activity may not be high enough to support the growth of the cells. To confirm the substrate range of AtdA, the Gibb's reagent screen was used to test the activity of AtdA on various aromatic amines having meta-position substituents. As shown in Figure 16, AtdA does not accept substrates with meta-position subtituents larger than an ethyl group as there are no colored halos around 2-IPA or any larger substrates. Hence directed evolution of AtdA to accept larger substrates was performed.

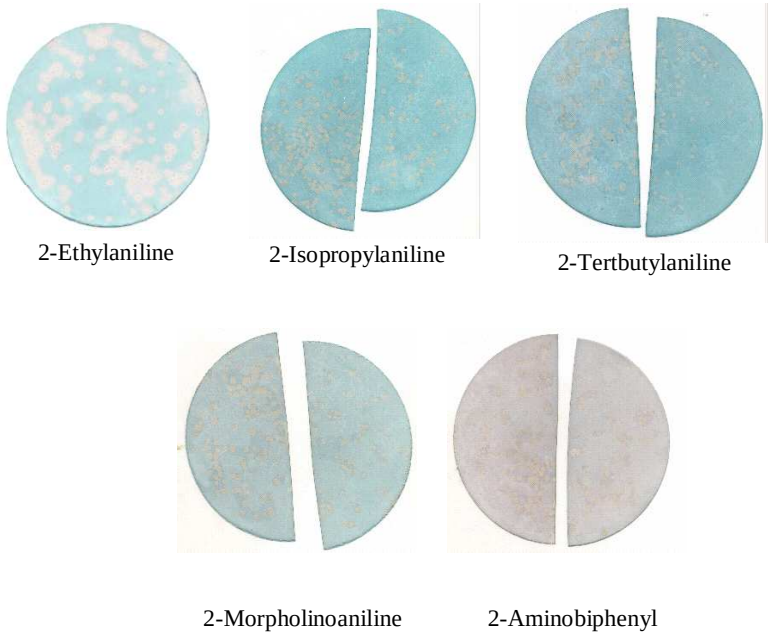


FIGURE 16. E.COLI JM109/pTRCA-2 COLONIES AFTER INCUBATION WITH VARIOUS SUBSTRATES AND GIBB'S REAGENT: RANDOM MUTAGENESIS OF *ATDA3*

The mutation rate of a gene is important in the directed evolution of AtdA3 because of the restrictions imposed by the library size that a screening process can handle. Typically, a

manageable library size for a screen is ~10^4. If the mutation rate is too high, the library size becomes large and it may not be possible to screen every one of the mutants. However, a low mutation rate may not be enough to cover the sequence space necessary to create an active mutant.

The $MnCl_2$ concentration in error prone PCR (epPCR) plays an important role in controlling the mutation rate of the gene. It was found that 0.2 mM of $MnCl_2$ in epPCR produced approximately 50% inactive mutants. This corresponds to about 1 amino acid change per gene. As the number of mutations normally accessible for each residue in an epPCR is 6, the number of possible mutants from AtdA3 (426 amino acids long) is about 2500. To have a comprehensive screen of all the mutants, the library size should be about 10 times the number of possible mutations. This gives a library size of ~2.5×10^4, which is a manageable number.

Homology modeling

Another approach to widen the substrate specificity of AtdA3 is via a semi-rational approach. In this method, region of interest on AtdA3 is narrowed down to the residues defining its binding pocket. Mutagenesis of the residues can reduce the steric hindrance in the binding pocket and hence allow for acceptance of larger substrates. To identify the binding pocket residues in AtdA3, a homology model of the protein was built using InsightII. The crystal structures of the α-subunits of naphthalene, biphenyl and cumene dioxygenases were used as templates. All three reference dioxyenases have at least 25% identity with AtdA3. The model and a close up of the substrate binding pocket is shown in Figure 17.

The substrate, 2-ethylaniline, was docked in the binding pocket of the model and residues, within van der Waals contact distance of 4 Å from the substrate, were identified as binding pocket residues. Fourteen residues were identified:

N	198	E	257
G	202	W	260
V	205	A	293
L	213	G	294
I	248	N	296
Q	250	L	304
K	256	F	348

Saturation mutagenesis of the binding pocket residues was used to widen the substrate specificity of AtdA3, but no derivatives capable of metabolizing 2-isopropylaniline or 2-ABP were

obtained. The fact that aniline dioxygenase can oxidize and deaminate aniline makes it likely that derivatives of aniline dioxygenase could eventually be obtained that can deaminate 2-ABP. However, this goal was not achieved within the time and resource constraints of this project.

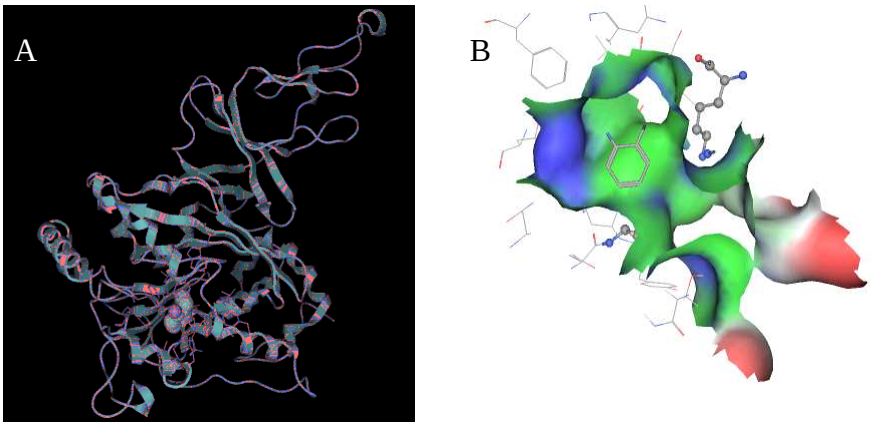


FIGURE 17. (A) HOMOLOGY MODEL OF ATDA3, WITH 2-EA DOCKED TO THE ACTIVE SITE. RESIDUES IN RED FORM THE BINDING POCKET; (B) CLOSE UP OF ATDA3 BINDING POCKET AND THE SUBSTRATE CHANNEL.

Genetic Studies

The reason that the isolation of a culture capable of utilizing 2-aminobiphenyl as a sole nitrogen source is a focus of enrichment culture experiments is because we ultimately want to construct a new metabolic pathway for the selective cleavage of both C-N bonds in carbazole. The first step in the carbazole degradation pathway accomplishes the cleavage of the first C-N bond in carbazole and yields a product 2-aminibiphenyl-2, 3-diol. Therefore an enzyme capable of selectively cleaving the C-N bond in 2-aminobiphenyl (since 2'-aminobiphenyl-2,3-diol is not commercially available) will be useful for the cleavage of the second C-N bond in carbazole. Genetic studies of the carbazole degradation pathway encoded by *Sphingomonas* sp. GTIN11 are essential to allow us to ultimately construct a novel pathway for the selective cleavage of both C-N bonds in carbazole and to facilitate the isolation of enzymes needed to complete this pathway.

Previous research by GTI scientists resulted in the isolation of the carbazole-degrading culture *Sphingomonas* sp. GTIN11 [31], and the cloning and sequencing a portion of the carbazole degrading genes of this culture. Prior to this project the *carA* and *carAc* genes of *Sphingomonas* sp. GTIN11 had been cloned and sequenced, but in order to have a fully functional carbazole 1,9a-dioxygenase enzyme (CarA) the *carAd* gene encoding the ferredoxin

reductase gene is also needed. In this project then we attempted to clone and sequence the *carAd* of *Sphingomonas* sp. GTIN11. DNA fragments contiguous with the *carA* and *carAc* genes were obtained by chromosome walking and eventually DNA sequence data was obtained for 5 kb in the vicinity of these *carA* and *carAc* genes. However, the *carAd* gene was not found in the vicinity of the *carA* and *carAc* genes in *Sphingomonas* sp. GTIN11. It is likely that the ferredoxin reductase of some other enzymatic pathway, encoded elsewhere on the chromosome, is capable of interacting with the products of the *carA* and *carAc* genes to produce a fully functional CarA enzyme. A map of the *Sphingomonas* sp. GTIN11 carbazole degradation operon is given below in Figure 18. While this work was in progress the DNA sequence of the *car* operon of a new carbazole-degrading culture, *Sphingomonas* sp. KA1 was published (Accession number AB095953.1) and it turns out that the DNA sequences of the *car* genes in *Sphingomonas* sp. KA1 and *Sphingomonas* sp. GTIN11 are identical.

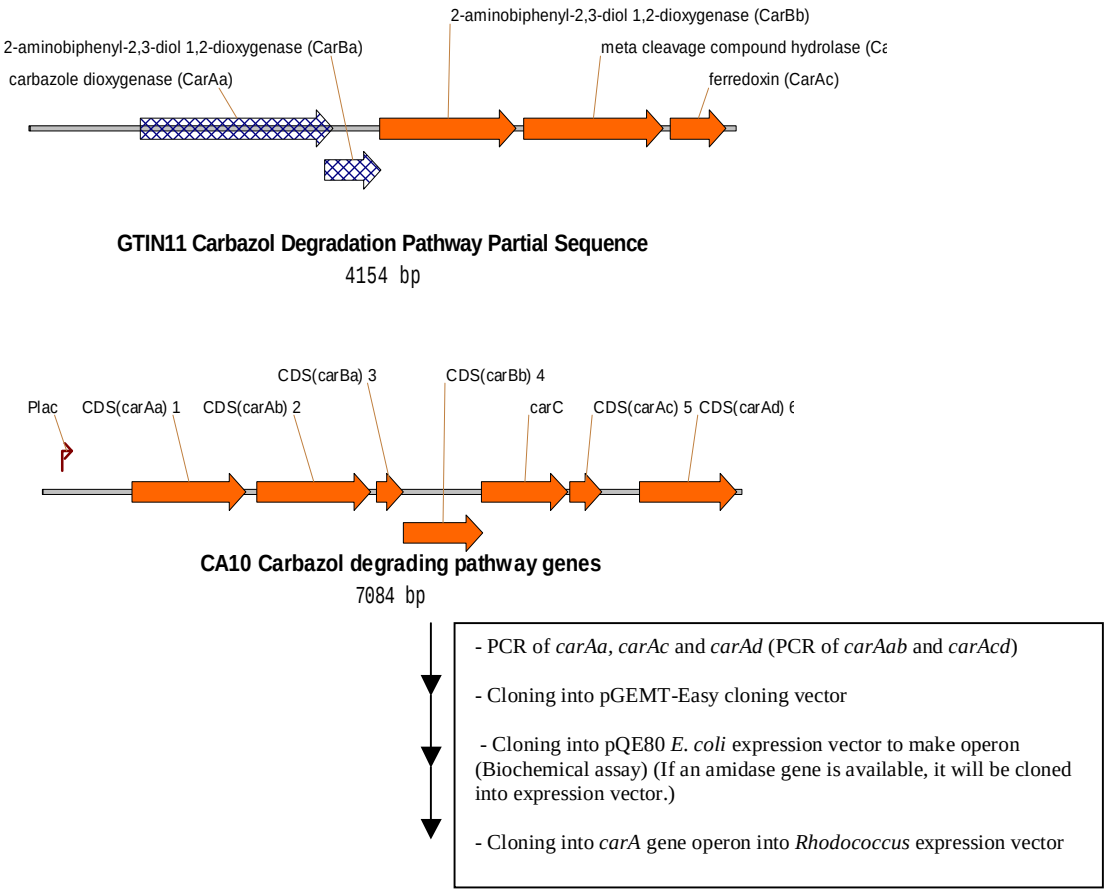
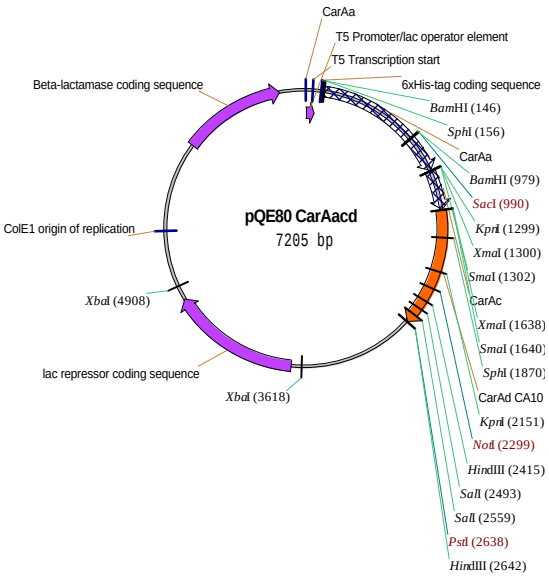
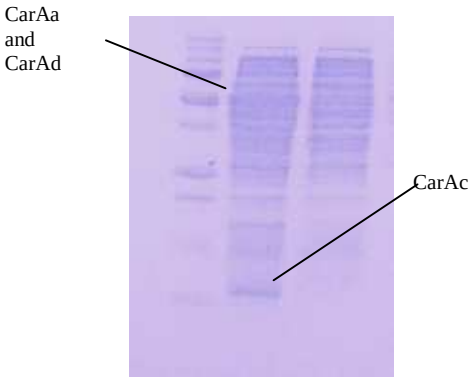


FIGURE 18. MAPS OF THE CARBAZOLE DEGRADATION OPERONS OF SPHINGOMONAS SP. GTIN11, AND PSEUDOMONAS RESINOVORANS CA10

The maps of the carbazole degradation operons of *Sphingomonas* sp. GTIN11, and *Pseudomonas resinovorans* CA10 are given in Figure 18. It is obvious that the gene arrangement of the carbazole operons in the *Sphingomonas* versus the *Pseudomonas* cultures is not identical. Not only is the gene arrangement different, but also the DNA sequences vary considerably. The percentage similarity between the amino acid sequences of the CarAa and CarAc proteins from *Sphingomonas* sp. GTIN11, and *Pseudomonas resinovorans* CA10 are 60% and less that 40% respectively. Therefore, while the CarA enzymes from these two species catalyze the same reaction they are rather different proteins. Accordingly, in this project we have made derivative clones containing the *carA* genes from both *Sphingomonas* sp. GTIN11 and *Pseudomonas resinovorans* CA10 in expression vectors both for *E. coli* and

for *R. erythropolis*. Since the *carAd* gene of *Sphingomonas* sp. GTIN11 could not be identified the *Sphingomonas* sp. GTIN11 *carAa* and *carAc* genes were combined with the *carAd* gene from *Pseudomonas resinovorans* CA10 to constitute the full complement of *carA* genes. Similarly, derivative clones containing only the *carAa, carAb, carAc,* and *carAd* genes from *Pseudomonas resinovorans* CA10 were constructed. These CarA expression vectors are devoid of functional genes for subsequent enzymatic steps in the carbazole degradation pathway so that carbazole will be converted to 2'aminobiphenyl-2,3-diol but will not be degraded further. Figure 19 depicts the vectors pQE80-CarAacd and pQE80-CarAabcd-CA10, which are *E. coli* expression vectors that contain *carA* genes from *Sphingomonas* sp. GTIN11 and from *Pseudomonas resinovorans* CA10 respectively. A protein gel is also shown illustrating the expression of CarA proteins from these constructs. Protein expression experiments demonstrated expression of the CarA proteins in the *E. coli* pQE80 expression vector. Similar constructs were also made with other *E. coli* vectors such as pACYCDuet and an example of that is shown in Figure 20.





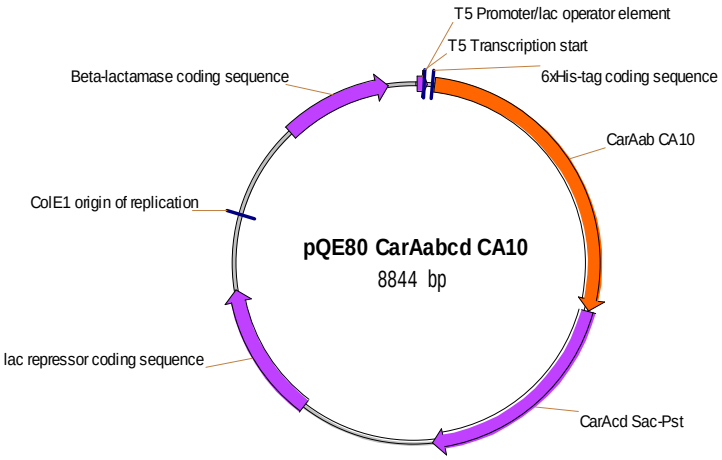


FIGURE 19. CONSTRUCTION OF *E. COLI* EXPRESSION VECTORS CONTAINING THE *CARA* GENES FROM *SPHINGOMONAS* SP. GTIN11 AND FROM *PSEUDOMONAS RESINOVORANS* CA10, AND A PROTEIN GEL ILLUSTRATING THE PRODUCTION OF THE CARA PROTEINS.

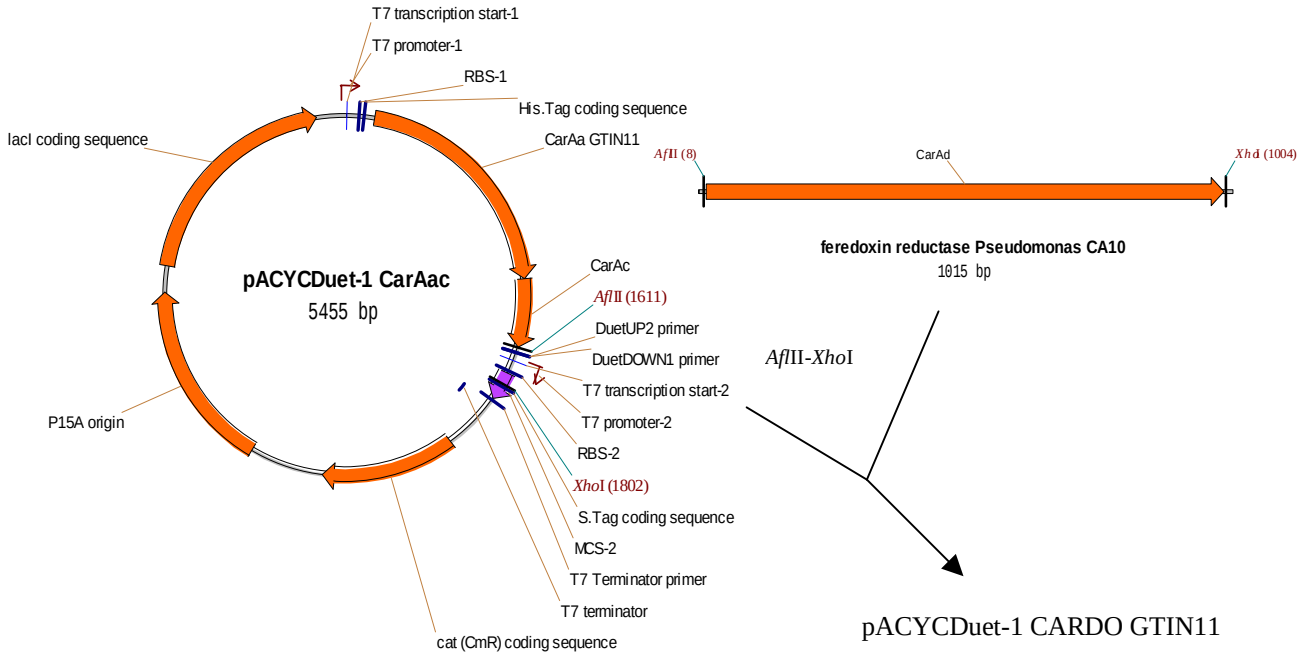


FIGURE 20. CONSTRUCTION OF pACYCDUET-1-CARDO-GTIN11 THAT CONTAINS THE *SPHINGOMONAS* SP. GTIN11 *CARA* GENES IN AN *E. COLI* EXPRESSION VECTOR.

Construction of a Derivative of *Sphingomonas* GTIN11 that Lacks CarB and CarC Enzymes

If a derivative of *Sphingomonas* sp. GTIN11 were available that had a functional CarA enzyme but lacked CarB and/or CarC enzymes it could be used to isolate the enzymes from suitable cultures that encode for a deaminase capable of cleaving the C-N bond in 2-aminobiphenyl. *Sphingomonas* sp. GTIN11 can utilize carbazole as both a source of nitrogen and carbon, but this is because nitrogen is ultimately liberated in the overall degradation of carbazole. What we want is to create a new metabolic pathway that is specific for the cleavage of C-N bonds in carbazole. If we retain the *carA* genes that encode for the cleavage of the first C-N bond in carbazole to form 2-aminobiphenyl-2,3-diol, but delete one or more of the other genes in the carbazole degradation pathway, then the culture will convert carbazole to 2-aminobiphenyl-2,3-

diol, but will not degrade it further. Accordingly, we performed experiments to delete the *carB* and *carC* genes. This was accomplished by the steps of: obtaining a DNA fragment containing the *carB* and *carC* genes as well as flanking regions, replacing the *carB* and *carC* genes with the kanamycin resistance gene from pSRKgfp, introduce this non-replicating DNA fragment into *Sphingomonas* sp. GTIN11 and select for kanamycin resistance to isolate derivatives of GTIN11 (resulting from homologous recombination at flanking DNA regions) that contain a deletion of *carB* and *carC*. A derivative of GTIN11 that lacks *carB* and *carC* could then be used as a host to clone deaminase genes that may be capable of utilizing 2-aminobiphenyl-2,3-diol as a substrate.

The strategy to construct a derivative of *Sphingomonas* GTIN11 that possesses a functional CarA enzyme (encoded by *carAa*, *carAc*, and *carAd*), but lacks subsequent steps in the carbazole degradation pathway is pictured in Figure 21. After obtaining kanamycin-resistant derivatives as pictured in Figure 21 we were unable to demonstrate the presence of a functional CarA enzyme. This may be due to a disruption of the expression of the *carAc* gene resulting from the deletion of the *carB* and *carC* genes. Accordingly, the use of cloned *carA* genes in combination with a suitable deaminase gene appears to provide the best means of constructing a metabolic pathway for the selective cleavage of both C-N bonds in carbazole.

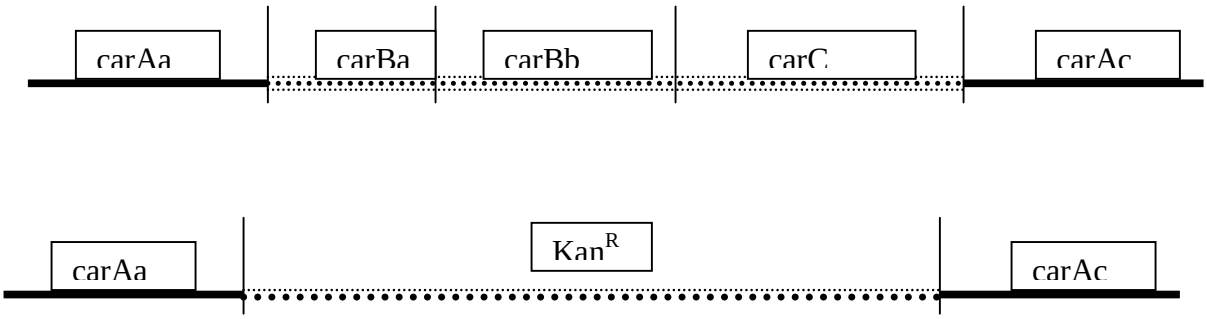


FIGURE 21. STRATEGY FOR THE REPLACEMENT OF THE *CARBA*, *CARBB*, AND *CARC* GENES OF SPHINGOMONAS GTIN11 WITH A GENE ENCODING KANAMYCIN RESISTANCE.

Genetic Transformation of *Rhodococcus*, *Sphingomonas*, and *Pseudomonas*

It is important that we place genes relevant to C-N bond cleavage into *Rhodococcus erythropolis* hosts to check the expression/functionality of those genes, and also to obtain bacterial cultures that can cleave C-S as well as C-N bonds. Similarly, the most useful cultures that we have found thus far that encode enzymes relevant to the cleavage of C-N bonds are the quinoline-metabolizing *Pseudomonas ayucida* IGTN9, and the carbazole-metabolizing *Sphingomonas* sp. GTIN11, so it is also important that we develop the ability to perform genetic manipulations of these cultures. We have several *R. erythropolis* strains: IGTS8, SQ1, CPE648 and CW25. Each of these has been used in electroporation experiments to introduce plasmids containing genes relevant to C-N bond cleavage. Genetic studies of *Pseudomonas ayucida* IGTN9 or *Sphingomonas* sp. GTIN11 have never been performed in the past so no protocols are available, nor is it known what antibiotic resistant markers or replicons can be used in these hosts. Accordingly, we performed electroporation experiments with several plasmid vectors in our possession to identify plasmids and genetic markers that may be used in these hosts. The kanamycin resistance gene encode by pKT230 appeared to be successfully transformed into both *Pseudomonas ayucida* IGTN9 and *Sphingomonas* sp. GTIN11.

However, gene expression in *Rhodococcus* hosts was of greatest interest in this project, as *Rhodococcus erythropolis* is known to tolerate exposure to crude oil and distilled petroleum products. Moreover, *Rhodococcus erythropolis* IGTS8 can selectively cleave carbon-sulfur bonds in organosulfur compounds typically found in petroleum. Therefore, it may be possible to construct a *R. erythropolis* strain that can simultaneously metabolize organosulfur and organonitrogen compounds.

Expression of CarA and AtdA in *Rhodococcus*

The *carAacd* genes from *Sphingomonas* sp. GTIN11 and *Pseudomonas resinovorans* CA10 have been cloned with their native promoters and with alternative promoters in several different vectors capable of replicating in *Rhodococcus* and *E. coli*. The vectors used include pEBC26, pSRKgfp, pRESQ, and pRESX. The plasmid pEBC26 [32] can replicate in both *E. coli* and *Rhodococcus*. It possesses antibiotic resistance genes for ampicillin and tetracycline for use in *E. coli* and *Rhodococcus* respectively. pEBC26 also contains a promoterless β-galactosidase gene downstream from a multiple cloning site so this plasmid can be used as a promoter probe vector to isolate DNA fragments capable of functioning in *Rhodococcus* as promoters. pEBC26

can also be used as a gene expression vector in *Rhodococcus* and a map of this plasmid is shown in Figure 22. A derivative of pEBC26 that contains a strong *Rhodococcus* promoter (K2) driving the expression of β-galactosidase was also used. This plasmid was called pYgalK2.

PSRKgfp [33] replicates in both *E. coli* and *Rhodococcus* and contains genes for resistance to ampicillin and kanamycin for use in *E. coli* and *Rhodococcus* respectively. The P45 promoter from a *Corynebacterium* is used to express the kanamycin resistance gene in pSRKgfp. This kanamycin resistance gene, with and without the P45 promoter, was also useful in constructing other *Rhodococcus* gene expression vectors. A map of pSRKgfp is shown in Figure 23.

The plasmids pRESQ and pRESX can replicate in *E. coli* and *Rhodococcus* and contain the *ccdB* gene and kanamycin resistance [34]. Expression of the *ccdB* gene in *E. coli* is lethal so it can be used as a positive selection cloning vector. Unique restriction sites within the *ccdB* gene are used to clone DNA fragments of interest and only DNA molecules or pRESQ or pRESX that successfully received inserts resulting in the inactivation of the *ccdB* gene will yield transformants in *E. coli*. The difference between pRESQ and pRESX is the promoter upstream of the *ccdB* gene. In pRESQ the *E. coli lacZ* promoter is present while in pRESQ the promoter of the *Rhodococcus kstD* gene is present. Both promoters function in *Rhodococcus* and genes cloned into pRESQ and pRESX can potentially be expressed in *Rhodococcus* from these promoters. A map of pRESQ is presented in Figure 24.

One of the problems faced in attempts to express genes in *Rhodococcus* hosts is that some of the genes we are interested in do not have selectable or easily assayed phenotypes. Moreover, genes of interest may come from hosts other than *Rhodococcus* and may not have promoters that are known to function in these hosts. Accordingly, we have constructed a gene expression vector in which the kanamycin resistance gene from pSRKgfp is used as a selectable genetic marker, but the promoter of the Kan^R gene has been separated from the structural gene encoding for kanamycin resistance by the insertion of a cloning site, Nsi1. This allows the cloning of genes of interest into the Nsi1 site so that the gene is placed immediately downstream from the promoter for the Kan^R gene, and if the Kan^R gene is functionally expressed then we can be sure that our gene of interest is being transcribed. In other words, we will form a transcriptional fusion with the Kan^R gene and any gene of interest and since the same promoter

will transcribe both the gene of interest and Kan^{R} then the selection of the culture for kanamycin resistance ensures that the gene of interest is also being expressed.

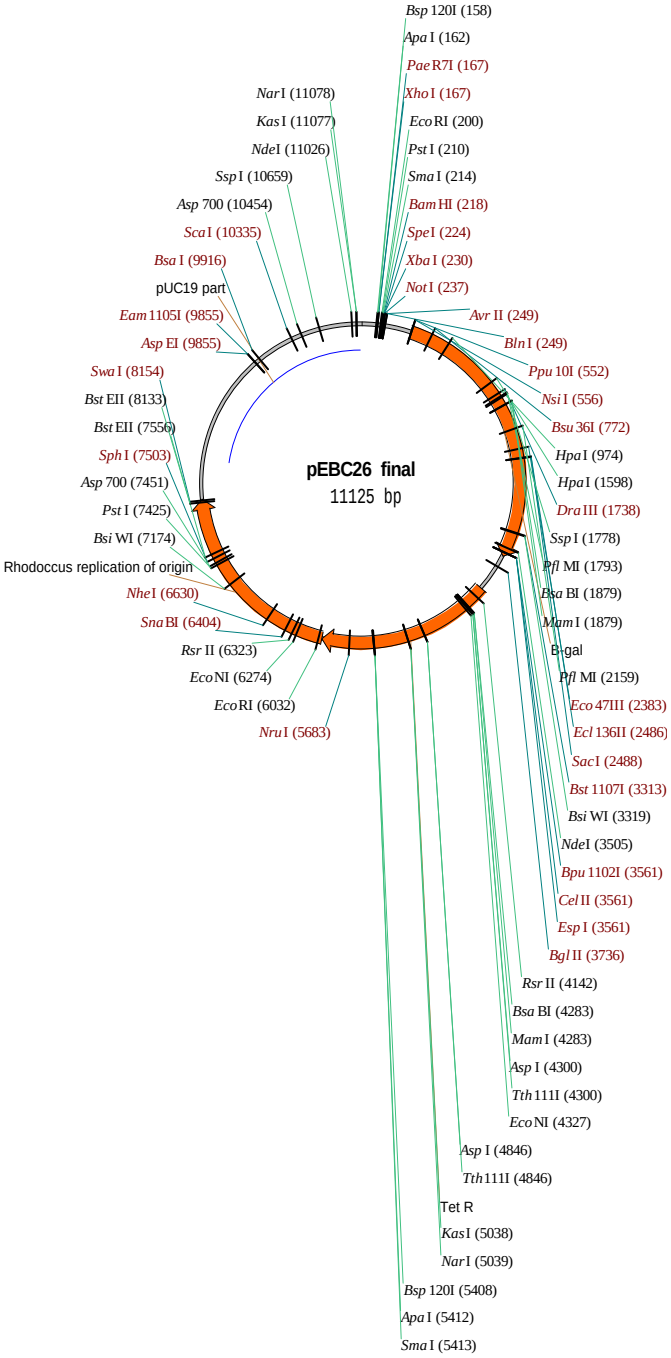


FIGURE 22. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR pEBC26.

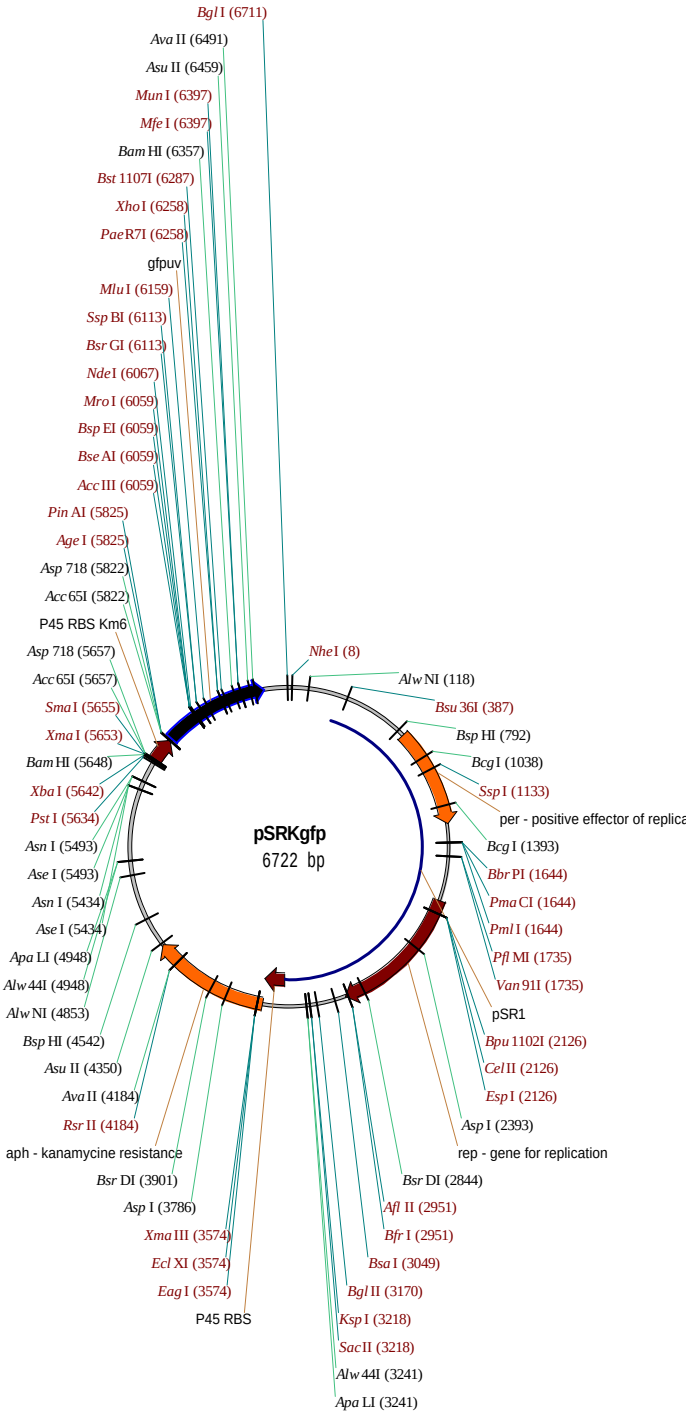


FIGURE 23. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR pSRKgfp.

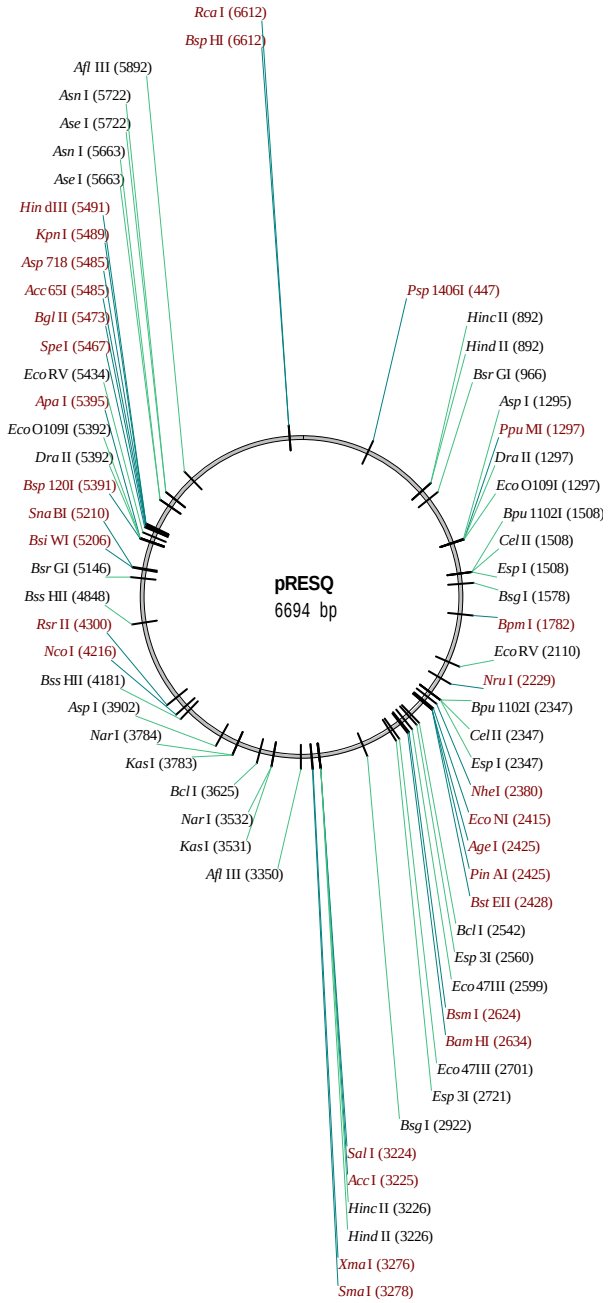


FIGURE 24. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR pRESQ.

This *Rhodococcus* gene expression vector is named pREP45boxBNsiKan, and an illustration of this plasmid is shown in Figure 25. This vector, pREP45boxBNsiKan, was constructed by replacing the beta-galactosidase gene (Xba/Bgl2 fragment) of pEBC26 with the kanamycin resistance gene of pSRKgfp. However the kanamycin resistance gene was modified using PCR to place two additional DNA sequences between the promoter (P45) and the structural gene encoding kanamycin resistance. These two DNA sequences were a boxB imperfect inverted repeat sequence and a Nsi1 site. The boxB sequence is derived from the *rrn* operon of *Rhodococcus erythropolis* IGTS8 and is thought to stabilize mRNA, most probably by providing protection from nucleases. It is hoped that the modified mRNA containing the boxB sequence will have a longer half-life in cells and will therefore result in higher levels of expression as compared to the original Kan^R gene. The Nsi1 site provides a cloning site for genes of interest allowing them to be inserted so that they are transcribed from the same promoter (P45) as the Kan^R gene. The P45 promoter can be removed from this plasmid by cleavage with Xba1 and Nsi1. This allows DNA fragments to be cloned as Xba1/Nsi1 fragments to serve as promoters to express kanamycin resistance in *Rhodococcus*. Additionally, a promoter along with a gene(s) of interest can be cloned as a Xba1/Nsi1 fragment to create an operon with kanamycin resistance as the distal gene in the operon so that expression of kanamycin resistance serves to indicate the successful transcription of genes located between the promoter and the kanamycin resistance gene.

The following constructs for the expression of genes in *Rhodococcus* relevant to the cleavage of C-N bonds were made: IGTS8/pYgalK2-TriA, SQ1/pEBC26-Cardo, IGTS8/pEBC26-Cardo, SQ1/pYgalK2-QorMSL, CPE648/pYgalK2-TriA, CW25/pygalK2-TriA, CPE684/pEBC26-Cardo, CW25/pEBC26-Cardo, and SQ1/pYgalK2-TriA, CW25/pRESQ-Cardo, CW25/pRESX-Cardo, pSRKgfp-Cardo, pREKan-Cardo, and pUC8B-AtdA.

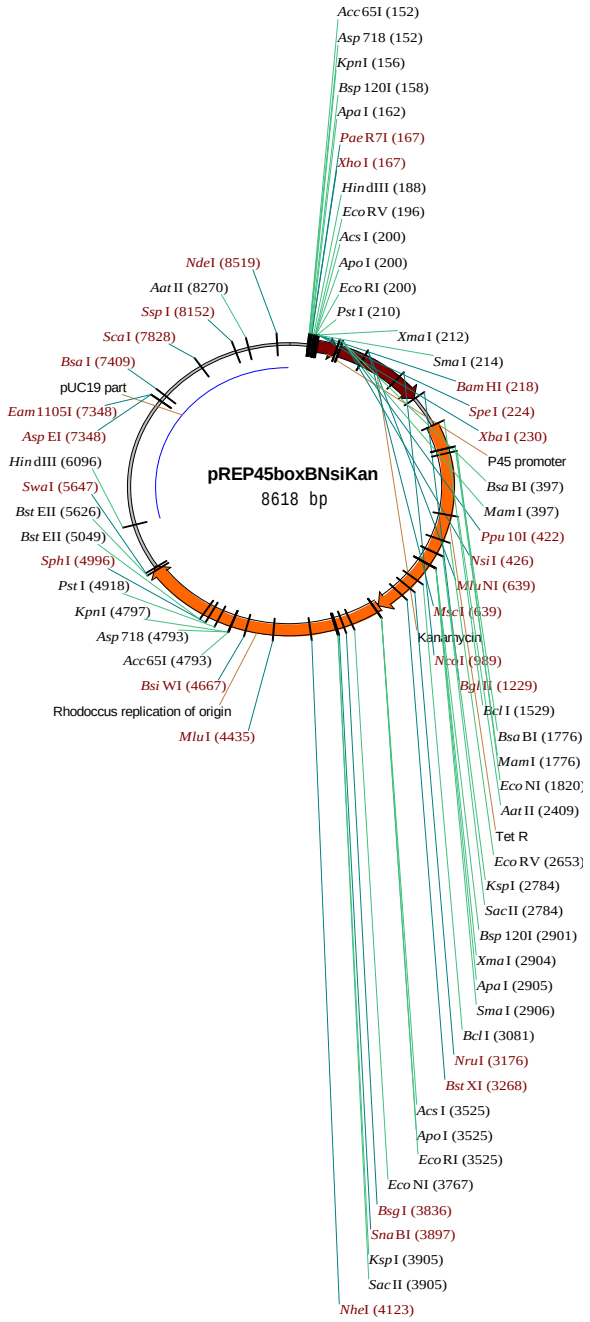


FIGURE 25. MAP OF *RHODOCOCCUS* EXPRESSION VECTOR pREP45BOXBNSIKAN.

Investigation of *Rhodococcus* Promoters

E. coli promoters typically possess conserved hexanucleotide sequences located approximately 10 basepairs (TATAAT) and 35 basepairs (TTGACA) upstream from the transcriptional initiation site [35]. These regions are generally spaced 17 bp apart. These conserved hexanucleotide sequences serve as promoter recognition sites for RNA polymerase, and the sigma subunit of RNA polymerase makes contact with the DNA in these regions. The promoters of a variety of other bacterial species have also been observed to have promoter regions that possess hexanucleotide sequences with high degrees of similarity to the *E. coli* sequences. However, *E. coli* promoters do not always function in other bacterial species and there is no DNA sequence that is universally recognized in all bacterial species as a promoter. Some bacterial species have been found to possess promoters that have conserved hexanucleotide sequences that are different from the *E. coli* consensus sequences, and the promoter regions of some bacterial species do not appear to have conserved regions [36, 37].

R. erythropolis promoters have not been examined thoroughly, so it is not known if *Rhodococcus* promoters share similar conserved DNA sequence regions. The spacing between transcriptional initiation site and the promoter recognition sites is thought to be fairly uniform in most bacterial species so if a collection of promoters obtained from a given bacterial species is aligned by their transcriptional initiation sites it should be possible to inspect the sequences to determine if conserved sequences occur in regions corresponding to the –10 and –35 bp hexanucleotide regions that have been observed in other bacterial species. One concern about this approach to identify conserved promoter sequences is that each bacterial species can have several forms of RNA polymerase that each recognizes unique DNA sequences. However, the method of selecting promoters in this study was designed to obtain constitutive promoters that will hopefully all belong to the same class, corresponding to the class of promoters recognized by the sigma 70 RNA polymerase in *E. coli* [38]. In order to obtain information regarding transcriptional signals in *Rhodococcus erythropolis* chromosomal DNA was cloned into the promoter probe vector pEBC26 and DNA fragments that were capable of functioning as promoters were identified. The transcriptional initiation sites of several of these *Rhodococcus erythropolis* promoters were mapped to facilitate comparison of their DNA sequences. The DNA sequences of several *R. erythropolis* promoters obtained in this study as well as *R. erythropolis*

promoters described in the literature, are shown in Figure 26 aligned by their transcriptional initiation sites.

An examination of the *R. erythropolis* promoters in Figure 26 show that, with the exception of the rrn promoter, do not possess regions that are homologous with the consensus sequences for *E. coli* or *B. subtilis* promoters, but several conserved motifs were found. Most of the promoters isolated in this study contained sequences in the –10 region that resemble the consensus sequence CGCGGA. There was no highly conserved –35 sequence detected in these *Rhodococcus* promoters but regions resembling the consensus sequence GGTGGC was found in most promoters at distances ranging from 12 to 26 basepairs from the presumed –10 region. A larger data set is needed before a consensus sequence for *R. erythropolis* promoters can be identified with certainty, but the presence of several conserved motifs that are present in at least some *Rhodococcus* promoters, including the DNA sequences of *R. erythropolis* genes obtained by other researchers, suggests that these sequences may contribute to promoter recognition and/or function.

	-50 -40 -30 -20 -10 +1
R16PRO-P1	TAACCGGCGCCAGTGCTCTGACCTGGCGATTT**GGTGTTG**TGTTGATCTTCGCCG**TAACTT**TGTCCAG
R16PRO-P2	CTGGCGAACACAATCCAGTTTTTCTTCTCGGGTTCC**GGTTTG**CGCCGGTGGCCAGG**TTGGGAT**AGGCTG
R16PRO-P3	CATCCGCGAATGATCAACCGCCTCTTGGGGTTT**GGTTGG**TTTGTGTGTGGTGTGAG**TATTTT**TTCGTCG
RCMPRO	CATGCCGTCTGGGATACGCGACGCCCCGAATC**GGTTGG**ACAATCACATCCGACATTGC**TAACTT**CGATGA**C**
pYgal **A1**	CCTGTGATGTCCATACGCCGTGTTCTCTCGCTTCCC**GGTGC**ACTGCGCACGGCCCG**ATTCGGC**GTAGCGA
pYgal **A2**	CTTTCCGAC**GTGGGCG**TAAGGCGATCGGTGG**ACACACC**GTTCCGC
pYgal **A3**	CGGCGTAGCGACTTTCCGAC**GTGGGCG**TAAGGCGATC**GGTGGAC**ACACCGTTCCGCC**TGCACA**GGGTC
pYgal **E**	GAGTAGTCCAGACCACGGCGGCTATC**GTGGC**ACGGCGTCATCTTCGGCCGGTG**TGACTG**CCGAGT
pYgal **I**	CGATCGACGGAGATGGCGTCCGCA**GTGGCG**ACTGCCGGCCGGAACCAGTTC**ACCGCGAA**CAGCTT
pYgal **K**	GCGTCCGCA**GTGGCG**ACGTCGGCCGGAACCAGTTC**ACCGCGAA**CAGCTT
pYgal **K2**	TGACGACGACGGAGATGGCGTCCGCA**GTGGCG**ACTGGCCGGAACCAGTTC**ACCGCGAA**CAGCTT
pYgal **15**	AGCCCGGGGGATCGGGGAACA**GGTGGTG**CTCGATC**AGTCGAG**TAGTCC
pYgal **18**	TCGGCATCGGAGACAGCGAGCTCGGCCAGGATC**GTGGCTG**CAGTTCTCTTTGTCGGCGG**TTTCTA**C
pYgal **21**	GGGTGCCCACCTCGACCGTTGAA**GTGGC**AGCGACACAGTCTTCTTCGCG**ACGACGG**CGGAAC
pYgal **22**	AGCGAGCTCGGCCAGGATC**GTGGCTG**CAGTTCTCTTTGTCGCGG**TTTCTA**C
dsz	TTGAAATCAGGTGAAGTTTAAC**GGTGGGC**ACACCCCGGGGGTGGGGGTGAG**ACTGCTTA**G

FIG. 26. *RHODOCOCCUS* PROMOTER NUCLEOTIDE SEQUENCES. The underlined bases in bold text show the positions of the potential -35 and -10 regions. The transcription initiation site is underlined at the 3' end of each sequence.

FIGURE 27. NUCLEOTIDE SEQUENCES OF THE *E. COLI LACZ* PROMOTER AND DERIVATIVE PROMOTERS.

Promoter 1-1: 170 bp, *lacZ* wild type
tcatctagatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggc*tttaca*ctttatgcttccggctcgtatg
ttgtgtggaAttgtgagcggataacaatttcacacaggaaacagct**atg**cattga

Promoter 1-2: 140 bp, WT *E. coli* –35 region, WT *E. coli* –10 region, shorter untranslated leader region, altered RBS
tcatctagatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggc*tttaca*ctttatgcttccggctcgtatg
ttgtgtggaAaggaa**atg**cattga

Promoter 1-3: 140 bp, WT *E. coli* –35 region, *Rhodococcus* –10 region, shorter untranslated leader region, altered RBS
tcatctagatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggc*tttaca*ctttatgcttccggctcgcgc
ggagtgtggaAaggaa**atg**cattga

Promoter 2-1: 100 bp, shorter *lacZ* wild type
tcatctagaggcaccccaggc*tttaca*ctttatgcttccggctcgtatgttgtgtggaAttgtgagcggataacaatttcacacaggaaacagct**atg**cattgat

Promoter 2-3: 65 bp, shorter upstream region, WT *E. coli* –35 region, *Rhodococcus* –10 region, shorter untranslated leader region, altered RBS
tcatctagaggcaccccaggc*tttaca*ctttatgcttccggctcgcgcggagtgtggaAaggaa**atg**cattga

Promoter 3-1: 80 bp, *Rhodococcus* –35 region, WT *E. coli* –10 region, untranslated leader region, and RBS
tcatctaga*ggtggc*ctttatgcttccggctcgtatgttgtgtggaAttgtgagcggataacaatttcacacaggaaacagct**atg**cattga

Promoter P45: *Corynebacterium* promoter expressing Km^R in pSRKgfp
cgcttggcttgaccgacttgatctgatcctcattcgccgtggcagcgtgggtcgaatgagaatacgaatggattggtcagggatttttcccgaagggcactaattttgctaaagtaagtgacgaagaa
gttcagcgggcacaggatcaagagacaggatgaggatcgtttcgc

The 5’ Xba1 restriction site (tctaga) is underline in each sequence. The 3’ Nsi 1 restriction site (**atg**cat), which contains the atg translational start site indicated in bold letters, is underlined in each sequence. The –35 hexanucleotide sequence is in italics and underlined. The –10 hexanucleotide sequence is underlined. The transcription initiation site is indicated with a capital letter “A”.

In order to determine if these proposed consensus sequences actually contribute to promoter function in *R. erythropolis* several synthetic DNA fragments were constructed based of the *E. coli lacZ* promoter. The *lacZ* promoter of *E. coli* has been reported to function in *R. erythropolis* [34] so it was used as a model promoter, which was subsequently modified to contain –10 or –35 regions of proposed *Rhodococcus* hexanucleotide sequences. Six DNA fragments were obtained by PCR amplification of the *lacZ* promoter from the *E. coli* vector pUC19 and these fragments were designated as 1-1, 1-2, 1-3, 2-1, 2-3, and 3-1. The DNA sequences of these six fragments, as well as the P45 promoter, are shown in Figure 27. The PCR primers use to obtain these six DNA fragments incorporated 5' Xba1 and 3' Nsi1 sites to allow directional cloning of the fragments into the *Rhodococcus* vector pREP45boxBNsiKan. Cleavage of pREP45boxBNsiKan by Xba1 and Nsi1 results in the excision of the P45 promoter that is immediately upstream from the kanamycin gene in this plasmid. In this way each of the six DNA fragments listed in Figure 27 was cloned upstream of the kanamycin gene in a vector that replicates in both *E. coli* and *R. erythropolis*. Investigations of promoter strength were performed in *E. coli* and *R. erythropolis* by quantifying resistance to kanamycin and by quantifying the amount of kanamycin gene mRNA using reverse transcription quantitative PCR (RT-qPCR). The results of these analyses are summarized in Tables 6 and 7 that show data obtained in *E. coli* and *R. erythropolis* respectively.

Promoter activity in *E. coli* was maintained despite modifications to the –10 region, but was eliminated by modifications to the –35 region. Promoter activity in *R. erythropolis* was only achieved in derivatives with a modified –10 region. The lack of activity of the wild type *lacZ* promoter in derivative pREKan1-1 in *R. erythropolis* is puzzling. Plasmid pREKan1-1 showed strong promoter activity in *E. coli* and the literature reports that this promoter (*E. coli lacZ*) functions in *R. erythropolis* [34]. However, when the DNA fragment containing the wild type *lacZ* promoter was placed upstream of the kanamycin gene in pREKan1-1 we detected no promoter activity in *R. erythropolis*. The vector used to test these promoter variants was pREP45boxBNsiKan (pREP45Kan) which contains the P45 promoter plus the kanamycin resistance gene from vector pSRKgfp cloned into pEBC26 resulting in the replacement of the B-galactosidase gene of pEBC26 being replaced with the kanamycin resistance gene. Surprisingly, the pREP45Kan vector expresses kanamycin resistance well in *E. coli* but not in *R. erythropolis*. Therefore, the local environment of the pEBC26 vector appears to affect the ability of DNA

fragments to function as promoters in *R. erythropolis* as both the P45 promoter and the wild type *lacZ* promoter failed to function in *R. erythropolis* when these promoters were cloned as Xba1/Nsi1 fragments into the pEBC26 vector. In any event, the observation that the hexanucleotide sequence CGCGGA appears to be conserved in the –10 region of DNA fragments that function as promoters in *R. erythropolis* and the substitution of this CGCGGA hexanucleotide for the normal –10 hexanucleotide sequence (TATAAT) in the *lacZ* promoter resulted in good promoter activity in *R. erythropolis* suggests that this sequence may contribute to promoter activity in *R. erythropolis* and warrants further investigation.

These data increase our understanding of the unique DNA sequence requirements for promoter activity in *Rhodococcus* and will benefit future projects that seek to obtain gene expression in *R. erythropolis* and related species.

Table 6. Promoter Strength in *Escherichia coli* DH5α: Growth on Agar Containing Different Kanamycin Concentrations and RT-qPCR Data

Promoter versions	Kan 20	Kan35	Kan 50 μg/ml	RT-qPCR
DH5/pREkan-1-1	+++	+++	+++	$6.2x10^5$
DH5/pREkan-1-2	+++	++	+	$2.9x10^5$
DH5/pREkan-1-3	+++	++	++	$6.0x10^4$
DH5/pREkan-2-1	+	+	+/-	$3.5x10^4$
DH5/pREkan-2-3	+	+/-	-	0
DH5/pREkan-3-1	-	-	-	0
DH5/pREP45kan	+	-	-	$1.2x10^5$
DH5/pSRKGFP	+++	+++	+++	$2.9x10^5$
DH5 only	-	-	-	0

-/+ poor: growth in the thick part of the streak only
+/- fair growth: no individual colonies
++ good growth including individual colonies
+++ robust growth
- no growth

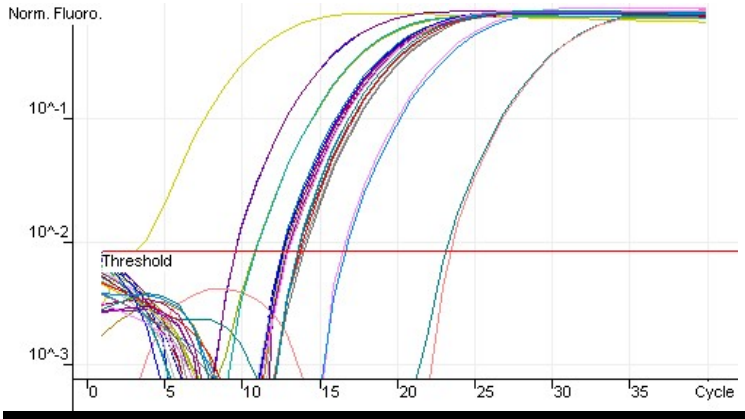
Table 7. Promoter Strength in *Rhodococcus erythropolis* CW25: Growth on Agar Containing Different Kanamycin Concentrations and RT-qPCR Data

Promoter versions	Kan 75	Kan100	Kan150	Kan 250 µg/ml	RT-qPCR
CW25/pREkan-1-1	++	++	+	_	0
CW25/pREkan-1-2	+	+	-/+	_	0
CW25/pREkan-1-3	+++	+++	+++	+++	$3.4x10^3$
CW25/pREkan-2-1	++	+	_	_	0
CW25/pREkan-2-3	+	+	_	_	0
CW25/pREkan-3-1	+	+	+/-	_	0
CW25/pREP45kan	+/-	-/+	-	-	0
CW25/pSRKGFP	+++	+++	+++	+++	$2.2x10^5$
CW25only	+/-	-/+	-	-	0

-/+ poor: growth in the thick part of the streak only
+/- fair growth: no individual colonies
++ good growth including individual colonies
+++ robust growth
- no growth

RT-qPCR data providing the details of data obtained for the quantification of messenger RNA for the kanamycin gene expressed *E. coli* and *R. erythropolis* are presented in Figures 28-30. An external control of luciferase mRNA was added to all experimental RT-qPCR samples. Figure 28 provides representative data for one such control amplification. A standard curve was prepared using a dilution series of luciferase mRNA. The amount of luciferase mRNA added to each experimental sample was the same so it was expected that the amount of *luc* mRNA detected in each sample should be the same unless there were variations introduced by the mRNA purification and quantification techniques. Indeed, as shown in Figure 28 all the experimental samples show highly similar results.

FIGURE 28. RT-QPCR DATA FOR LUCIFERASE EXTERNAL mRNA CONTROL IN E. COLI

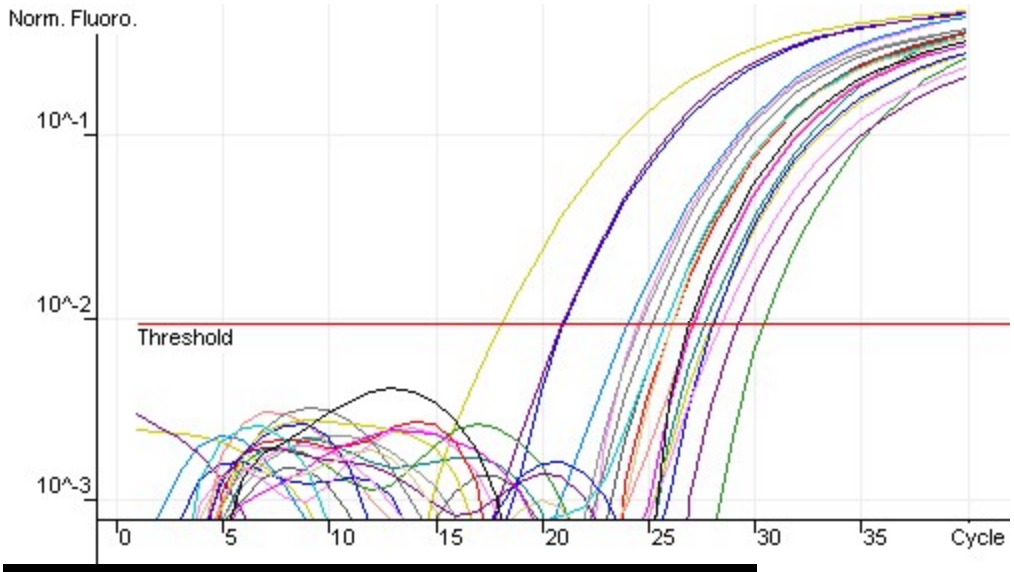


No.	Colour	Name	Type	Ct	Calc Conc (copies/reaction)	% Var
A2		luc, 1x10^8 (luc)	Standard	3.27	9.33E+07	6.7%
A3		luc, 1x10^7 (luc)	Standard	9.63	1.07E+07	6.9%
A4		luc, 1x10^7 (luc)	Standard	9.65	1.06E+07	6.1%
A5		luc, 1x10^6 (luc)	Standard	16.55	1.01E+06	0.8%
A6		luc, 1x10^6 (luc)	Standard	16.73	9.49E+05	5.1%
A7		luc, 1x10^5 (luc)	Standard	23.19	1.05E+05	5.0%
A8		luc, 1x10^5 (luc)	Standard	23.51	9.40E+04	6.0%
B1		luc NTC	Positive Control	0.00	1.00E+00	0.0%
C4		DH5/1-1(Luc)	Unknown	12.91	3.49E+06	0.0%
C5		DH5/1-1(Luc)	Unknown	13.09	3.28E+06	0.0%
C6		DH5/1-2(Luc)	Unknown	10.90	6.92E+06	0.0%
C7		DH5/1-2(Luc)	Unknown	10.96	6.79E+06	0.0%
C8		DH5/1-3(Luc)	Unknown	12.62	3.85E+06	0.0%
D1		DH5/1-3(Luc)	Unknown	12.69	3.76E+06	0.0%
D2		DH5/2-1(Luc)	Unknown	12.91	3.49E+06	0.0%
D3		DH5/2-1(Luc)	Unknown	12.87	3.54E+06	0.0%
D4		DH5/2-3(Luc)	Unknown	12.87	3.54E+06	0.0%
D5		DH5/2-3(Luc)	Unknown	12.95	3.44E+06	0.0%
D6		DH5/3-1(Luc)	Unknown	13.91	2.48E+06	0.0%
D7		DH5/3-1(Luc)	Unknown	14.06	2.36E+06	0.0%
D8		DH5 ONLY(Luc)	Unknown	13.68	2.68E+06	0.0%
E1		DH5 ONLY(Luc)	Unknown	13.80	2.57E+06	0.0%

E2		DH5/pREP45(Luc)	Unknown	12.79	3.63E+06	0.0%
E3		DH5/pREP45(Luc)	Unknown	12.74	3.70E+06	0.0%
E4		DH5/pSRKgfp(Luc)	Unknown	13.01	3.37E+06	0.0%
E5		DH5/pSRKgfp(Luc)	Unknown	13.09	3.28E+06	0.0%
E6		DH5/1-2 (2) (Luc)	Unknown	13.62	2.74E+06	0.0%
E7		DH5/1-2 (2) (Luc)	Unknown	13.66	2.71E+06	0.0%

The data in Figure 29 were obtained from the RT-qPCR analysis of mRNA samples extracted from *E. coli* cells containing promoter variants pREKan1-1, pREKan1-2, pREKan1-3, pREKan2-1, pREKan2-3, pREKan3-1, and control plasmids pREP45Kan, and pSRKgfp. DNA prepared from pREP45Kan was used to generate a standard curve for the kanamycin gene and the level of kan mRNA in each experimental sample was quantified. The data in Figure 29 illustrate that a good standard curve was obtained for the quantification of kanamycin, and this result is representative of multiple similar experiments that were performed. The levels of kan mRNA in *E. coli* and in *Rhodococcus* cultures was determined on at least four occasions so that accurate and reproducible values could be obtained.

FIGURE 29. RT-QPCR DATA FOR EXPRESSION OF KANAMYCIN IN E. COLI BY PROMOTER VARIATIONS

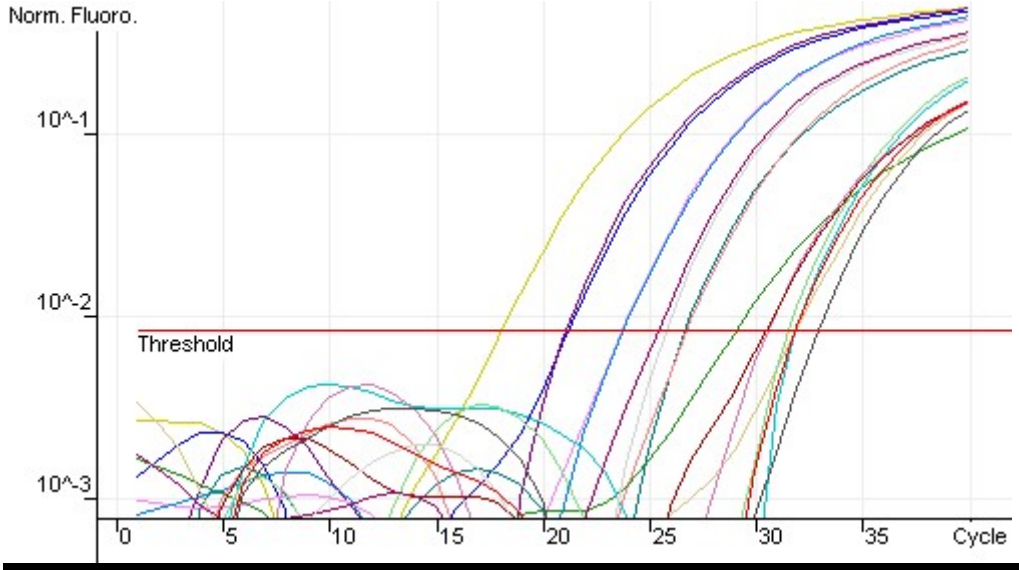


No.	Colour	Name	Type	Ct	Calc Conc (copies/reaction)	% Var

No.	Colour	Name	Type	Ct	Calc Conc (copies/reaction)	% Var
A2		pREP45 1x10^8 (kan)	Standard	18.20	9.27E+07	7.3%
A3		pREP45 1x10^7 (kan)	Standard	21.12	1.06E+07	6.1%
A4		pREP45 1x10^7 (kan)	Standard	21.06	1.11E+07	10.6%
A5		pREP45 1x10^6 (kan)	Standard	24.58	8.15E+05	18.5%
A6		pREP45 1x10^6 (kan)	Standard	24.13	1.14E+06	13.9%
A7		pREP45 1x10^5 (kan)	Standard	27.72	7.96E+04	20.4%
A8		pREP45 1x10^5 (kan)	Standard	27.14	1.22E+05	22.1%
B1		pREP45 1x10^4 (kan)	Standard	30.49	1.02E+04	1.9%
B2		pREP45 1x10^4 (kan)	Standard	0.00	1.00E+00	0.0%
B3		pREP45 1x10^3 (kan)	Standard	0.00	1.00E+00	0.0%
D6		DH/1-1 (KAN)	Unknown	24.72	7.34E+05	0.0%
D7		DH/1-1 (KAN)	Unknown	25.21	5.10E+05	0.0%
D8		DH/1-2 (KAN)	Unknown	25.82	3.24E+05	0.0%
E1		DH/1-2 (KAN)	Unknown	26.17	2.51E+05	0.0%
E2		DH/1-3 (KAN)	Unknown	28.06	6.17E+04	0.0%
E3		DH/1-3 (KAN)	Unknown	28.12	5.90E+04	0.0%
E4		DH/2-1 (KAN)	Unknown	29.28	2.49E+04	0.0%
E5		DH/2-1 (KAN)	Unknown	28.49	4.47E+04	0.0%
E6		DH/2-3 (KAN)	Unknown	0.00	1.00E+00	0.0%
E7		DH/2-3 (KAN)	Unknown	0.00	1.00E+00	0.0%
E8		DH/3-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
F1		DH/3-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
F2		DH/pREP45 (KAN)	Unknown	27.22	1.15E+05	0.0%
F3		DH/pREP45 (KAN)	Unknown	27.02	1.33E+05	0.0%
F4		DH/pSRKGFP (KAN)	Unknown	25.84	3.20E+05	0.0%
F5		DH/pSRKGFP (KAN)	Unknown	26.16	2.53E+05	0.0%
F6		DH ONLY (KAN)	Unknown	0.00	1.00E+00	0.0%
F7		DH ONLY (KAN)	Unknown	0.00	1.00E+00	0.0%
F8		DH/1-1 no rt (KAN)	Negative Control	0.00	1.00E+00	0.0%
G1		DH/1-1 no rt (KAN)	Negative Control	0.00	1.00E+00	0.0%
G2		NTC (KAN)	NTC	0.00	1.00E+00	0.0%

The data in Figure 30 was obtained from RT-qPCR experiments with mRNA purified from *Rhodococcus erythropolis* CW25 cultures containing various promoter configurations. A good standard curve for kanamycin was obtained and the expression of the kanamycin gene was only detected in mRNA from samples containing the 1-3 promoter or the control culture containing pSRKgfp.

FIGURE 30. RT-QPCR DATA FOR EXPRESSION OF KANAMYCIN IN R. ERYTHROPOLIS BY PROMOTER VARIATIONS



No.	Colour	Name	Type	Ct	Calc Conc (copies/reaction)	% Var
A2		pPREP45KAN 1x10^8	Standard	18.21	1.20E+08	20.1%
A3		pPREP45KAN 1x10^7	Standard	21.32	8.72E+06	12.8%
A4		pPREP45KAN 1x10^7	Standard	21.18	9.78E+06	2.2%
A5		pPREP45KAN 1x10^6	Standard	23.83	1.05E+06	4.5%
A6		pPREP45KAN 1x10^6	Standard	23.86	1.02E+06	1.8%
A7		pPREP45KAN 1x10^5	Standard	26.73	8.99E+04	10.1%
A8		pPREP45KAN 1x10^5	Standard	26.83	8.24E+04	17.6%
B1		pPREP45KAN 1x10^4	Standard	29.16	1.15E+04	14.7%
B4		pPREP45KAN 1x10^3	Standard	31.96	1.08E+03	8.0%
B5		NTC (kan std)	Negative Control	31.96	1.08E+03	0.0%
B6		NTC (kan std)	Negative Control	31.60	1.46E+03	0.0%
B7		CW/1-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
B8		CW/1-1 (KAN)	Unknown	0.00	1.00E+00	0.0%

No.	Colour	Name	Type	Ct	Calc Conc (copies/reaction)	% Var
C1		CW/1-2 (KAN)	Unknown	0.00	1.00E+00	0.0%
C2		CW/1-2 (KAN)	Unknown	0.00	1.00E+00	0.0%
C3		CW/1-3 (KAN)	Unknown	30.63	3.31E+03	0.0%
C4		CW/1-3 (KAN)	Unknown	30.58	3.46E+03	0.0%
C5		CW/2-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
C6		CW/2-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
C7		CW/2-3 (KAN)	Unknown	0.00	1.00E+00	0.0%
C8		CW/2-3 (KAN)	Unknown	0.00	1.00E+00	0.0%
D1		CW/3-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
D2		CW/3-1 (KAN)	Unknown	0.00	1.00E+00	0.0%
D3		CW/pSRKGFP (KAN)	Unknown	25.51	2.52E+05	0.0%
D4		CW/pSRKGFP (KAN)	Unknown	25.86	1.87E+05	0.0%
D5		CW25 ONLY (KAN)	Unknown	0.00	1.00E+00	0.0%
D6		CW25 ONLY (KAN)	Unknown	0.00	1.00E+00	0.0%
D7		NTC	Negative Control	0.00	1.00E+00	0.0%
D8		NO RT (1-3) (KAN)	Negative Control	33.01	4.44E+02	0.0%
E1		NO RT (1-3) (KAN)	Negative Control	31.94	1.10E+03	0.0%

CONCLUSIONS AND PROJECT IMPLICATIONS

Enrichment culture experiments designed to isolate a culture capable of cleaving the carbon-nitrogen bond in 2-aminobiphenyl, and thus being capable of providing an enzyme for the cleavage of the second carbon-nitrogen bond in carbazole, resulted in the isolation of a unique microbial culture *Pseudomonas* sp. GTIN-G4. *Pseudomonas* sp. GTIN-G4 is capable of metabolizing 2-aminobiphenyl and related compounds, but does not appear to be capable of cleaving the carbon-nitrogen bond. Instead it has the novel biochemical ability to modify organonitrogen compounds by replacing a hydrogen bound to the nitrogen atom with a formaldehyde group. Aminobiphenyl compounds are toxic and carcinogenic [26]. 4-aminobiphenyl has been shown to react with DNA and cause mutagenesis. 2-aminobiphenyl and 3-aminobiphenyl are less mutagenic and their primary mechanism of toxicity is thought to be reactions with proteins and interference with the electron transport system in cells [26, 27]. No culture capable of utilizing 2-ABP as a sole source of nitrogen and/or carbon was obtained from enrichment culture experiments, nor are there reports of the biodegradation of 2-ABP in the literature. This serves to highlight the toxicity of 2-ABP and its importance as an environmental contaminant. While the formylation of aniline is known this is the first report of the biodegradation of 2-ABP. The formation of N-formamide derivatives of 2-aminobiphenyl, diphenylamine, aniline, and o-toluidine is a unique capability of *Pseudomonas* sp. GTIN-G4 and this culture may be of use in the remediation of soil and water contaminated with aromatic amine compounds.

Directed evolution experiments were performed with several amidase, deaminase, and dioxygenease genes, however; no enzyme capable of selectively removing nitrogen from 2-aminobiphenyl was obtained. Nonetheless, the fact that aniline dioxygenase can deaminate aniline makes it probable that with further research directed evolution experiments could succeed in developing a derivative enzyme that could deaminate 2-aminobiphenyl.

Research performed in this project also included gene expression studies in *Rhodococcus erythropolis*. The genes that encode carbazole dioxygenase from *Sphingomonas* sp. GTIN11 and *Pseudomonas resinovorans* CA10 (*carAacd*), aniline dioxygenase from *Acinetobacter* sp. YAA (*atdA1, atdA2, atdA3, atdA4*, and *atdA5*), and resistance to the antibiotic kanamycin were cloned and introduced into *Rhodococcus erythropolis*, a bacterial host that tolerates exposure to petroleum. It was hoped that by simultaneously expressing the genes for carbazole dioxygenase

and aniline dioxygenase in the same bacterial host then enrichment culture experiments employing carbazole as a sole nitrogen source may result in the isolation of derivative cultures with altered abilities to cleave C-N bonds. However, no culture capable of selectively cleaving both C-N bonds in carbazole was obtained. Because no gene encoding an enzyme that could selectively cleave the C-N bond in 2-aminobiphenyl was available, it was not possible to construct a new metabolic pathway for the selective removal of nitrogen from carbazole.

In the course of gene expression studies an increased understanding of gene transcription in *Rhodococcus erythropolis* was developed. Chromosomal DNA fragments were screened for promoter activity, transcription initiation sites were mapped, and promoter regions were examined to detect conserved sequences. No DNA sequences that resemble the two conserved hexanucleotide sequences typically found in *E. coli* promoters were present in DNA fragments that functioned as promoters in *R. erythropolis*. However, possible conserved DNA sequence motifs were observed and synthetic DNA fragments were constructed to test the importance of these motifs for promoter activity in both *E. coli* and *R. erythropolis*. A DNA sequence about 10 bp upstream from the transcriptional initiation site was demonstrated to be important for promoter activity in *R. erythropolis*, but this sequence differs significantly from typical *E. coli* promoters. This information will aid in our understanding of gene regulation/expression in the genus *Rhodococcus*.

Rhodococcus cultures are capable of metabolizing a wide range of substrates, including hydrophobic substrates like petroleum, and may find widespread use in the development of biotechnology processes in the future. The results of this project will directly benefit the development of future biotechnology processes/projects employing *Rhodococcus* hosts.

Three manuscripts were prepared and submitted for publication based on the data obtained in this project:

"Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4" submitted to Biochemical and Biophysical Research Communications,

"Screening and Analysis of DNA Fragments that Show Promoter Activities in *Rhodococcus erythropolis*" manuscript in preparation,

"Microbial Biocatalyst Developments to Upgrade Fossil Fuels" Current Opinion in Biotechnology (In Press).

REFERENCES

[1] Swaty, T.E. (2005) Global refining industry trends: the present and future. Hydrocarbon Processing. 84: 35-46.

[2] Shin, S., K. Sakanashi, I. Mochida, D. A. Grudoski, and J. H. Shin (2000) Identification and reactivity of nitrogen molecular species in gas oil. Energy & Fuels. 14: 539-544.

[3] Hegedus, L.L., and McCabe, R. W. (1981) Catalyst Poisoning. Catalyst Reveiw. 23: 377-476.

[4] Speight, J.G., *The Chemistry and Technology of Petroleum*. 1980, New York: Marcel Dekker, Inc.

[5] Hsu, C.S., K. Qian, and W. K. Robbins (1994) Nitrogen speciation of polar petroleum compounds by compound class separation and on-line liquid chromatography-mass spectometry (LC-MS). J. High Resolution Chromatography. 17: 271-276.

[6] Creaser, C.S., F. Krokos, K. E. O'neill, M. J. C. Smith and P. G. McDowell (1993) Selective chemical ionization of nitrogen and sulfur heterocycles in petroleum fractions by ion trap mass spectrometry. J. American Society of Mass Spectrometry. 4: 322-326.

[7] Kilbane II, J.J., and S. LeBorgne, Petroleum Biorefining: the Selective Removal of Sulfur, Nitrogen, and Metals, in *Studies in Surface Science and Catalysis: Petroleum Biotechnology: Developments and Perspectives*, a.R.Q.-R. R. Vasquez-Duholt, Editor. 2004, Elsevier. p. 29-65.

[8] Kilbane II, J.J., A. Daram, J. Abbasian, and K. J. Kayser (2002) Isolation and characterization of carbazole-degrading bacterium *Sphingomonas* sp. GTIN11. Biochem. Biophys. Res. Comm. 297: 242-248.

[9] Technology, I.o.G. (1991) Energy Statistics. Energy Statistics. 14.

[10] Kassler, P., World energy demand outlook, in *Energy exploration and exploitation*, G. Jenkins, Editor. 1996, Multi-Science Publishing Co: Berkshire, United Kingdom. p. 229-242.

[11] Sambrook, J., E. F. Fritsch, and T. Maniatis (eds.) (1989) Molecular Cloning: a Laboratory Manual. Cold Spring Harbour, NY, USA.

[12] Watson, G.K. and Cain, R.B. (1975) Microbial metabolism of the pyridine ring. Metabolic pathways of pyridine biodegradation by soil bacteria. Biochem J. 146: 157-172.

[13] Miller, F.D. and Hershberger, C.L. (1984) A quantitative beta-galactosidase alpha-complementation assay for fusion proteins containing human insulin B-chain peptides. Gene. 29: 247-250.

[14] Kayser, K.J. and Kilbane, J.J., 2nd (2001) New host-vector system for Thermus spp. based on the malate dehydrogenase gene. J Bacteriol. 183: 1792-1795.

[15] Kayser, K.J., Kwak, J.H., Park, H.S., and Kilbane, J.J., 2nd (2001) Inducible and constitutive expression using new plasmid and integrative expression vectors for Thermus sp. Lett Appl Microbiol. 32: 412-418.

[16] Johnson, D.R., Lee, P.K., Holmes, V.F., and Alvarez-Cohen, L. (2005) An internal reference technique for accurately quantifying specific mRNAs by real-

time PCR with application to the tceA reductive dehalogenase gene. Appl Environ Microbiol. 71: 3866-3871.
[17] Bressler, D.C. and Fedorak, P.M. (2000) Bacterial metabolism of fluorene, dibenzofuran, dibenzothiophene, and carbazole. Can J Microbiol. 46: 397-409.
[18] Bressler, D.C., and M. R. Gray (2002) Hydrotreating chemistry of model products from bioprocessing of carbazoles. Energy & Fuels. 16: 1076-1086.
[19] Takeo, M., T. Fujii, and Y. Maeda (1998) Sequence analysis of the genes encoding a multicomponent dioxygenase involved in oxidation of aniline and o-toluidine in *Acinetobacter* sp. strain YAA. J. Fermentation and Bioengineering. 85: 17-24.
[20] McClure, N.C. and Venables, W.A. (1986) Adaptation of Pseudomonas putida mt-2 to growth on aromatic amines. J Gen Microbiol. 132: 2209-2218.
[21] Lyons, C.D., Katz, S., and Bartha, R. (1984) Mechanisms and pathways of aniline elimination from aquatic environments. Appl Environ Microbiol. 48: 491-496.
[22] Schnell, S., and B. Schink (1991) Anaerobic aniline degradation via reductive deamination of 4-aminobenzoyl-CoA in *Desulfobacterium anilini.* Arch Microbiol. 155: 183-190.
[23] Schnell, S., Bak, F., and Pfennig, N. (1989) Anaerobic degradation of aniline and dihydroxybenzenes by newly isolated sulfate-reducing bacteria and description of Desulfobacterium anilini. Arch Microbiol. 152: 556-563.
[24] Biegert, T., Fuchs, G., and Heider, J. (1996) Evidence that anaerobic oxidation of toluene in the denitrifying bacterium Thauera aromatica is initiated by formation of benzylsuccinate from toluene and fumarate. Eur J Biochem. 238: 661-668.
[25] Schuhle, K. and Fuchs, G. (2004) Phenylphosphate carboxylase: a new C-C lyase involved in anaerobic phenol metabolism in Thauera aromatica. J Bacteriol. 186: 4556-4567.
[26] Feng, Z., Hu, W., Rom, W.N., Beland, F.A., and Tang, M.S. (2002) 4-aminobiphenyl is a major etiological agent of human bladder cancer: evidence from its DNA binding spectrum in human p53 gene. Carcinogenesis. 23: 1721-1727.
[27] Chung, K.T. and Adris, P. (2003) Growth inhibition of intestinal bacteria and mutagenicity of 2-, 3-, 4-aminobiphenyls, benzidine, and biphenyl. Toxicol In Vitro. 17: 145-152.
[28] Trott, S., Burger, S., Calaminus, C., and Stolz, A. (2002) Cloning and heterologous expression of an enantioselective amidase from Rhodococcus erythropolis strain MP50. Appl Environ Microbiol. 68: 3279-3286.
[29] Seffernick, J.L., de Souza, M.L., Sadowsky, M.J., and Wackett, L.P. (2001) Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. J Bacteriol. 183: 2405-2410.
[30] Kauppi, B., Lee, K., Carredano, E., Parales, R.E., Gibson, D.T., Eklund, H., and Ramaswamy, S. (1998) Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase. Structure. 6: 571-586.
[31] Kilbane, J.J., A. Daram, J. Abbasian, and K. J. Kayser (2002) Isolation and characterization of a carbazole degrading bacterium Sphingomonas sp. GTIN11. Appl Environ Microbiol. Submitted.

[32] Kayser, K.J., C.-O. Yun, and J. J. Kilbane II (1996) The construction and use of promoter probe vectors for Rhodococcus sp. Actinomycetes. 7: 55-65.
[33] Vesely, M., M. Patek, J. Nesvera, A. Cejkova, J. Masak, and V. Jirku (2003) Host-vector system for phenol-degrading *Rhodococcus erythropolis* based on *Corynebacterium* plasmids. Appl. Microbiol. Biotechnol. 61: 523-527.
[34] van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., and Dijkhuizen, L. (2002) Molecular and functional characterization of kshA and kshB, encoding two components of 3-ketosteroid 9alpha-hydroxylase, a class IA monooxygenase, in Rhodococcus erythropolis strain SQ1. Mol Microbiol. 45: 1007-1018.
[35] Hawley, D.K. and McClure, W.R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. 11: 2237-2255.
[36] Bannantine, J.P., Barletta, R.G., Thoen, C.O., and Andrews, R.E., Jr. (1997) Identification of *Mycobacterium paratuberculosis* gene expression signals. Microbiology. 143 (Pt 3): 921-928.
[37] Patek, M., Nesvera, J., Guyonvarch, A., Reyes, O., and Leblon, G. (2003) Promoters of *Corynebacterium glutamicum.* J Biotechnol. 104: 311-323.
[38] Busby, S. and Ebright, R.H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. Cell. 79: 743-746.

APPENDIX: PUBLICATIONS RESULTING FROM THIS PROJECT

"Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4" submitted to Biochemical and Biophysical Research Communications,

"Screening and Analysis of DNA Fragments that Show Promoter Activities in *Rhodococcus erythropolis*" manuscript in preparation,

"Microbial Biocatalyst Developments to Upgrade Fossil Fuels" Current Opinion in Biotechnology (In Press).

FORMYLATION OF 2-AMINOBIPHENYL BY *PSEUDOMONAS* SP. STRAIN GTIN-G4

John J. Kilbane II*, and John Robbins

Gas Technology Institute, Des Plaines, IL;

*Corresponding author. Gas Technology Institute, 1700 S. Mt. Prospect Rd., Des Plaines IL 60018, john.kilbane@gastechnology.org (847-768-0723 phone, 847-768-0546 Fax)

Running Title: Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4

Keywords: formylation, organonitrogen, biodegradation, 2-aminobiphenyl

"Screening and Analysis of DNA Fragments that Show Promoter Activities in *Rhodococcus erythropolis*" manuscript in preparation,

"Microbial Biocatalyst Developments to Upgrade Fossil Fuels" Current Opinion in Biotechnology (In Press).

FORMYLATION OF 2-AMINOBIPHENYL BY *PSEUDOMONAS* SP. STRAIN GTIN-G4

John J. Kilbane II*, and John Robbins

Gas Technology Institute, Des Plaines, IL;

*Corresponding author. Gas Technology Institute, 1700 S. Mt. Prospect Rd., Des Plaines IL 60018, john.kilbane@gastechnology.org (847-768-0723 phone, 847-768-0546 Fax)

Running Title: Formylation of 2-aminobiphenyl by *Pseudomonas* sp. strain GTIN-G4

Keywords: formylation, organonitrogen, biodegradation, 2-aminobiphenyl

ABSTRACT

A bacterial culture was isolated from enrichment culture experiments based on its ability to metabolize the nitrogen-containing aromatic compound 2-aminobiphenyl. The culture was identified as a *Pseudomonas sp*., based on DNA sequence analysis of a portion of its 16S rRNA gene, and was given the designation GTIN-G4. *Pseudomonas sp*. GTIN-G4 was shown to be a facultative nitrogen-fixing microbial culture that metabolizes 2-aminobiphenyl and related organonitrogen compounds under conditions where nitrogen fixation was required for growth. GC-MS analyses demonstrated that 2-aminobiphenyl is converted to 2-formamidylbiphenyl. Similarly, o-toluidine, aniline, and diphenylamine were metabolized to their corresponding formylated derivative by *Pseudomonas sp*. GTIN-G4. However, further degradation products of these organonitrogen compounds were not detected.

INTRODUCTION

Aminobiphenyl compounds are used in the manufacture of chemicals and pharmaceuticals, are components of pesticides and antimicrobials, and are produced as combustion byproducts [1-3]. Consequently, aminobiphenyl compounds are commonly found as environmental contaminants. Aminobiphenyl compounds are toxic and carcinogenic [3]. 4-aminobiphenyl has been shown to react with DNA and cause mutagenesis. 2-aminobiphenyl and 3-aminobiphenyl are less mutagenic and their primary mechanism of toxicity is thought to be reactions with proteins and interference with the electron transport system in cells [2, 3].

Despite the widespread use of aminobiphenyl compounds in industry and their common occurrence as environmental contaminants surprisingly little is known about their biodegradation. There is an abundant literature concerning the biodegradation of biphenyl and polychlorinated biphenyls under aerobic and anaerobic conditions [4], but there are no reports of the microbial biodegradation of 2-aminobiphenol or 2-hydroxybiphenyl. This is undoubtedly due to the toxicity of these compounds. When benzidine and benzidine analogues were tested for toxicity to eighteen different microbial cultures 2-aminobiphenyl was found to be the most inhibitory organonitrogen compound

[1]. 2-aminobiphenyl was also shown to inhibit the growth of most intestinal bacteria [2]. When ten ortho-substituted biphenyls were tested for oxidation by the type II methanotroph strain CSC1 the only compounds that were not oxidized were 2-aminobiphenyl, 2-nitrobiphenyl, and 2-hydroxybiphenyl illustrating that even enzymes with extremely broad substrate ranges, like methane monooxygenase, fail to metabolize 2-aminobiphenyl [5].

In this manuscript we report the isolation of a facultative nitrogen-fixing culture, *Pseudomonas* sp. GTIN-G4, which metabolizes 2-aminobiphenyl and related compounds by addition of a formaldehyde group to the nitrogen atom.

MATERIALS AND METHODS

Chemicals

Aniline, 2-aminobiphenyl, o-toluidine, and diphenylamine were purchased from Aldrich Chemical Company (Milwaukee, WI). Ammonium Chloride was purchased from Fisher Scientific (Fair Lawn, NJ).

Media and Growth Conditions

Nitrogen-free medium A2 is composed of 5.0 g glucose, 6.3 g KH_2PO_4, 8.0 g K_2HPO_4, 0.2 g $MgSO_4$, 1.0 ml vitamin solution and 2.0 ml metal solution per liter. The vitamin solution contains 400 mg Ca-pantothenate, 400 mg niacin, 400 mg pyridoxine-HCL, 200 mg inositol, 200 mg p-aminobenzoate, 200 mg D-biotin, 50 mg riboflavin, 50 mg folic acid, 15 mg thiamine HCL and 50 mg cyanocobalamin per liter of distilled water. The metal solution contains 100 mg $MnSO_4$, 31.7 mg $CuCl_2$, 25 mg $(NH_4)_6Mo_7O_{24}$, 20 mg H_3BO_3, 25 mg $Co(NO_3)_2$, 25 mg $ZnCl_2$ and 10 mg NH_4VO_3 per liter of distilled water. Carbon sources used included a mixed carbon source consisting of 5 g/l each of glucose, glycerol, and succinate, as well as lactate or ethanol at 20 mM. Ammonium chloride and ammonium nitrate were added at concentrations ranging from 0 to 1 mM using sterile aqueous stock solutions. Organonitrogen compounds were added as neat liquids or as crystals to avoid the introduction of alternative carbon sources in the form of solvents. The concentration of organonitrogen compounds present in microbial growth or biodegradation experiments ranged from approximately 1 micromolar to 1 mM. Cultures were grown at 30°C in Erlenmeyer flasks for aerobic incubations, and in sealed serum bottles containing various headspace gases for anaerobic incubations. For growth under

conditions when a defined headspace gas composition was required 50 ml of A2 media in a 250 ml serum bottle was first sparged with either nitrogen or argon, and then crimp-sealed with a butyl rubber stopper. Sterile syringes were used to withdraw gas and to add certified gases of known composition into the sealed serum bottles using All experiments were performed using an initial gas pressure of 1 atmosphere. Nutrient agar and A2 medium solidified with 1.5% agar was used to obtain microbial colonies.

Species Identification

Bacterial species identification was done by determining the DNA sequence of a 500 bp portion of the 16S-rRNA gene (MIDI Labs, Newark NJ) and comparison with a library of known DNA sequences in the ribosomal database project [6].

Sample Preparation

The presence of metabolites in inoculated samples and sterile controls was investigated by solvent extraction. Some samples were first processed by centrifugation at 10,000 x g to remove bacterial cells, and for other samples the entire cell suspension was subjected to solvent extraction. The samples were adjusted to pH 3.0 using 6N HCl, then extracted twice with 15 ml of ethyl acetate. Water was removed from the ethyl acetate extracts using Na_2SO_4, and then each sample was evaporated to dryness under a nitrogen atmosphere using a Zymark TurboVap LV (Hopkinton, MA). Each sample was then resuspended in 1ml of acetonitrile and analyzed by thin layer chromatography and/or GC-MS.

Analytical Methods

Thin layer chromatography (TLC) was performed on Whatman Silica C-18 plates by the method described by Watson and Cain [7]. GC-MS analysis was performed using a Hewlett Packard 5971 mass selective detector and 5890 series II GC with HP 7673 auto sampler tower and a 30 meter Rezteck XTI-5 column. The final oven temperature was maintained at 300°C. The detection limit was 1 ng or 1 mg/ml with a 1 ml injection. Mass spectrographs were compared with various libraries of mass spectrograph data prepared from known standard compounds. The presence or absence of nitrogen in various compounds was also determined by GC-AED using the nitrogen-specific wavelength of 174.2 nm for detection.

RESULTS AND DISCUSSION

Enrichment culture experiments were performed under aerobic and anaerobic conditions to obtain cultures capable of utilizing 2-aminobiphenyl (2-ABP) as a sole source of carbon and/or nitrogen. Petroleum-contaminated soil and activated sludge from a municipal wastewater treatment plant were used to inoculate enrichment culture experiments. The toxicity of 2-ABP was readily apparent as it was challenging to obtain cultures that tolerate exposure to 2-ABP. The concentration of 2-ABP was gradually increased from 1 micromolar to 1 mM and a mixed culture capable of growing well in the presence of 2-ABP was obtained after nearly a year. The mixed culture could not be demonstrated to be capable of utilizing 2-ABP as a sole nitrogen and/or carbon source, but the culture became pigmented. Colored products were never observed in sterile cultures containing 2-ABP, nor when the culture was grown in the presence of 2-ABP but under conditions where nitrogen fixation was not required. It seemed likely that the discoloration of the mixed culture growing in the presence of 2-ABP under nitrogen fixation conditions was due to the production of metabolites of 2-ABP.

After repeated streaking of the 2-ABP-degrading mixed culture onto A2 agar plates containing a mixture of glucose/glycerol/succinate, lactate, or ethanol as a carbon source and 2-ABP as a nitrogen source, eventually a pure culture was obtained that produced colored metabolites when grow in the presence of 2-ABP. Lactate and ethanol were found to be preferred carbon sources for the growth of this pure culture. DNA sequence analysis of a portion of the 16S rRNA gene of this culture identified it as a *Pseudomonas* species most closely related (4.6% difference) to *Pseudomonas fluorescens*. This culture was designated as *Pseudomonas* sp. GTIN-G4. The metabolites of 2-ABP, and related compounds, formed by *Pseudomonas* sp. GTIN-G4 were investigated.

Pseudomonas sp. GTIN-G4 was found to be rather sensitive to exposure to oxygen. It prefers to grow under microaerophilic or anaerobic conditions. *Pseudomonas* sp. GTIN-G4 would not tolerate oxygen concentrations above 20%. Even momentary exposure to pure oxygen when preparing a gas headspace with a final oxygen concentration of 20% or less was lethal if oxygen was added first. However, if helium is

added first and then the headspace is adjusted to 20% oxygen *Pseudomonas* sp. GTIN-G4 can grow (data not shown).

Pseudomonas sp. GTIN-G4 produced copper, light brown, brown, and purple products when grown in A2 medium with 2-ABP, o-toluidine, aniline, or diphenylamine respectively. However, colored products were not observed when *Pseudomonas* sp. GTIN-G4 was grown in the presence of carbazole. Moreover, colored products were only observed under conditions where nitrogen fixation was required for growth. *Pseudomonas* sp. GTIN-G4 could not be demonstrated to utilize 2-ABP, or related organonitrogen compounds, as a sole source of carbon. The requirement for growth under nitrogen fixing conditions in order to observe the metabolism of 2-ABP made it extremely challenging to assess from growth studies if 2-ABP was contributing nitrogen to the growth of this culture.

To more fully define the growth capabilities of *Pseudomonas* sp. GTIN-G4 and its ability to metabolize organonitrogen compounds experiments were performed using cultures initially sparged with helium and subsequently adjusted so the gas composition was 20% oxygen. Similar experiments were performed using 20% oxygen and 80% nitrogen. These experiments demonstrated that when growth of *Pseudomonas* sp. GTIN-G4 occurs in the absence of N_2 gas, colored metabolites of 2-ABP, o-toluidine, aniline, or diphenylamine were not observed. (data not shown).

Pseudomonas sp. GTIN-G4 metabolizes 2-aminobiphenyl, diphenylamine, o-toluidine, and aniline in liquid cultures to produce colored metabolites that are not observed in sterile controls or with cultures of other bacteria incubated with these substrates. Thin layer chromatography (TLC) conformed that metabolites of the test organonitrogen compounds were being produced as a consequence of metabolism by *Pseudomonas* sp. GTIN-G4. GC-MS analyses identified some unique products that appear to result from the metabolism of *Pseudomonas* sp. GTIN-G4. Mass spectral analysis was performed on extracts of bacterial samples and samples eluted from TLC plates in an attempt to identify potential biodegradation byproducts formed by *Pseudomonas* sp. GTIN-G4 from 2-aminobiphenyl, diphenylamine, aniline, and o-toluidine. Compounds were found in all four sample types suggesting a substitution of a hydrogen on the nitrogen atom with a CHO group. GC-MS data is shown in Figure 1

illustrating the metabolite derived from the metabolism of 2-ABP by *Pseudomonas* sp. GTIN-G4 was identified as 2-formamidylbiphenyl. The compounds obtained from the metabolism of o-toluidine, and diphenylamine by *Pseudomonas* sp. GTIN-G4 are shown in Figure 2. Similar results were also obtained for the metabolism of aniline by *Pseudomonas* sp. GTIN-G4 (data not shown). In each case a formaldehyde group was added to the nitrogen atom in substitution for a hydrogen atom. No other metabolites were observed in sufficient quantity to enable identification. The proposed structure was confirmed for the diphenylamine by comparing it to a standard of N,N-diphenylformamide. The structures related to the 2-aminobiphenyl and o-toluidine samples are based on the similarities of their mass spectra to that of N,N-diphenylformamide.

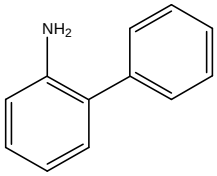
The microbial modification of 2-ABP by the addition of a formaldehyde group is without precedent, and is an interesting biochemical ability. Since there are no other publications describing the biodegradation of 2-ABP the most similar studies with which to compare these data are studies of the biodegradation of aniline. The predominant mechanism for the biodegradation of aniline is by aerobic oxidation whereby aniline dioxygenase oxidatively deaminates aniline yielding catechol [8, 9]. However, minor degradation products consisting of acylated or formylated derivatives of aniline were also observed derived from anaerobic metabolism [10]. The derivitization of compounds is commonly observed in the microbial degradation of aromatic compounds by anaerobic microorganisms. The anaerobic degradation of aniline involves the carboxylation of aniline to 4-aminobenzoate which is converted to 4-aminobenzoyl-CoenzymeA and subsequently deaminated and further degraded [11, 12].The metabolism of toluene under anaerobic conditions by the denitrifying bacterium *Thauera aromatica* is initiated by the formation of benzylsuccinate from the combination of toluene and fumarate catalyzed by benzyl succinate synthase [13]. Similarly, the biodegradation of phenol by *Thauera aromatica* involves phosphorylation and carboxylation [14].

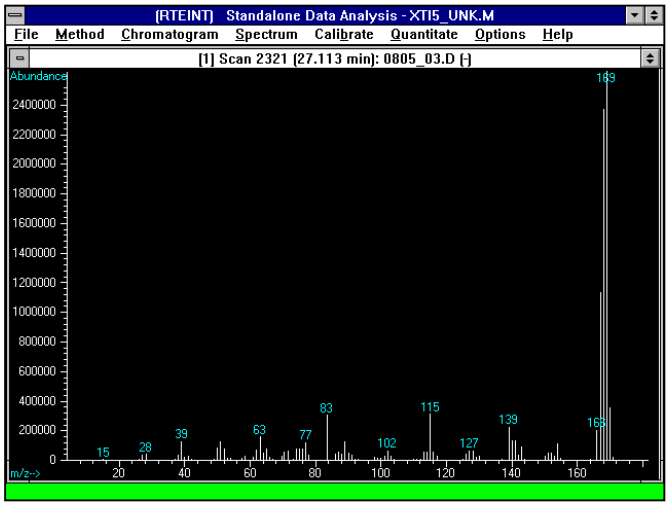
No culture capable of utilizing 2-ABP as a sole source of nitrogen and/or carbon was obtained from enrichment culture experiments, nor are there reports of the biodegradation of 2-ABP in the literature. This serves to highlight the toxicity of 2-ABP and its importance as an environmental contaminant. While the formylation of aniline is

known this is the first report of the biodegradation of 2-ABP. The formation of N-formamide derivatives of 2-aminobiphenyl, diphenylamine, aniline, and o-toluidine is a unique capability of *Pseudomonas* sp. GTIN-G4.

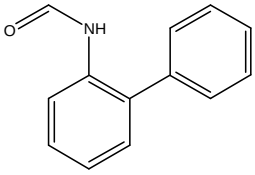
FIGURE 1. MASS SPECTRA OF 2-AMINOBIPHENYL AND A METABOLITE PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4.

2-aminobiphenyl





2-formamidylbiphenyl



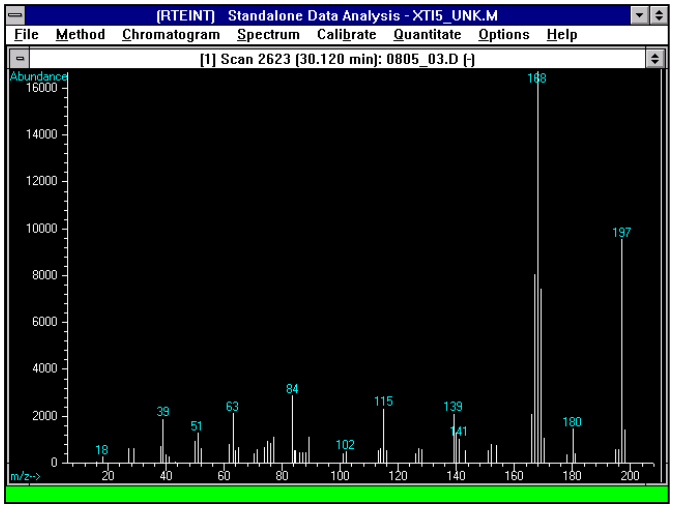
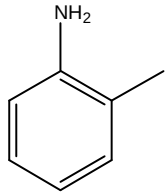
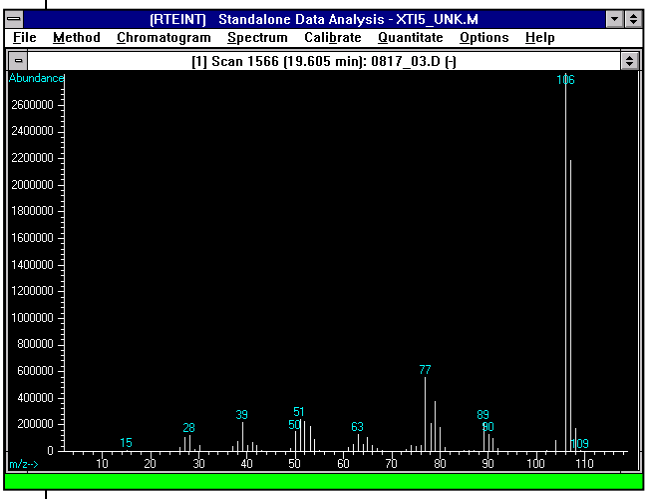


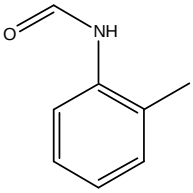
FIGURE 2. MASS SPECTRA OF O-TOLUIDINE AND DIPHENYLAMINE AND THE CORRESPONDING METABOLITES PRODUCED BY *PSEUDOMONAS* SP. GTIN-G4.

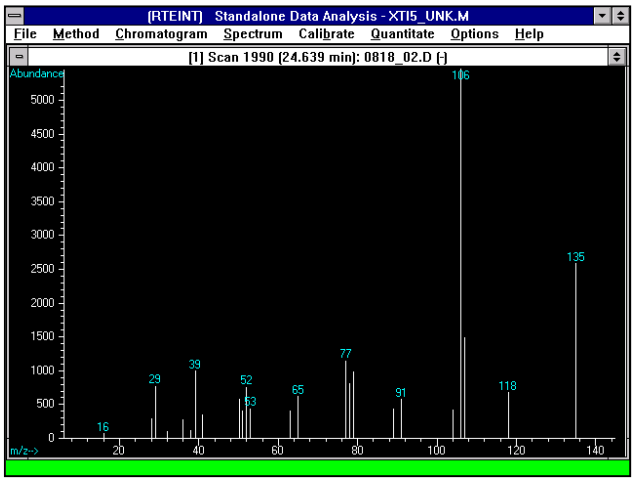
o-toluidine



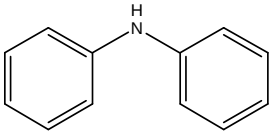


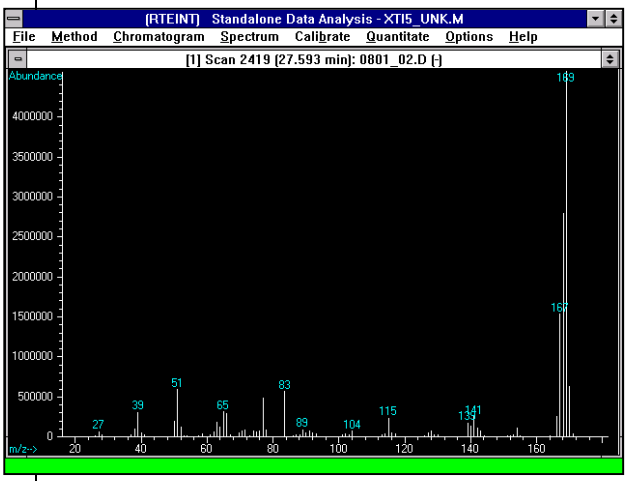
2-methylbenzylformamide



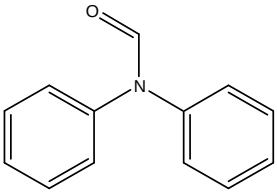


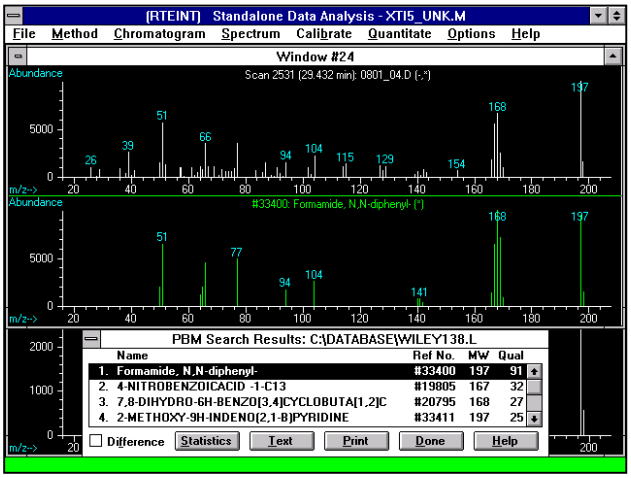
Diphenylamine





N,N-diphenyl formamide





REFERENCES

[1] Chung, K.-T., S-C. Chen, T. Y. Wong, and C-I. Wei (1998) Effects of benzidine and benzidine analogues on growth of bacteria including Azotobacter vinlandii. Environmental Toxicology and Chemistry. 17: 271-275.

[2] Chung, K.T. and Adris, P. (2003) Growth inhibition of intestinal bacteria and mutagenicity of 2-, 3-, 4-aminobiphenyls, benzidine, and biphenyl. Toxicol In Vitro. 17: 145-152.

[3] Feng, Z., Hu, W., Rom, W.N., Beland, F.A., and Tang, M.S. (2002) 4-aminobiphenyl is a major etiological agent of human bladder cancer: evidence from its DNA binding spectrum in human p53 gene. Carcinogenesis. 23: 1721-1727.

[4] Patureau, D. and Trably, E. (2006) Impact of Anaerobic and Aerobic Processes on PolyChloroBiphenyl Removal in Contaminated Sewage Sludge. Biodegradation. 17: 9-17.

[5] Lindner, A.S., Semrau, J.D., and Adriaens, P. (2005) Substituent effects on the oxidation of substituted biphenyl congeners by type II methanotroph strain CSC1. Arch Microbiol. 183: 266-276.

[6] Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker, C.T., Saxman, P.R., Farris, R.J., Garrity, G.M., Olsen, G.J., Schmidt, T.M., and Tiedje, J.M. (2001) The RDP-II (Ribosomal Database Project). Nucl. Acids. Res. 29: 173-174.

[7] Watson, G.K. and Cain, R.B. (1975) Microbial metabolism of the pyridine ring. Metabolic pathways of pyridine biodegradation by soil bacteria. Biochem J. 146: 157-172.

[8] Takeo, M., T. Fujii, and Y. Maeda (1998) Sequence analysis of the genes encoding a multicomponent dioxygenase involved in oxidation of aniline and o-toluidine in *Acinetobacter* sp. strain YAA. J. Fermentation and Bioengineering. 85: 17-24.

[9] McClure, N.C. and Venables, W.A. (1986) Adaptation of Pseudomonas putida mt-2 to growth on aromatic amines. J Gen Microbiol. 132: 2209-2218.

[10] Lyons, C.D., Katz, S., and Bartha, R. (1984) Mechanisms and pathways of aniline elimination from aquatic environments. Appl Environ Microbiol. 48: 491-496.

[11] Schnell, S., and B. Schink (1991) Anaerobic aniline degradation via reductive deamination of 4-aminobenzoyl-CoA in *Desulfobacterium anilini.* Arch Microbiol. 155: 183-190.

[12] Schnell, S., Bak, F., and Pfennig, N. (1989) Anaerobic degradation of aniline and dihydroxybenzenes by newly isolated sulfate-reducing bacteria and description of Desulfobacterium anilini. Arch Microbiol. 152: 556-563.

[13] Biegert, T., Fuchs, G., and Heider, J. (1996) Evidence that anaerobic oxidation of toluene in the denitrifying bacterium Thauera aromatica is initiated by formation of benzylsuccinate from toluene and fumarate. Eur J Biochem. 238: 661-668.

[14] Schuhle, K. and Fuchs, G. (2004) Phenylphosphate carboxylase: a new C-C lyase involved in anaerobic phenol metabolism in Thauera aromatica. J Bacteriol. 186: 4556-4567.

Microbial biocatalyst developments to upgrade fossil fuels

John J Kilbane II

Steady increases in the average sulfur content of petroleum and stricter environmental regulations concerning the sulfur content have promoted studies of bioprocessing to upgrade fossil fuels. Bioprocesses can potentially provide a solution to the need for improved and expanded fuel upgrading worldwide, because bioprocesses for fuel upgrading do not require hydrogen and produce far less carbon dioxide than thermochemical processes. Recent advances have demonstrated that biodesulfurization is capable of removing sulfur from hydrotreated diesel to yield a product with an ultra-low sulfur concentration that meets current environmental regulations. However, the technology has not yet progressed beyond laboratory-scale testing, as more efficient biocatalysts are needed. Genetic studies to obtain improved biocatalysts for the selective removal of sulfur and nitrogen from petroleum provide the focus of current research efforts.

Addresses

Gas Technology Institute, 1700 South Mount Prospect Road, Des Plaines, IL 60018, USA

Corresponding author: Kilbane II, John J (john.kilbane@gastechnology.org)

Current Opinion in Biotechnology 2006, **17**:XXX–XXX

This review comes from a themed issue on Energy biotechnology

Edited by Jonathan R Mielenz

0958-1669/$ – see front matter
© 2006 Elsevier Ltd. All rights reserved.

Introduction

The interest in bioprocesses for upgrading petroleum stems from the fact that the average sulfur content of globally available petroleum is steadily increasing, and the environmental regulations concerning the sulfur content of petroleum-derived products have become increasingly strict [1••]. Treating heavier more sour crude oil is the reality for refineries of the future, while at the same time facing increased pressure to reduce carbon dioxide emissions and make ultra-low sulfur transportation fuels. The maximum allowable sulfur and aromatic concentration for diesel in the United States is currently 15 ppm and 36%, respectively, with targets of 10 ppm sulfur and 15% aromatics by 2010 [1••]. About 72% of oil is refined into transportation fuels in the US and this yield will be increasingly hard to obtain as the quality of crude declines. This means that additional/increased processing will be needed just to produce the same quality and volume of products currently produced. Refinery equipment constraints dictate the ability to process heavy crude and to produce low-sulfur transportation fuels. So that although major improvements in hydrotreating technology have occurred in recent years, which enable the production of low-sulfur transportation fuels [2], demand for transportation fuels has expanded more rapidly than existing refinery capacity can handle, especially while maintaining product sulfur concentrations within regulatory limits [1••]. This trend is expected to continue and will be driven by expanding economies in China, India and elsewhere.

This article highlights research progress on bioprocessing to upgrade fossil fuels over the past few years and will focus on the development of bioprocesses to selectively remove sulfur, nitrogen and metals. Other recent reviews that address petroleum microbiology [3•] and bioprocesses for upgrading petroleum [4•] are good sources of additional information. This area of research was addressed in *Current Opinion in Biotechnology* in 2000 [5] and in *Current Opinion in Microbiology* in 2003 [6]. These articles focused exclusively on biodesulfurization and gave a concise account of the status of research at that time.

Biodesulfurization

About 15 to 20 years ago biodesulfurization research targeted coal with an interest in the removal of both inorganic (pyrite) and organic sulfur by microorganisms; however, in recent years coal biodesulfurization research has become less common [7,8]. The microbial leaching of inorganic sulfur from coal has progressed to the level of industrial-scale tests, but the removal of organic sulfur from coal remains problematic [9]. The difficulties in accessing sulfur in coal as compared with petroleum, combined with increasingly stringent environmental regulations regarding the sulfur content of transportation fuels [1••], has resulted in most biodesulfurization research being focused on organic sulfur in petroleum. Petroleum contains a complex mixture of chemical structures, but the organosulfur compounds of greatest interest for biodesulfurization are benzothiophenes and DBTs, primarily because of their relative recalcitrance to removal by refinery processes, specifically hydrodesulfurization [4•].

DBT has become the model compound for biodesulfurization research. Numerous microbial cultures, including Gram-positive and Gram-negative bacterial cultures and yeast, have been isolated based on their ability to utilize DBT (or related molecules) as sole sources of sulfur. Although occasional reports of anaerobic bacteria that can desulfurize DBT have occurred, the rate and extent of petroleum desulfurization by currently available anaerobic cultures and the lack of knowledge of the biochemistry and genetics of such microorganisms makes the development of a commercial process unlikely [10]. Much more is known about the pathway for the metabolism of DBT by aerobic microorganisms, and the pathway for the oxidative desulfurization of DBT is illustrated in Figure 1. An industrial-scale process for petroleum biodesulfurization using aerobic microorganisms has not yet been demonstrated either, but through an improved understanding of the biochemistry and genetics of the desulfurization pathway it

is anticipated that improved biocatalysts with activities suitable for an industrial process will be developed. Even though alternative pathways for desulfurization of benzothiophenes exist, the oxidative desulfurization pathway pictured in Figure 1 has been the focus of the majority of research.

The biodesulfurization pathway

The *dszC* gene encodes the dibenzothiophene monooxygenase (DszC) that catalyzes the conversion of DBT into DBT sulfone (DBTSO2). The *dszA* gene encodes the dibenzothiophene-5,5-dioxide monooxygenase (DszA) that catalyzes the conversion of DBTSO2 into 2-hydroxybiphenyl-2-sulfinate (HBPSi). The *dszB* gene encodes 2-hydroxybiphenyl-2-sulfinate sulfinolyase (DszB) that catalyzes the conversion of HBPSi into 2-hydroxybiphenyl (2-HBP) and sulfite. The *dszABC* genes are transcribed as an operon found on a large plasmid, pSOX, in many desulfurization-competent bacteria [11]. An unlinked fourth gene, the *dszD* gene encoding an NADH–FMN oxidoreductase (DszD), is an accessory component of the desulfurization pathway and allows the regeneration of the cofactors needed for the monooxygenase reactions catalyzed by DszC and DszA. The enzymology of the oxidative desulfurization pathway has been firmly established using purified enzymes from several desulfurization competent bacterial species and from the results of genetic analyses [6,12]. Benzothiophene and DBT in petroleum chiefly exist as alkylated derivatives and, generally, as the number and size of alkyl group substitutions increase these compounds are less suitable as substrates for the desulfurization enzymes [13]. Moreover, the accumulation of oxidized intermediates (sulfones and sultines) when crude oil is biodesulfurized indicates that DszA and DszB have a more restricted substrate range than DszC [13]. Furthermore, DszB has been shown to catalyse the rate-limiting step in the desulfurization pathway for *Rhodococcus erythropolis* IGTS8 [12], although the enzyme analogous with DszC (BdsC) appears to be the rate-limiting enzyme in the desulfurization pathway of *Bacillus subtilis* WU-S2B [14].

The role of biodesulfurization in nature

Insight as to what role the desulfurization trait plays in nature is provided by the finding that in most soils sulfur is present predominantly as sulfonates and sulfate esters, rather than as sulfate, and the utilization of organosulfur compounds may be crucial for the survival of many bacterial cultures [15•]. The desulfurization trait is widely dispersed in the microbial kingdom and diverse species have often been found to possess identical, or nearly identical, *dsz* gene sequences [4•,11•,16] — with generally >99% identity to the *dsz* genes of *R. erythropolis* IGTS8 (ATCC 53968), which was the first desulfurization-competent bacterial culture to be isolated and characterized. The *dsz* genes of *Gordonia alkinovorans* strain 1B show about 88% identity with the *dsz* genes of *R. erythropolis* IGTS8, but the overall gene order and structure is preserved including an overlap of the *dszA* and *dszB* genes, which is 4 bp in *R. erythropolis* IGTS8 but 64 bp in *G. alkinovorans* strain 1B [17]. Distinct variants of the *dsz* genes have also been found, particularly among thermophilic bacteria. Thus far, only moderately thermophilic cultures have been isolated that have the ability to desulfurize DBT at temperatures up to 52°C. The first such culture to be identified was *Paenibacillus* sp. A11-2 and the DNA sequences of its desulfurization genes *tds*ABC are 73%, 61% and 52% homologous with the *dszABC* genes, respectively [18]. More recently, other desulfurization-competent moderately thermophilic cultures have been isolated, and although the desulfurization genes of *B. subtilis* WU-S2B and *Mycobacterium phlei* WU-F1 are identical to each other, they exhibit 61% and 58% homology with the *dszABC* and *tdsABC* genes, respectively [19•].

It is not surprising that different species of desulfurization-competent bacteria that contain different desulfurization genes should possess different desulfurization rates and substrate ranges, but a most interesting recent finding is that bacterial cultures that possess identical *dsz* gene sequences can have vastly different desulfurization phenotypes. This was clearly illustrated by the characterization of the moderate thermophile *Mycobacterium phlei* GTIS10, which was shown to have *dszABC* gene sequences identical to *R. erythropolis* IGTS8; however, the temperature at which maximum desulfurization activity was detected in the cultures was about 50°C and 30°C, respectively, and the concentrations of metabolites produced by the two cultures varied [11•]. Similarly, the characterization of five bacterial cultures isolated from enrichment cultures on the basis of their ability to utilize DBT as a sole source of sulfur revealed that four of the cultures had identical *dsz* genes, but the cultures differed significantly as regards their substrate range, desulfurization activity, yield of metabolites, and other parameters [20•]. However, the most profound demonstration of the effect of bacterial host contributions to the functioning of the desulfurization pathway involved the use of the *dszABCD* genes on a transposable element. When the *dszABCD* transposon was introduced into 22 *Rhodococcus* or *Mycobacterium* cultures a wide range of desulfurization phenotypes were found [21]. Although all the cultures could desulfurize DBT, 9 of the 22 could desulfurize 4,6-diethylDBT, and only one culture could desulfurize 4,6-dipropylDBT. The desulfurization pathway requires NADH, $FMNH_2$ and oxygen in order to complete the conversion of DBT to 2-HBP. The host must supply these factors and the range of desulfurization phenotypes observed in different cultures could reflect the ability of each bacterial species to provide cofactors and reaction substrates under the conditions tested. Transport of substrates and products might also contribute to desulfurization activity, as demonstrated by the fact that cell-free lysates of desulfurization cultures can exhibit broader substrate range than the intact cell culture [22].

The issue of substrate transport in desulfurization was investigated in *Pseudomonas putida* IFO13696, which is a recombinant strain containing cloned *dsz* genes [23•]. This culture can desulfurize DBT in water, but not in tetradecane unless the *hbcABC* genes (encoding substrate transport proteins) were also provided. Another recent study sought to improve the desulfurization performance of *Pseudomonas*

delafieldii R-8 by adding Al_2O_3 nanoparticles to the surface of the bacterial cells to facilitate adsorption of DBT. This treatment was shown to double the rate of desulfurization of DBT in *n*-dodecane by *P. delafieldii* R-8 [24••]. However, the ability of *Pseudomonas* cultures versus *Rhodococcus* cultures to access hydrophobic substrates, particularly as regards to biodesulfurization, is known to differ [4•,5]. No DBT transport genes or proteins have been identified in *Rhodococcus* or *Mycobacterium* cultures, and it is not yet known to what degree substrate and/or product transport functions contribute to the desulfurization phenotype of most cultures. However, when considering substrate/product transport in desulfurization it is good to keep in mind that desulfurization-competent bacterial cultures can remove organic sulfur from coal [4•,8]. There is no evidence that desulfurization enzymes are excreted from desulfurization-competent cells, but the size of substrates metabolized and the ability of other bacterial species to successfully compete for sulfur liberated from organosulfur substrates by desulfurization-competent cultures make it likely that desulfurization in some cultures occurs in association with the external surface of cells [4•].

The observation that the *dsz* operon had two apparent temperature maximums in two different bacterial hosts suggests that if the *dsz* operon were to be expressed in a thermophilic bacterial host, the desulfurization enzymes could function at even higher temperatures. This possibility was tested in gene expression studies employing the extreme thermophile *Thermus thermophilus* [25••]. The *dszC* gene from *R. erythropolis* IGTS8 was successfully expressed in *T. thermophilus*, but the thermolability of the other enzymes prevented the functional expression of the full desulfurization pathway in this host. With the increased availability of desulfurization-competent bacterial cultures it is likely that the complete desulfurization pathway could eventually be expressed in *T. thermophilus*. Also, because this host grows at temperatures ranging from 55°C to 85°C it might be possible to employ directed evolution and selective pressure to gradually evolve desulfurization enzymes capable of functioning at higher temperatures so that they are suitable for a high-temperature bioprocess [25••,26].

Engineering strains for biodesulfurization

The main goal of biodesulfurization research is to develop a commercial biodesulfurization process for petroleum, and it has been estimated that a successful commercial process would require a biocatalyst with desulfurization activity of 1.2 mM DBT/g DCW/h [5,27•] to 3 mM DBT/g DCW/h [19•]. The desulfurization activity of various cultures is listed in Table 1. The desulfurization activity of naturally occurring bacterial cultures is low in comparison to the requirements of a commercial process and genetic manipulation has been used to achieve higher desulfurization rates. Efforts to provide multiple copies of the *dsz* genes of *R. erythropolis* KA 2-5-1 and to place them under the control of alternative promoters resulted in derivative cultures in which the maximum desulfurization activity was increased from ~50 to ~250 μmol/g DCW/h [28,29•,30]. *Escherichia coli–Rhodococcus* shuttle vectors have been available for the genetic manipulation of *Rhodococcus* cultures for many years [4•,31•], but more recently gene expression vectors for *Rhodococcus* have become available [32•]. Inducible gene expression in *Rhodococcus* has been demonstrated as well as the expression of several heterologous genes that could not be efficiently expressed in *E. coli* [32•].

An interesting feature of genetic experiments involving the desulfurization genes is that the level of desulfurization activity achieved by genetic manipulation appears to be limited by factors that are not yet understood. What this statement means is that the level of increased gene expression accomplished/anticipated by the use of alternative promoters and/or through supplying multiple copies of the *dsz* genes exceeds the amount of desulfurization activity observed in the resulting cultures [29•,30,33–36]. As observed in studies of the expression of *dsz* genes in various hosts, it is evident from genetic experiments employing alternative promoters and/or increased copy numbers for the expression of the *dsz* genes that host contributions are important to the functioning of the desulfurization pathway. In order to achieve very high levels of expression of the desulfurization pathway, a better understanding of the host factors that contribute to the functioning of the pathway is needed. The genome of *Rhodococcus* sp. strain RHA1 is currently being sequenced [31•] and the availability of this and related genome sequences will enable future studies to characterize genome-wide gene expression and metabolic fluxes related to the expression of the desulfurization pathway. The desulfurization enzymes might be the rate-limiting factors in achieving desulfurization activity up to a certain level, but in order to achieve even higher activities alterations in genes and regulatory sequences other than (or in addition to) the *dsz* genes are needed. It is known from the study of the expression of metabolic pathways in other microbial species that alterations of transcriptional regulators/sequences as well as mutations in enzyme-encoding sequences are needed to enhance metabolic flow through a pathway [37]. Transcriptional regulators that are responsive to the concentration of various organosulfur compounds and that impact the expression of multiple genetic loci are known for some species [38], and an improved understanding of transcriptional and post-transcriptional gene regulation in *Rhodococcus* and other desulfurization-competent hosts will greatly benefit efforts to create biocatalysts with ever-higher desulfurization levels.

The levels of desulfurization activity already achieved with various cultures are quite impressive, as shown in Table 1.

The highest desulfurization rates were obtained in tests employing DBT in model systems rather than in the treatment of actual petroleum products. A genetically modified strain of *R. erythropolis* KA 2-5-1 was capable of achieving desulfurization activities of ~250 μmol/g DCW/h, but it is important to note that no significant increase in the desulfurization activity of this culture as a result of genetic modification has been reported in the past five years [29•,30]. In a similar study, the desulfurization rate of a

Mycobacterium sp. G3 culture was observed to increase from 178 to 211 μmol/g DCW/h as a consequence of genetic manipulation [35]. The study of desulfurization by *Mycobacterium* sp. G3 was also interesting, because it revealed that even after the expression of the desulfurization genes were placed under the control of a strong promoter (P*hsp60*) not known to be regulated by sulfur, the desulfurization activity declined dramatically in the presence of sulfate [35]. Whether transcriptional, post-transcriptional or a combination of effects was responsible for the decreased level of desulfurization activity in the presence of sulfate in these experiments is not known. As discussed above, desulfurization activities of about 250 μmol/g DCW/h could be the maximum levels that can be achieved through the manipulation of the desulfurization genes alone, and alterations to other host functions may be required in order to achieve higher levels of desulfurization.

The desulfurization rates obtained with *R. erythropolis* XP [39••] and *Mycobacterium* sp. X7B [40] (listed in Table 1) are particularly noteworthy, because these rates were derived from experiments employing petroleum products. Both studies employed diesel oil that had previously been treated by hydrodesulfurization so that only organosulfur compounds that were relatively recalcitrant to hydrodesulfurization remained. The desulfurization rates obtained with diesel oil are about 4 μmol/g DCW/h, but significantly biodesulfurization was capable of achieving a final sulfur concentration of 14 ppm. This work demonstrates that a biodesulfurization process can be used to produce diesel having ultra-low sulfur content, as required by current environmental regulations. Moreover, this impressive level of biodesulfurization was achieved even though only organosulfur compounds that were relatively recalcitrant to hydrodesulfurization were present. Other diesel biodesulfurization studies were not capable of achieving such low residual sulfur concentrations and compounds that were refractory to biodesulfurization were seen to accumulate [41•]. *Mycobacterium* sp. X7B was also shown to be capable of desulfurizing gasoline [42•]. This is important because the organosulfur compounds most commonly present in diesel oil are DBTs, while benzothiophenes predominate in gasoline. Gasoline is also more toxic to bacterial cells than diesel so the demonstration that gasoline can be successfully biodesulfurized is a significant achievement.

The ability of biodesulfurization catalysts to desulfurize DBT under standardized laboratory conditions provides a convenient way to obtain desulfurization rates that can be meaningfully compared, but the desulfurization rate obtained with crude oil or a petroleum product is the truly important measurement in determining the suitability of a biocatalyst for an industrial biodesulfurization process. Therefore, if the target desulfurization activity for a successful industrial biodesulfurization process is 1.2 to 3 mmol/g DCW/h [5,19•], as measured with petroleum products, then the currently available biocatalysts require an increase in desulfurization rate of about 500-fold.

Because biodesulfurization rates are still relatively low compared with the requirements of an industrial process, little research has focused on process development. The most common conception of a biodesulfurization process is for the treatment of diesel oil [5]. A biodesulfurization process could complement hydrotreating processes, as DBTs are recalcitrant in hydrodesulfurization but are the preferred substrates for biodesulfurization. However, the appropriate placement of a biodesulfurization process within a refinery is debatable. Some have proposed that hydrodesulfurization should occur first, then the final removal of organosulfur compounds that are recalcitrant to hydrotreating would be accomplished by biodesulfurization [5]. The increasingly stringent environmental regulations regarding the sulfur content of diesel, and other transportation fuels, requires that the final desulfurization step be capable of reliably achieving ultra-low sulfur levels. As biodesulfurization is unable to work effectively on highly alkylated compounds [13,41•], and because the composition of organosulfur compounds present in the diesel cut during the life of a refinery can vary significantly [1••], biodesulfurization does not provide the level of reliability needed to consistently achieve ultra-low sulfur concentrations. An alternative is to employ biodesulfurization before hydrotreating. In this way a significant portion of organosulfur compounds that are recalcitrant to hydrotreating can be removed, allowing the hydrotreating process to function more efficiently and consume less hydrogen. At the same time, improved hydrodesulfurization catalysts can be used to reliably achieve environmentally mandated levels of sulfur in the final product.

Energy BioSystems Corporation (EBC) was the only commercial venture dedicated to the development of biodesulfurization technology. EBC's concept for a biodesulfurization process was to treat diesel, but also to produce a value-added surfactant byproduct to achieve a more economical process [5]. This was to be accomplished through the use of biodesulfurization catalysts that lacked DszB, so that DBT and related compounds would be converted to sulfones and sulfinates that could be recovered and used as surfactants. There was a plan to construct a demonstration-scale biodesulfurization process at the Petro Star refinery in Valdez, Alaska. The date for the construction of a demonstration plant was progressively postponed, EBC changed their name to Enchira and eventually went out of business. Petro Star continued the biodesulfurization process development work in collaboration with Diversa and other partners [43••]. Diversa employed their Gene Site Saturated Mutagenesis process to obtain improved derivatives of DszA and DszC enzymes and they explored the use of multiple alternative hosts to express a truncated version of the desulfurization pathway. Although mutants with improved desulfurization activity were found, no biocatalyst with activity sufficient for an industrial process were developed. The project has now ended and Petro Star will employ improved hydrodesulfurization catalysts in any planned modifications of their refinery (D Nunn, personal communication). Ultimately, there was no serious interest in the surfactant byproduct and significant improvements in

the performance of catalysts used in hydrotreating [1••] prevented this demonstration-scale biodesulfurization process from being developed.

Potentially, the most attractive option for a biodesulfurization process is to integrate bioprocessing into existing industrial practices to the greatest degree possible and to target the treatment of heavy crude oils. A bioprocess for upgrading crude oil can potentially remove sulfur, nitrogen and metals and simultaneously reduce the viscosity/molecular weight of oil because of carbon–sulfur and carbon–nitrogen bond cleavage. If biorefining can be used in conjunction with desalting and dewatering steps in oil production operations [44] and can reduce the sulfur, nitrogen, and metal content of crude oil prior to the oil being sent to refineries, then existing refining technologies could be used with an expanded range of low-quality oils that could not otherwise be treated. Biodesulfurization could fit well with current practices in the petroleum industry if performed in conjunction with desalting and dewatering operations; however, this would require thermophilic cultures. A potentially attractive means of implementing a petroleum biodesulfurization process could be to treat heavy oil on site in the production field prior to the initial separation of petroleum from produced water [45]. A schematic illustration of a biorefining process used to treat crude oil in association with production operations is shown in Figure 2.

Biodenitrogenization of petroleum

Organonitrogen compounds in petroleum can poison catalysts used in refinery processes making the removal of nitrogen from petroleum desirable, but not subject to environmental regulation [4•]. The removal of organically bound nitrogen from crude oil, without the loss of significant calorific value, requires the selective cleavage of carbon–nitrogen bonds. The cleavage of carbon–nitrogen bonds resulting in the conversion of quinoline to 8-hydroxycoumarin and ammonia has been demonstrated [46], and the genes that encode the enzymes participating in the quinoline degradation pathway have been identified and sequenced [47]. The removal of nitrogen from crude oil by a quinoline-degrading culture, *Pseudomonas ayucida* IGTN9m, has also been demonstrated [46]. However, the abundance of quinoline relative to other organonitrogen compounds in crude oil is low and existing quinoline degradation enzymes have a narrow substrate range. Consequently, even though removal of 68% of quinoline from crude oil was demonstrated, the total nitrogen content was reduced by only 5%.

Carbazole is a good model compound that is representative of the nitrogen-containing compounds present in the greatest abundance in many petroleum fractions. None of the presently known carbazole-degrading cultures is particularly appropriate for a biological process for the removal of nitrogen from petroleum, because nitrogen is only removed in the course of complete degradation. The carbazole biodegradation pathway begins with the oxidative cleavage of the heterocyclic nitrogen ring of carbazole to form 2′-aminobiphenyl-2,3-diol. This compound is then oxidized through *meta* cleavage to yield 2-hydroxy-6-oxo-6-hexa-2e,4z-dienoate. The next metabolic steps result in the degradation of one of the aromatic rings to release carbon dioxide. In existing pathways, nitrogen is released from carbazole only after substantial carbon degradation. Figure 3 illustrates the carbazole degradation pathway employed by currently known carbazole-utilizing cultures. A potential pathway for the selective removal of nitrogen from carbazole is also shown, which could be created using metabolic engineering to combine the CarA enzyme from carbazole degraders such as *Sphingomonas* sp. GTIN11 with a suitable deaminase. There are no known deaminases that can metabolize 2′-aminobiphenyl-2,3-diol and accomplish the cleavage of the final C–N bond [4•,48•]. It is important that complete removal of nitrogen from carbazole and other organonitrogen compounds be achieved in a biodenitrogenization process, because microbial metabolites derived from the partial degradation of carbazole (e.g. 2-aminobiphenyl-2,3-diol) experience condensation, ring closure and polymerization when processed by hydrotreatment [49••]. Therefore, a combination of biotreatment and hydrotreatment will not work to achieve the removal of nitrogen from petroleum and an effective biocatalyst for the selective removal of nitrogen from petroleum remains to be developed. However, a microbial culture, *Gordonia* sp. strain F.5.25.8, was recently reported that has the ability to simultaneously metabolize DBT and carbazole [50•]. The use of enzymes to remove metals from petroleum has also been reported [4•,51•]. Cytochrome *c* and various peroxidases can be used to oxidize porphyrin compounds, resulting in the release of chelated metals, chiefly nickel and vanadium from petroleum [51•]. This suggests the possibility that future biocatalysts for the simultaneous removal of sulfur, nitrogen and metals from petroleum could be developed.

Conclusions

The development of bioprocesses for upgrading fossil fuels ranks among the greatest challenges addressed by biotechnology. The vast majority of commercial bioprocesses produce pharmaceuticals or fine chemicals where process costs are higher and product volumes are orders of magnitude lower than the requirements for a bioprocess to upgrade fossil fuels. However, as all industries mature, making large volumes of product at ever lower per unit cost becomes the dominant activity. So, when viewed from this perspective, efforts to develop bioprocesses for upgrading fossil fuels can be seen as addressing the issues important to the development of the biotechnology industry of the future. The key research needs for bioprocesses to upgrade fossil fuels are the development of desulfurization biocatalysts with higher specific activity, broader substrate range, and higher thermal tolerance. An improved understanding of host contributions to the functioning of the desulfurization pathway will be needed to achieve these goals. For the denitrogenization of petroleum, we need to identify a deaminase that efficiently cleaves the carbon–nitrogen bond in 2-aminobiphenyl and related compounds so that a metabolic pathway for the complete removal of nitrogen from carbazole can be constructed.

Acknowledgements

The support of the US Department of Energy funding through contract DE-FC26-02NT15382 is gratefully acknowledged.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

••1. Swaty TE: **Global refining industry trends: the present and future.** *Hydrocarbon Processing* 2005, **84**:35-46.

This article summarizes the trends in sulfur content and viscosity of crude oil over the past few decades and extrapolates the trends to the future. Environmental regulations and future regulatory targets for sulfur and aromatic content for petroleum products throughout the world are listed. Current and future petroleum refinery capacity is discussed, along with current and future petroleum demand.

2. Song C: **An overview of new approaches to deep desulfurization for ultra-clean gasoline, diesel fuel and jet fuel.** *Catal Today* 2003, **86**:211-263.

•3. Van Hamme JD, Singh A, Ward OP: **Recent advances in petroleum microbiology.** *Microbiol Mol Biol Rev* 2003, **67**:503-549.

The diverse applications of microbiology to the petroleum industry are summarized in this review article. Topics addressed include hydrocarbon bioremediation, fuel upgrading, microbiologically enhanced oil recovery, fine chemical production, and pollutant biosensors.

•4. Kilbane JJ II, LeBorgne S: **Petroleum biorefining: the selective removal of sulfur, nitrogen, and metals**. In *Studies in Surface Science and Catalysis: Petroleum Biotechnology: Developments and Perspectives*. vol 151. Edited by Vasquez-Duhalt R, Quintero-Ramirez R: Elsevier; 2004:29-65.

A thorough review of petroleum biorefining is provided summarizing the use of bioprocesses for the selective removal of sulfur, nitrogen and metals from petroleum.

5. Monticello DJ: **Biodesulfurization and the upgrading of petroleum distillates.** *Curr Opin Biotechnol* 2000, **11**:540-546.

6. Gray KA, Mrachko GT, Squires CH: **Biodesulfurization of fossil fuels.** *Curr Opin Microbiol* 2003, **6**:229-235.

7. Demirbas A, Balat M: **Coal desulfurization via different methods.** *Energy Sources* 2004, **26**:541-550.

8. Baek K, Kim C-S, Lee H-H, Shin H-J, Yang J-W: **Microbial desulfurization of solubilized coal.** *Biotechnol Lett* 2002, **24**:401-405.

9. Cara J, Carballo MT, Moran A, Bonilla D, Escolano O, Garcia Frutos FJ: **Biodesulphurization of high sulphur coal by heap leaching.** *Fuel Proc Technol* 2005, **84**:1905-1910.

10. Bahrami A, Shojaosadati SA, Mohebali G: **Biodegradation of dibenzothiophene by thermophilic bacteria.** *Biotechnol Lett* 2001, **23**:899-901.

11. Kayser KJ, Cleveland L, Park H-S, Kwak J-H, Kolhatkar A, Kilbane JJ II: **Isolation and characterization of a moderate thermophile, *Mycobacterium phlei* GTIS10, capable of dibenzothiophene desulfurization.** *Appl Microbiol Biotechnol* 2002, **59**:737-746.

12. Watkins LM, Rodriguez R, Schneider D, Broderick R, Cruz M, Chambers R, Ruckman E, Cody M, Mrachko GT: **Purification and characterization of the aromatic desulfinase 2-(2′-hydroxybiphenyl)benzensulfinate desulfinase.** *Arch Biochem Biophys* 2003, **415**:14-23.

13. Prince RC, Grossman MJ: **Substrate preferences in biodesulfurization of diesel range fuels by rhodococcus sp. strain ECRD-1.** *Appl Environ Microbiol* 2003, **69**:5833-5838.

14. Ohshiro T, Ishii Y, Matsubara T, Ueda K, Izumi Y, Kino K, Kirimura K: **Dibenzothiophene desulfurizing enzymes from moderately thermophilic bacterium *Bacillus subtilis* WU-S2B: purification, characterization and overexpression.** *J Biosci Bioeng* 2005, **100**:266-273.

•15. Mirleau P, Wogelius R, Smith A, Kertesz MA: **Importance of organosulfur utilization for survival of *Pseudomonas putida* in soil and rhizosphere.** *Appl Environ Microbiol* 2005, **71**:6571-6577.

In this study, it was shown that sulfur in most soils is present as sulfonates and sulfate esters, making the ability to metabolize organosulfur compounds essential for the survival of various microbial species in a range of natural habitats. This observation can, at least in part, explain why the desulfurization pathway evolved and why it is found in diverse microbial species.

16. Castorena G, Suarez C, Valdez I, Amador G, Fernandez L, LeBorgne S: **Sulfur-selective desulfurization of dibenzothiophene and diesel oil by newly isolated rhodococcus sp. strains.** *FEMS Microbiol Lett* 2002, **215**:157-161.

17. Alves L, Melo M, Mendonca D, Simoes F, Matos J, Tenreiro R, Girio FM: **Dibenzothiophene desulfurization using recombinant *Escherichia coli* carrying the *dsz* genes of *Gordonia alkanivorans* strain 1B**. *Appl Biochem Biotechnol* 2005, **120**:199-208.

18. Ishii Y, Konoshi J, Okada H, Hirasawa K, Onaka T, Suzuki M: **Operon structure and functional analysis of the genes encoding thermophilic desulfurizing enzymes of *Paenibacillus* sp. A11-2.** *Biochem Biophys Res Commun* 2000, **270**:81-88.

•19. Kirimura K, Harada K, Iwasawa H, Tanaka T, Iwasaki Y, Furuya T, Ishii Y, Kino K: **Identification and functional analysis of the genes encoding dibenzothiophene-desulfurizing enzymes from thermophilic bacteria.** *Appl Microbiol Biotechnol* 2004, **65**:703-713.

The DNA sequences of the desulfurization genes encoded by the moderate thermophiles *B. subtilis* WU-S2B and *M. phlei* WU-F1 were determined and shown to be identical to each other; however, they share only 61% and 58% identity with the desulfurization genes of *R. erythropolis* IGTS8 and *Paenibacillus,* respectively.

•20. Abbad-Andaloussi S, Lagnel C, Warzywoda M, Monot F: **Multi-criteria comparison of resting cell activities of bacterial strains selected for biodesulfurization of petroleum compounds.** *Microbiol Technology* 2003, **32**:446-454.

Four out of five bacterial cultures isolated on the basis of their ability to utilize DBT as a sole source of sulfur were shown to have identical *dsz* gene sequences, yet they possessed different phenotypes as regards substrate range and desulfurization activity.

•21. Noda K, Watanabe K, Maruhashi K: **Isolation of a recombinant desulfurizing 4,6-diproply dibenzothiophene in n-tetradecane.** *J Biosci Bioeng* 2003, **95**:354-360.

The *dszABCD* genes were introduced on a transposon into 22 *Rhodococcus* or *Mycobacterium* strains. The substrate range and enzymatic activities of the cultures was found to vary considerably.

22. Konishi J, Okada H, Hirasawa K, Ishii Y, Maruhashi K: **Comparison of the substrate specificity of the two bacterial desulfurization systems.** *Biotechnol Lett* 2002, **24**:1863-1867.

23. Noda K, Watanabe K, Maruhashi K: **Recombinant *Pseudomonas putida* carrying both the *dsz* and *hcu* genes can desulfurize dibenzothiophene in *n*-tetradecane.** *Biotechnol Lett* 2003, **25**:1147-1150.

••24. Guobin S, Huaiying Z, Weiquan C, Jianmin X, Huizhou L: **Improvement of biodesulfurization rate by assembling nanosorbents on the surfaces of microbial cells.** *Biophys J* 2005, **89**:L58-L60.

The addition of Al_2O_3 nanoparticles to *Pseudomonas delafieldii* R-8 cells doubled the rate of DBT biodesulfurization.

••25. Park HS, Kayser KJ, Kwak JH, Kilbane JJ II: **Heterologous gene expression in *Thermus thermophilus*: β-galactosidase, dibenzothiophene monooxygenase, PNB carboxy esterase, 2-aminobiphenyl-2,3-diol dioxygenase, and chloramphenicol acetyl transferase.** *J Ind Microbiol Biotechnol* 2004, **31**:189-197.

Functional expression of the *dszC* gene from *R. erythropolis* IGTS8 and the *carBaBb* genes from *Sphingomonas* sp. GTIN11 was demonstrated in the thermophilic culture *T. thermophilus*. This work demonstrates that the development of high-temperature biorefining processes for the removal of sulfur and/or nitrogen from petroleum might be possible.

26. Nakamura A, Takakura Y, Kobayashi H, Hoshino T: ***In vivo* directed evolution for thermostabilization of *Escherichia* coli hygromycin B phosphotransferase and the use of the gene as a selection marker in the host-vector system of *Thermus thermophilus*.** *J Biosci Bioeng* 2005, **100**:158-163.

•27. Okada H, Nomura N, Nakahara T, Saitoh K, Uchiyama H, Maruhashi K: **Analyses of microbial desulfurization reaction of alkylated dibenzothiophenes dissolved in oil phase.** *Biotechnol Bioeng* 2003, **83**:489-497.

The specific desulfurization activity needed for a commercial biodesulfurization process is stated as >1200 μmol/g DCW/h. The authors examined the activity of *Mycobacterium* sp. G3 on DBT and various alkylated DBTs and found that the maximum activity was 49 μmol/g DCW/h – rates declined with increasing degrees of alkylation.

28. Kobayashi M, Onaka T, Ishii Y, Konishi J, Takaki M, Okada H, Ohta Y, Koizumi K, Suzuki M: **Desulfurization of alkylated forms of both dibenzothiophene and benzothiophene by a single bacterial strain.** *FEMS Microbiol Lett* 2000, **187**:123-126.

•29. Konishi M, Kishimoto M, Omasa T, Katakura Y, Shioya S, Ohtake H: **Effect of sulfur sources on specific desulfurization activity of *Rhodococcus erythropolis* KA2-5-1 in exponential fed-batch culture.** *J Biosci Bioeng* 2005, **99**:259-263.

Fed-batch cultivation of *R. erythropolis* KA2-5-1 engineered to contain multiple copies of the *dsz* genes resulted in a maximum desulfurization activity of 250 μmol 2-HBP/g DWC/h and a cell yield of 20 g DCW/L, but this level of activity is stated to be too low for a commercial biodesulfurization process.

30. Hirasawa K, Ishii Y, Kobayashi M, Koizumi K, Maruhashi K: **Improvement of desulfurization activity in *Rhodococcus erythropolis* KA2- 5-1 by genetic engineering.** *Biosci Biotechnol Biochem* 2001, **65**:239-246.

•31. van der Geize R, Dijkhuizen L: **Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications.** *Curr Opin Microbiol* 2004, **7**:255-261.

This article provides the website for the *Rhodococcus* sp. strain RHA1 genome sequencing project (http://www.bcgs.bc.ca/cgibin/rhodococcus/blast_rha1.pl) together with information about the metabolic diversity and range of biotechnology applications for rhodocci.

•32. Nakashima N, Tamura T: **A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C.** *Biotechnol Bioeng* 2004, **86**:136-148.

The construction and use of gene expression vectors for *R. erythropolis* and related species are described. Various heterologous genes that cannot be efficiently expressed in *E. coli* were found to express in *R. erythropolis*.

33. Franchi E, Rodriguez F, Serbolisca L, de Ferra F: **Vector development, isolation of new promoters and enhancement of the catalytic activity of the Dsz enzyme complex in Rhodococcus sp. strains.** *Oil Gas Science Technology Review IFP* 2003, **58**:515-520.

34. Reichmuth DS, Blanch HW, Keasling JD: **Dibenzothiophene biodesulfurization pathway improvement using diagnostic GFP fusions.** *Biotechnol Bioeng* 2004, **88**:94-99.

35. Takada M, Nomura N, Okada H, Nakajima-Kambe T, Nakahara T, Uchiyama H: **De-repression and comparison of oil-water separation activity of the dibenzothiophene desulfurizing bacterium, Mycobacterium sp. G3.** *Biotechnol Lett* 2005, **27**:871-874.

36. Noda K, Watanabe K, Maruhashi K: **Cloning of a rhodococcal promoter using a transposon for dibenzothiophene biodesulfurization.** *Biotechnol Lett* 2003, **25**:1875-1882.

37. Jimenez A, Santos MA, Pompejus M, Revuelta JL: **Metabolic engineering of the purine pathway for riboflavin production in ashbya gossypii.** *Appl Environ Microbiol* 2005, **71**:5743-5751.

38. Erwin KN, Nakano S, Zuber P: **Sulfate-dependent repression of genes that function in organosulfur metabolism in *Bacillus subtilis* requires Spx.** *J Bacteriol* 2005, **187**:4042-4049.

••39. Yu B, Xu P, Shi Q, Ma C: **Deep desulfurization of diesel oil and crude oils by a newly isolated *Rhodococcus erythropolis* strain.** *Appl Environ Microbiol* 2006, **72**:54-58.

R. erythropolis XP was shown to desulfurize hydrotreated diesel from 259 ppm sulfur to 14 ppm in 24 h, and crude oil was desulfurized from 3210 ppm sulfur to 1210 ppm sulfur in 72 h. The specific desulfurization rate for crude oil was equivalent to about 3 μmol/g DCW/h.

40. Li FL, Xu P, Ma CQ, Luo LL, Wang XS: **Deep desulfurization of a hydrodesulfurization-treated diesel oil by a facultative thermophilic mycobacterium sp. X7B.** *FEMS Microbiol Lett* 2003, **223**:301-307.

•41. Labana S, Pandey G, Jain RK: **Desulphurization of dibenzothiophene and diesel oils by bacteria.** *Lett Appl Microbiol* 2005, **40**:159-163.

Diesel oil containing 7600 ppm sulfur was hydrodesulfurized to yield a product that contained 170 ppm sulfur and this was further desulfurized to 80 ppm sulfur by a *Rhodococcus* sp. biocatalyst.

•42. Li F, Xu P, Feng J, Meng L, Zheng Y, Luo L, Ma C: **Microbial desulfurization of gasoline in a *Mycobacterium goodii* X7B immobilized-cell system.** *Appl Environ Microbiol* 2005, **71**:276-281.

The desulfurization of gasoline from 227 ppm to 71 ppm sulfur was achieved using *Mycobacterium goodii* X7B. The vast majority of biodesulfurization studies target diesel rather than gasoline, so this demonstration of gasoline desulfurization is important. Benzothiophenes are the most abundant organosulfur compounds in gasoline, whereas DBTs predominate in diesel. Gasoline is also considerably more toxic to bacteria than diesel.

••43. Nunn D: *The Biocatalytic Desulfurization Project: DOE contract FC26-02NT15340.* Diversa and Petro Star; 2006.

Diversa was working with Petro Star on a DOE-funded project to further develop the biodesulfurization technology acquired from Energy BioSystems/Enchira. The goal was to construct a biodesulfurization demonstration plant at Petrostar's Valdez Alaska refinery; however, insufficient biocatalyst activity resulted in the cancellation of plans for the construction of a demonstration plant. They considered chemical oxidation combined with biodesulfurization, and conversion of DBTs to sulfonates/sulfinates that could be recovered and used as surfactants.

44. Al-Otabi MB, Elkamel A, Nassehi V, Abdul-Wahab SA: **A computational intelligence based approach for the analysis and optimization of a crude oil desalting and dehydration process.** *Energy Fuels* 2005, **19**:2526-2534.

45. Shong RG: **Bioprocessing of crude oils**. *Div Fuel Chem Amer Chem Soc* 1999, **44**:1-4.

46. Kilbane JJ, Ranganathan R, Kayser KJ, Cleveland L, Ribiero C, Linhares MM: **Selective removal of nitrogen from quinoline and petroleum by *Pseudomonas ayucida* IGTN9m.** *Appl Environ Microbiol* 2000, **66**:688-693.

47. Frerichs-Deeken U, Goldenstedt B, Gahl-Janben R, Kappl R, Huttermann J, Fretzner S: **Functional expression of the quinoline 2-oxidoreductase genes (*qorMSL*) in *Pseudomonas putida* KT2440 pUF1 and in *Pseudomonas putida* 86-1 (Δ) qor pUF1 and analysis of the Qor proteins.** *Eur J Biochem* 2003, **270**:1567-1577.

48. Kilbane JJ II, Daram A, Abbasian J, Kayser KJ: **Isolation and characterization of carbazole-degrading bacterium *Sphingomonas* sp. GTIN11.** *Biochem Biophys Res Commun* 2002, **297**:242-248.

49. Bressler DC, Gray MR: **Hydrotreating chemistry of model products from bioprocessing of carbazoles.** *Energy Fuels* 2002, **16**:1076-1086.

•50. Santos SC, Alviano DS, Alviano CS, Padula M, Leitao AC, Martins OB, Ribeiro CM, Sassaki MY, Matta CP, Bevilaqua J *et al.*: **Characterization of *Gordonia* sp. strain F.5.25.8 capable of dibenzothiophene desulfurization and carbazole utilization.** *Appl Microbiol Biotechnol* 2005, 1-8. DOI: 10.1007/s00253-005-0154-z.

Gordonia sp. strain F.5.25.8 is the first microbial culture to be reported that has the ability to both desulfurize DBT and to degrade carbazole.

Such strains could potentially be used for the simultaneous removal of both sulfur and nitrogen from petroleum.

•51. Garcia-Arellano H, Buenrostro-Gonzalez E, Vazquez-Duhalt R: **Biocatalytic transformation of petroporphyrins by chemical modified cytochrome *c*.** *Biotechnol Bioeng* 2004, **85**:790-798.

Cytochrome *c* was demonstrated to be capable of releasing vanadium and nickel complexed by porphyrin compounds and, therefore, could be useful in a bioprocess for the removal of metals from petroleum.

52. Luo MF, Xing JM, Gou ZX, Li S, Liu HZ, Chen JV: **Desulfurization of dibenzothiophene by lyopholized cells of *Pseudomonas delafieldii* R-8 in the presence of dodecane.** *Biochem Eng J* 2003, **13**:1-6.

53. Guobin S, Xing J, Zhang H, Liu H: **Deep desulfurization of hydrotreated diesel oil by *Pseudomonas delafieldii* R-8.** *J Chem Technol Biotechnol* 2005, **80**:420-424.

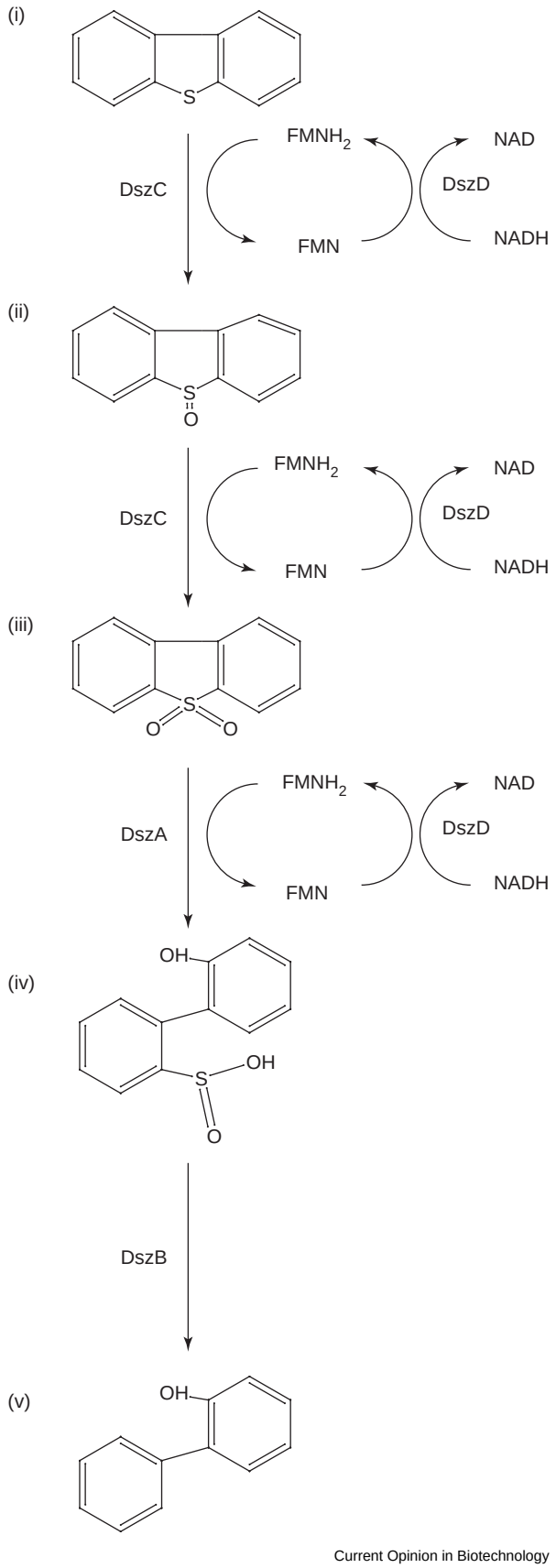


Figure 1

The oxidative metabolic pathway for DBT desulfurization. DszC, DBT monooxygenase; DszA, DBT sulfone monooxygenase; DszB, HPBSi desulfinase; DszD, flavin reductase. I, DBT; II, DBT sulfoxide; III, DBT sulfone; IV, hydroxyphenylbenzenesulfinate; V, 2-hydroxybiphenyl.

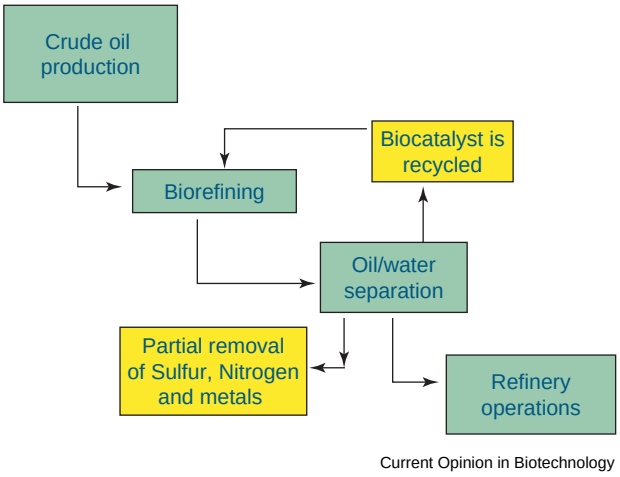


Figure 2

As petroleum production includes desalting and dewatering steps, then performing biodesulfurization/biorefining prior to, or in combination with, these oil/water separation steps would allow minimal alterations to existing industry practices. A biorefining process could potentially reduce the sulfur, nitrogen, and metal content of crude oil prior to being sent to a refinery.

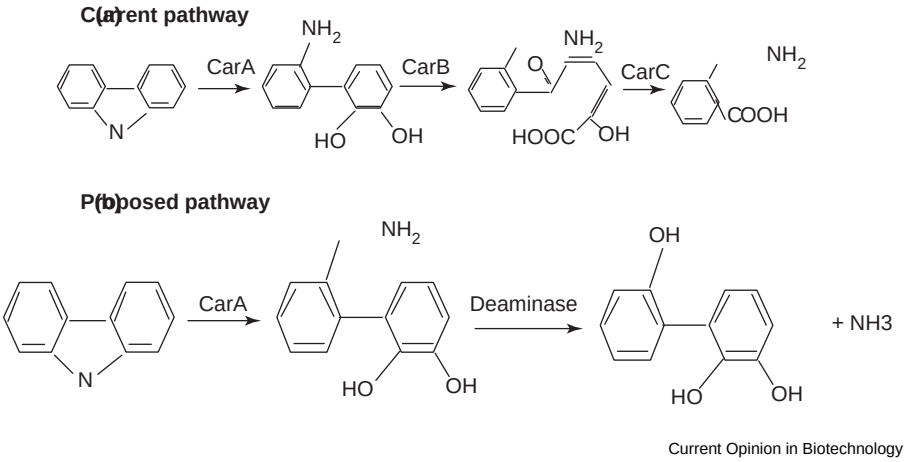


Figure 3

Carbazole degradation pathways. **(a)** The existing pathway in which carbazole is initially converted to 2'-aminobiphenyl-2,3-diol and ultimately results in overall degradation. **(b)** A potential pathway for the selective removal of nitrogen from carbazole that could be developed using metabolic engineering.

Table 1

Specific activity of various biodesulfurization cultures.

Microbial culture	Specific activity µmol/g DCW/h	References
Rhodococcus erythropolis KA 2-5-1, wild type	74	[28]
Rhodococcus erythropolis KA 2-5-1, wild type	50	[30]
Rhodococcus erythropolis KA 2-5-1, cloned *dsz* genes	280	[30]
Rhodococcus erythropolis KA 2-5-1, cloned *dsz* genes	250	[29•]
Rhodococcus erythropolis IGTS8	72	[4•]
Rhodococcus erythropolis XP	4	[39••]
Mycobacterium sp. GB	49	[27•]
Mycobacterium sp. X7B	4	[40]
Mycobacterium sp. G3, wild type	178	[35]
Mycobacterium sp. G3, *hsp60* promoter plus sulfate	35	[35]
Mycobacterium sp. G3, *hsp60* promoter	211	[35]
Pseudomonas delafieldii R-8	11	[52]
Pseudomonas delafieldii R-8, wild type	16	[24••]
Pseudomonas delafieldii R-8, with Al_2O_3	40	[24••]

Pseudomonas delafieldii R-8	13	[53]

The desulfurization activities of several cultures described in the literature are expressed in terms of μmol/g DCW/h so that they can be readily compared, however, not all of the studies were performed under equivalent conditions. Some studies employed model systems comprised of DBT dissolved in a solvent/oil phase, whereas others employed actual petroleum products. In some instances the concentration of desulfurized product, 2-HBP, was quantified and used to calculate desulfurization rates, whereas in other cases the concentration of sulfur was determined and it was necessary to convert milligrams of sulfur or some other unit into μmol/g DCW/h. When such conversions were necessary it was assumed that the molecule being desulfurized was DBT with a molecular weight of 184, even when petroleum samples containing complex mixtures of organosulfur compounds were present. All specific activity values listed were rounded up to the nearest integer.

Hazardous Substance Report: DE-FC26-02NT15382 "Metabolic Engineering to Develop a Pathway for the Selective Cleavage of Carbon-Nitrogen Bonds"
April 28, 2006

Description of Substance/Chemical

The hazardous chemicals used in this project were:

Pyridine.	CAS# 110-86-1
Aniline,	CAS# 62-53-3
Ethyl acetate	CAS# 141-78-6

EPA Hazardous Waste Number

Pyridine.	U196
Aniline,	K083
Ethyl acetate	F003

EPA Hazard Code

Pyridine.	T (toxic)
Aniline,	T (toxic)
Ethyl acetate	I (ignitable)

Quantity Utilized or Generated

Approximately 2 grams of pyridine, 5 grams of aniline, and 2 liters of ethyl acetate was utilized in microbiological and biochemical experiments during this project.

Hazardous Waste Transporter

The portion of chemicals utilized in this project that was not consumed/degraded in the course of the project and resulted in hazardous waste requiring disposal was transported by:

Hazchem Environmental Corporation
1115 National Drive
Addison, IL 60101
EPA # ILD 981 002 074

Hazardous Waste Disposal Facility Contractor and Location

Environmental Services of America, Inc.
604 S. Scott Street
South Bend, IN 46601
EPA # IND 980 590 947

Treatment Method

Thermal oxidation