This is to certify that the

thesis entitled

The Initial Steps in 2,4-D Degradation in Soil Bacteria

presented by

Timothy Martin Sassanella

has been accepted towards fulfillment
of the requirements for
M.S. degree in **Microbiology**

Major professor

Date 5/13/97

THE INITIAL STEPS OF 2,4-DICHLOROPHENOXYACETIC ACID (2,4-D) DEGRADATION IN SOIL BACTERIAL ISOLATES

By

Timothy Martin Sassanella

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Microbiology

1998

ABSTRACT

THE INITIAL STEPS OF 2,4-DICHLOROPHENOXYACETIC ACID (2,4-D) DEGRADATION IN SOIL BACTERIAL ISOLATES

By

Timothy M. Sassanella

The initial step of the canonical 2,4-D degradation pathway is carried out by 2,4-D/alpha-ketoglutarate dioxygenase (TfdA), a member of a diverse superfamily of mechanistically related enzymes. Phylogenetic analysis of these enzymes indicated that this superfamily is probably polyphyletic. A continuous, quantitative assay using 4-nitrophenoxyacetic acid (4-NPAA) to detect TfdA-like activities was developed for the screening of intact bacterial cells and cell lysates. A survey using this assay indicated substantial diversity among a diverse collection of 2,4-D degrading soil bacteria. *Nitrobacter winogradskyi* M1 was examined further, and was found to carry a plasmid borne atypical TfdA-like activity. Differential response to the 4-NPAA assay of several engineered constructs indicated that the permeability of 4-NPAA varied among soil bacteria, and

was confirmed using ^{14}C-2,4-D uptake and incorporation assays. Transposon mutagenesis of the 2,4-D degradation plasmid pJP4 indicated that the plasmid encodes an unknown factor that influences the permeation of 2,4-D in some strains.

To my parents and family, whose unflagging support made me the person I am today, to Brenda Knotts, for being the true friend that she is, and to my best friend and partner - Dean.

ACKNOWLEDGMENTS

A thesis is never generated in isolation, so I would like to extend my thanks to a few people. Thanks to all of the graduate students that I have spent hours speaking with about this and that, especially those of the incoming class of 1991 - all of whom were more than just colleagues. Thanks to my friends and associates that participated in the ROME lab, where a great deal of intriguing and thoughtful work was done that inspired my personal work. Thanks to Dr. Tiedje and the NSF Center for Microbial Ecology for many years of financial and institutional support. I would like to thank my committee members: Dr. Pat Oriel, Dr. Robert Hausinger, Dr. Mike Thomashow, and Dr. Frans DeBruijn for guiding my work.

Finally, I would like to thank my mentor Dr. Micheal Bagdasarian, who has many qualities that I may only aspire to have, for his time, patience, and wisdom.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Microorganisms play a significant part in the removal and detoxification of xenobiotic compounds in nature (Alexander,1981; Chaudhry,1991; Ghosal, 1985; Häggblom, 1992; van der Meer, 1992). Degradation of these compounds, which include halogenated aromatics, is often slow due to limited degradation by unfavorable physiochemical conditions (Evans, 1959; World Health Organization, 1984 and 1989) and the inability of most microorganisms to metabolize them due to unusual chemical structure. Microbial communities have been shown to adapt, developing the ability to utilize many halogenated aromatic compounds upon exposure to xenobiotic compounds (Pemberton, 1981). This presumably occurs by induction of appropriate members of the community, shifts in the population that increase the number of degrading bacteria present in the community (Spain,1983), or by mutation, gene transfer, or other mechanisms of bacterial evolution that result in a state where degradation is possible. In this last case, a novel catabolic function is created from a specific 'parental' enzyme or enzymes. Genes expressing these enzymes would be expected to continue to evolve rapidly under selection of the xenobiotic, particularly if the compound can serve as a source of carbon and energy.

The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D), a synthetic auxin, is one of the least chemically complex members of the chlorinated aromatic family that includes PCP, PCBs, and dioxins. 2,4-D has been in wide commercial and agricultural use for approximately forty years, with 40 million pounds being applied to 54.8 million acres of U.S. agricultural land during 1971 alone (Bovey, 1980). Unexpectedly, this compound, though halogenated and moderately toxic, has not been found to accumulate in soil or water (Perkins, 1988). Individual environmental isolates as well as microbial communities from around the world have been found to mineralize 2,4-D readily (Friedrich, 1983; Fulthorpe 1995 and 1996; Pemberton, 1981; Sinton, 1986; Spain, 1983). The diverse geographic and temporal application of 2,4-D to soils, combined with the widespread degradability of this xenobiotic, should provide a maximal diversity of microorganisms and catabolic systems in which the evolution of novel catabolic genes and pathways can be studied.

The canonical pathway for the complete mineralization of 2,4-D has been well characterized (Don, 1985; Fukumori, 1993a and 1993b; Harker, 1989; Kaphammer, 1990a and 1990b; Kasberg, 1995; Perkins, 1988 and 1990; Pieper, 1988; Streiber, 1987; You, 1995). The genes are encoded on the 90 kb, IncP, broad-host-range plasmid pJP4 originally isolated from *Alcaligenes eutrophus* (Don, 1985). There are six known structural and two

identical regulatory genes responsible for the conversion of 2,4-D to b-ketoadipate (Don, 1985; Kukor, 1989; Kasberg, 1995). The initial three steps of the pathway are performed by three oxygenases: TfdA, an α-ketoglutarate-dependent dioxygenase that converts 2,4-D to 2,4-DCP (Fukumori, 1993a and 1993b); TfdB, a dichlorophenol monooxygenase that converts 2,4-DCP to 3,5-dichlorocatechol (Perkins, 1988); and TfdC, an intradiol chlorocatechol dioxygenase that begins a typical modified *ortho*-cleavage pathway through 2-chloromaleylacetate to TCA cycle compounds (Perkins, 1990; Figure A). The pathway genes are clustered in a 20 kb region of pJP4. The functional and regulatory genes are found in three clusters that are separated by regions that contain non-functional partial duplications of pathway genes (Figure B). The overall genetic structure of the pathway, including the abundance of partial fragments and the presence of insertion sequences, suggests that this pathway was assembled from diverse genetic elements. The genes *tfdA* and that encoding 2,4-dichlorophenol hydroxylase (TfdB) seem to encode the most atypical enzymes of the group, and appear to have been recruited to expand the chlorocatechol degradation pathway (Harker, 1989; Kaphammer, 1990a and 1990b; Perkins, 1990). An alternate 2,4-D pathway has also been found that retains a TfdA-like side chain cleavage function as in the canonical pathway, but where the 2,4-dichlorophenol is degraded in a similar manner as

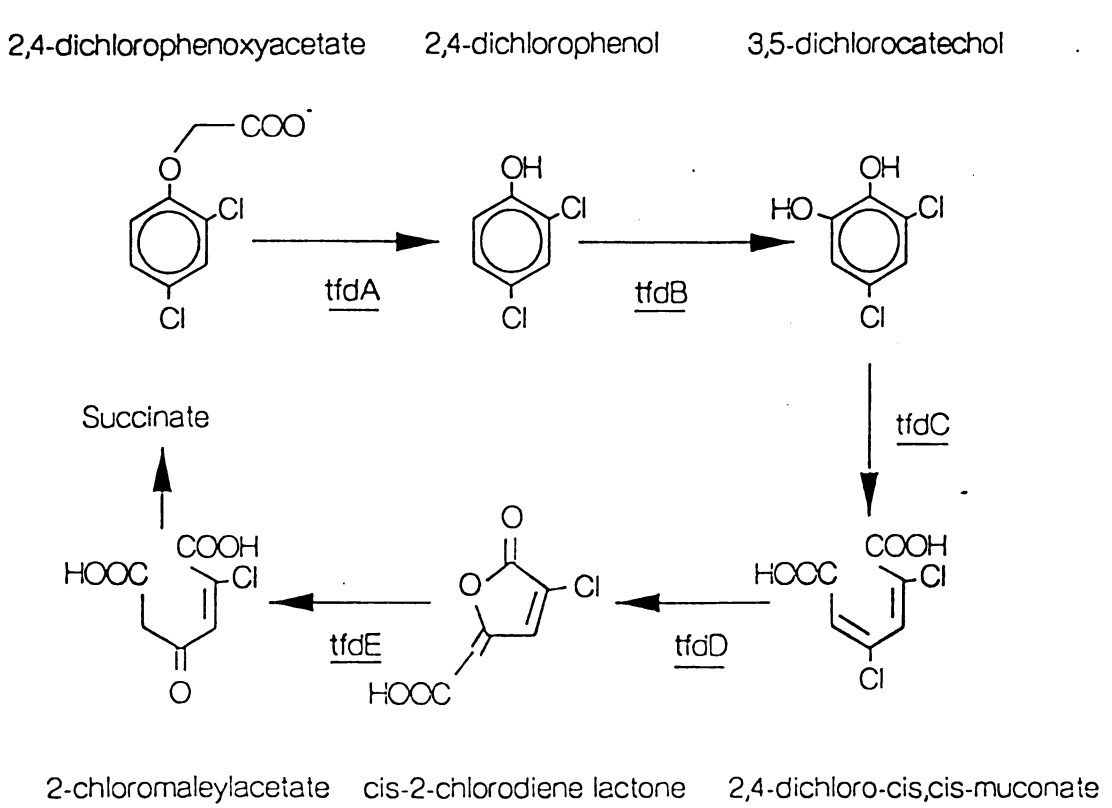


Figure A. The canonical pathway for 2,4-D degradation as occurs in *Alcaligenes eutrophus* JMP134.

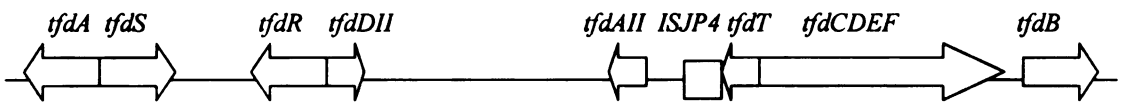


Figure B. Genetic structure of the canonical 2,4-D degradation pathway of plasmid pJP4 isolated from *Alcaligenes eutrophus* JMP134. The known 2,4-D degradation genes are located in an approximately 20 kb region of the 90 kb plasmid. *TfdAII* and *TfdDII* are highly similar non-functioning truncated copies of the primary pathway gene. *TfdT* is a non-functional partial copy of a LysR regulatory gene that is closely related to the identical pathway regulators *tfdR* and *tfdS*. Compiled from Don, 1985, Harker, 1989, and Levaeu, 1996.

pentachlorophenol (O. Maltseva, personal communication). An analogous compound, 2,4,5-trichlorophenoxyacetic acid (2,4,5-T), is also degraded by the pentachlorophenol pathway. The chlorophenol is produced by a NAPH-dependent monooxygenase (TftA) that performs an analogous reaction to that of TfdA (Haughland, 1991, Xun,1995) and will cleave both 2,4-D and 2,4,5-T, though at a markedly reduced rate with the dichlorinated substrate (Fukumori, 1993b; Haughland, 1991). Though they share a similar biochemical function, TfdA and TftA are evolutionarily distinct.

Various surveys of 2,4-D degrading isolates indicate that there are many plasmids that closely resemble the pathway encoded on pJP4 among Gram-negative soil bacteria that can use 2,4-D as a sole carbon and energy source (Tonso, 1996; Top, 1996). Hybridization studies indicate that 2,4-D degradation pathways utilizing different but closely related (>60% identity) genes have been assembled independently in different organisms (Fulthorpe, 1995). This mosaic pattern for the assembly of catabolic pathways is consistent with the sequence information for pJP4. These experiments also indicate that there is significant diversity among these bacteria, with a number of strains failing to hybridize to probes for *tfdA, tfdB,* or *tfdC*. Many of the strains tested, including *Alcaligenes eutrophus* JMP134, were β-proteobacterial isolates, and those strains that are phylogenetically more

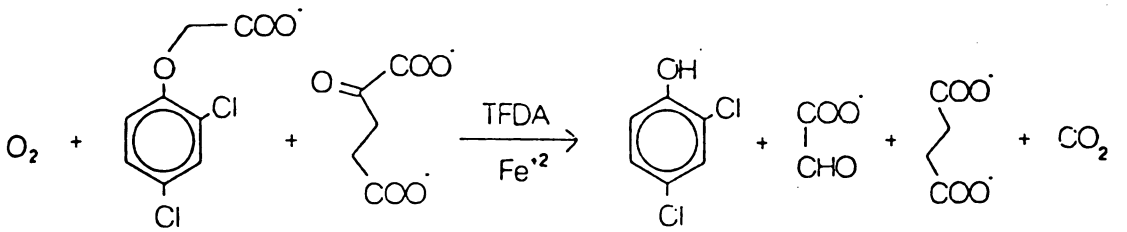
distant rarely hybridized with the canonical genes (McGowan, 1995). Further, plasmid capture experiments using a *tfdA*⁻ deletion mutant of pJP4 indicated that there is a wide variety of plasmids in soil that can complement the lack of TfdA activity in the *A. eutrophus* construct (Top, 1995). A number of these plasmids carry only a TfdA-like activity (Top, 1996), suggesting that there may be significant genetic diversity found among incomplete 2,4-D pathways.

The utilization of a-ketoglutarate places TfdA into the highly diverse superfamily of α-ketoglutarate-dependent dioxygenases, a group of oxygenases that are thought to share a common mechanism of catalysis (Prescott, 1993). A number of sequences of enzymes with this requirement from both eucaryotic and procaryotic organisms are known, and the phylogenetic relationships within these sequences and among other enzymes of similar function are unclear. Members of this group are as diverse as eucaryotic prolyl-4-hydroxylase (Helaakoski, 1989), bacterial gamma-butyrobetaine hydroxylase (Englard, 1985), and fungal isopenicillin N synthase (Aharonowitz, 1992). TfdA, several slightly divergent variants, and a putative α-ketoglutarate-dependent dioxygenase *tauD* (Leisinger, 1997) show conservation of a number of residues, particularly histidine residues which have been implicated as essential for activity (Fukumori, 1993b). Though clearly related to each other, these enzymes exhibit distant

similarity to another cluster of enzymes that are gamma-butyrobetaine hydroxylases, but not with other members of the superfamily (Prescott, 1993; this work, Chapter 1).

TfdA converts 2,4-D, O_2, and α-ketoglutarate to 2,4-dichlorophenol, glyoxylate, carbon dioxide, and succinate (Fukumori, 1993b; Figure C). This enzyme is a non-heme iron dependent dioxygenase has utilizes reducing agent, usually ascorbate, that is thought to help maintain the metallocenter in its active state. This enzyme has the highest affinities for 2,4-D and α-ketoglutarate among a series of phenoxyacids and α-ketoacids examined (Fukumori, 1993a and 1993b), but may utilize other similar compounds to a lesser degree (Figure D). This property was exploited in the development of a chromogenic assay for TfdA-like activity that has significant advantages to previous methods, and allows rapid, large scale screening of bacteria (Chapter 2).

Two virtually identical *tfdA*-like genes have recently been cloned from *Burkholderia* species, both of which have greater than 75% homology to *tfdA* and 91% identity at the amino acid level (Suwa, 1996; Matheson, 1996). The product of $tfdA_{rasc}$ has been partially characterized biochemically, and it functions in a similar fashion to the benchmark enzyme. The three-dimensional structure of these enzymes is unknown. Recent genome sequence



2,4-D α-keto glutarate 2,4-DCP glyoxylate succinate

Figure C. Reaction mediated by 2,4-D/α-ketoglutarate dioxygenase (TfdA). This diagram is from Fukumori, 1993b.

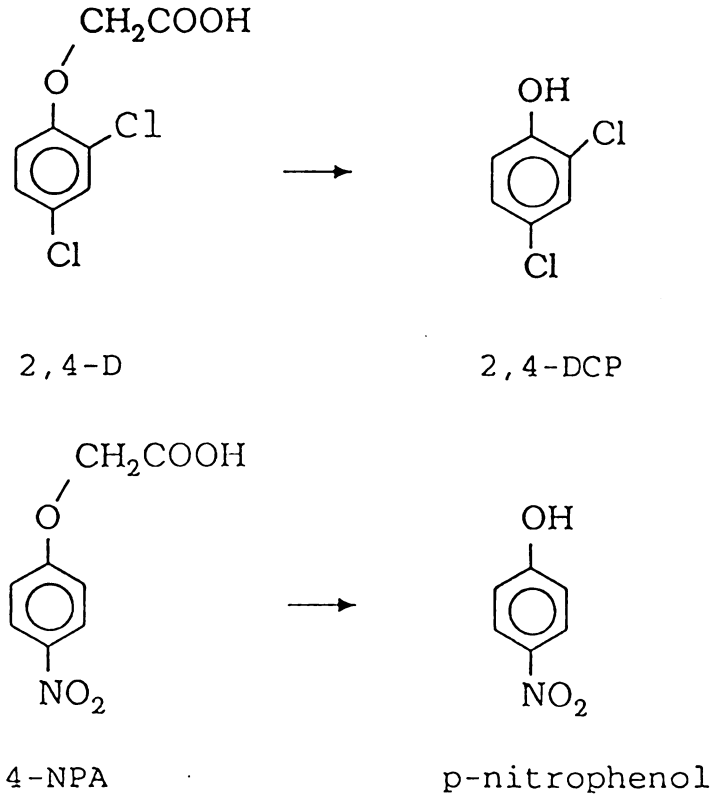


Figure D. Two of the known reactions catalyzed by 2,4-D/α-ketoglutarate dioxygenase (TfdA). 2,4-dichlorophenoxyacetic acid (2,4-D) is converted to 2,4-dichlorophenol (2,4-DCP) preferentially by TfdA. The conversion of 4-nitrophenoxyacetic acid produces 4-nitrophenol, an intensely yellow compound. This chromogenic compound can not be utilized by the rest of the 2,4-D pathway and accumulates, a property which was exploited for the development of an assay for TfdA (Chapter 2).

projects have produced several complete sequences that are similar enough to *tfdA* to infer relatedness. Three sequences - *scox1* from *Saccharomyces cereviseae*, and *mtox1* and *mtox2* from *Mycobacterium tuberculosis* - share more than 27% identity with *tfdA*, but have not been biochemically characterized. The gene product of *tauD* from *E. coli* is taurine dioxygenase, and shares 31% identity with *tfdA*. It is part of a sulfur scavenging pathway, and has yet to be well characterized (Leisinger, 1997). Since the phylogenetically related *tauD* gene product can catalyze alternate hydroxylation reactions, yet the unrelated *tmo* gene product will hydroxylate 2,4-D at C-2 in a TfdA-like reaction, then is not unfeasible that there are *tfdA*-like genes that are less homologous than have been examined, and there may be other phylogenetically distinct TfdA-like enzymes in soil bacteria. More than 50 Gram-negative bacterial 2,4-D degrading strains were examined for alternate activities, and the identification of two strains expressing biochemically distinct TfdA-like enzymes is covered in Chapter 3.

Though the characterization of the canonical 2,4-D pathway has been extensive, another degradation factor has been overlooked. Since bacterial cytoplasmic membranes are relatively impermeable to most solutes, permeases are required for the transport of most substances used for catabolic metabolism (Nikaido, 1985). This transport step would be a crucial initial step

in the biodegradation of xenobiotic compounds. The transport of 2,4-D into bacterial cells was found vary significantly among bacterial strains, and the expression of a putative permease encoded on the plasmid pJP4 is critical factor in the degradation of 2,4-D in some environmental strains. The transport of 2,4-D into some bacterial cells is examined in Chapter 4.

REFERENCES

Aharonowitz, Y., Cohen, G., and Martin, J.F. 1992. Penicillin and cephalosporin biosynthetic genes: structure, organization, regulation, and evolution. Ann. Rev. Microbiol. **46**, 461-95.

Alexander, M. 1981. Biodegradation of chemicals of environmental concern. Science **211**, 132-8.

Bovey, R.W. and A. L. Young. 1980. The Science of 2,4,5-T and Associated Phenoxy Herbicides. John Wiley & Sons, New York.

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

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

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

Englard, S., Balnchard, J.S., and C. F. Middlefort. 1985. Gamma-

butyrobetaine hydroxylase: stereochemical course of the hydroxylation reaction. Biochemistry. **24**: 1110-1116.

Evans, W.C. and B.S.W. Smith. 1959. The photochemical inactivation and microbial metabolism of the chlorophenoxyacetic acid herbicides. Proc. Biochem. Soc. **26**, p. 126.

Friedrich, B., Meyer, M., and H. G. Schlegel. 1983. Transfer and expression of the herbicide-degrading plasmid pJP4 in aerobic autotrophic bacteria. Arch. Microbiol. **134**, 92-7.

Fukumori, F., and R.P. Hausinger. 1993a. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetic acid monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175**:2083-2086.

Fukumori, F., and R. P. Hausinger. 1993b. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J. Biol. Chem. **268**:24311-24317.

Fulthorpe, R.R., C. McGowan, O.V. Maltseva, W.E. Holben, and J.M. Tiedje. 1995. 2,4-dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. **61**: 3274-3281.

Fulthorpe, R.R., A. N. Rhodes, and J. M. Tiedje. 1996. Pristine soils

mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. Appl. Environ. Microbiol. **62**:1159-1166.

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

Harker, A.R., Olsen, R.H., and R. J. Seidler. 1989. Phenoxyacetic acid degradation by the 2,4-dichlorophenoxyacetic acid (TFD) pathway of plasmid pJP4: mapping and characterization of the TFD regulatory gene, *tfdR*. J. Bacteriol. **171**, 314-20.

Häggblom, M. M. 1992. Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. **103**: 29-72.

Haughland, R.A., Sangodkar, U.M.X., Sferra, P.R., and A. M. Chakrabarty. 1991. Cloning and characterization of a chromosomal DNA region required for growth on 2,4,5-T by *Peeudomonas cepacia* AC1100. Gene **100**, 65-73.

Helaakoski, T., Vuori, K., Myllylä, R., Kivirikko, K. I., and T. Pihlajaniemi. 1989. Molecular cloning of the alpha-subunit of human prolyl 4-hydroxylase: the complete cDNA-derived amino acid sequence and

evidence for alternative splicing. Proc. Natl. Acad. Sci. USA. **86**: 4392-4396.

Ka, J.O., W.E. Holben, and J.M. Tiedje. 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. Appl. Environ. Microbiol. **60**:1106-1115.

Kaphammer, B., Kukor, J.J., R. H. Olsen. 1990a. Regulation of *tfdCDEF* by *tfdR* of the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. **172**, 2280-6.

Kaphammer, B., Kukor, J.J., R. H. Olsen. 1990b. Cloning and characterization of *tfdS*, the repressor/activator of *tfdB*, from the 2,4-dichlorophenoxyacetic acid degradation plasmid pJP4. J. Bacteriol. **172**, 5856-62.

Kasberg, T., Daubaras, D. L., Chakrabarty, A. M., Kinzelt, D., and W. Reineke. 1995. Evidence the operons *tcb*, *tfd*, and *clc* encode maleylacetate reductase, the fourth enzyme of the modified *ortho* pathway. J. Bacteriol. **177(13)**: 3885-3889.

Kukor, J.J., Olsen, R.H., and J. S. Siak. 1989. Recruitment of a chromosomally encoded maleylacetate reductase for degradation of

2,4-dichlorophenoxyacetic acid by plasmid pJP4. J. Bacteriol. **171**, 3385-90.

Levaeu, J. H. J., and J. R. van der Meer. 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134 (pJP4). J. Bacteriol. **178(23)**: 6824-6832.

Matheson, V. G., Forney, L. J., Suwa, Y., Nakatsu, C. H., Sexstone, A. J., and W. E. Holben. 1996. Evidence for acquisition in nature of a chromosomal 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase gene by different *Burkholderia* spp. Appl. Environ. Microbiol. **62(7)**: 2457-2463.

Nikaido, H. and M. Vaara. 1985. Molecular basis of bacterial outer membrane permeability. Microbiol. Rev. **49**: 1-32.

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

Perkins, E.J., M.P. Gordon, O. Caceres, and P.F. Lurquin. 1990.

Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. **172**:2351-2359.

Pemberton, J.M. 1981. Genetic engineering and biological detoxification of environmental pollutants. Residue Rev. **78**, 1-11.

Pieper, D.H., Reineke, W., Engesser, K.H., and H. J. Knackmuss. 1988. Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid, and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP134. Arch. Microbiol. **150**: 95-102.

Prescott, A. G. 1993. A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). J. Exp. Bot. **44**:849-861.

Sinton, G. L., L. T. Fan, L. E. Erickson, and S. M. Lee. 1986. Biodegradation of 2,4-D and related xenobiotic compounds. Enzyme Microbiol. Technol. **8**:395-403.

Spain, J.C. and P. A. van Veld. 1983. Adaption of natural microbial communities to degradation of xenobiotic compunds: effects of concentration, exposure, time, inoculum, and chemical structure. Appl. Env. Microbiol. **45**, 428-35.

Streber, W.R., K.N. Timmis, and M.H. Zenk. 1987. Analysis, cloning, and

high level expression of 2,4-dichlorophenoxyacetic acid monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169**:2950-2955.

Suwa, Y., A. D. Wright, F. Fukumori, K. A. Nummy, R. P. Hausinger, W. E. Holben, and L. J. Forney. 1996. Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetate (2,4-D)/ α-ketoglutarate dioxygenase. Appl. Environ. Microbiol. **62**: 2464-2469.

Tonso, N.L., V.G. Matheson, and W.E. Holben. 1995. Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microbial Ecol. **30**:3-24.

Top, E. M., Holben, W. E., and L. J. Forney. 1995. Characterization of diverse 2,4-D degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. **61**: 1691-1698.

Top, E. M., O. M. Maltseva, and L. J. Forney. 1996. Capture of a catabolic plasmid that encodes only 2,4-dichlorophenoxyacetic acid/α - ketoglutarate dioxygenase (TfdA) by genetic complementation. Appl. Environ. Microbiol. **62(7)**: 2470-2476.

van der Meer, J.R., de Vos, W.M., Harayama, S., A.J.B. Zehnder. 1992. Microbiol. Rev. **56**: 677-94.

World Health Organization. 1984. 2,4-dichlorophenoxyacetic acid (2,4-D). World Health Organization, Geneva, Switzerland.

World Health Organization. 1989. 2,4-dichlorophenoxyacetic acid (2,4-D) - environmental aspects. World Health Organization, Geneva, Switzerland.

You, I. S., and D. Ghosal. 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetic acid catabolic plasmid pJP4. Molec. Microbiol. **16**:321-331.

Xun, L., and K. B. Wagnon. 1995. Purification and properties of component B of 2,4,5-trichlorophenoxyacetic acid oxygenase from *Pseudomonas cepacia* AC1100. Appl. Env. Microbiol. **61**: 3499-3502.

Chapter 1:

The Phylogenetic Analysis of the Alpha-ketoglutarate-dependent Dioxygenase Superfamily

ABSTRACT

Genetically, the α-ketoglutarate-dependent dioxygenase superfamily is an extremely diverse group of enzymes that are thought to share a similar biochemical mechanism. Using protein sequences of the known members of the superfamily it was possible to perform a series of analyses that identified clusters of related enzymes using current phylogenetic techniques. This superfamily contains at least five distinct groups of enzymes that retain a number of conserved residues and regions of high similarity. Since the sequences of the distal members of the groups can share as little as 15% identity, this analysis provides a clearer picture of the evolutionary relationships of this superfamily than is revealed by general identity or similarity analysis.

INTRODUCTION

The α-ketoglutarate-dependent dioxygenase superfamily is formed of a highly diverse group of non-heme iron-(II) enzymes most of which are thought to share a common or highly similar biochemical mechanism of catalysis (Prescott,1993). These enzymes typically incorporate oxygen into organic substrates, producing intermediates that may retain the oxygen atoms or rearrange to produce one or more products. This mechanism is initiated by the binding of ferrous iron, oxygen, a substrate, and, in most cases, α-ketoglutarate. In the case of many of these enzymes including TfdA, oxidative decarboxylation of the α-ketoglutarate occurs to yield carbon dioxide, succinate, and a reactive iron-oxygen species. The insertion of oxygen into the other substrate the occurs (reviewed in Prescott, 1993). Members of this superfamily have been cloned from bacteria, fungi, plants, and animals. They exhibit moderate to very low sequence identity at the DNA and protein sequence level, raising the possibility that this group is polyphyletic. A clarification of the relatedness of the members of this group could be useful in further biochemical and evolutionary investigations.

Multiple sequence alignment of divergent sequences allows the detection of conserved residues and regions of proteins. These regions often

adopt a similar three-dimensional conformation, presenting specific residues for structural considerations or for the binding of substrates and co-factors. All of the enzymes in this superfamily are non-heme, iron metalloproteins (Prescott, 1993). Typical iron ligands are charged at physiological pH, including histidine, lysine, aspartic acid, glutamic acid, and cysteine, though the tyrosine residue is uncharged. Conservation of these residues may indicate that these amino acids are important for iron binding. Further, these enzymes bind at least one other substrate, and may require other conserved residues to maintain steric constraints and provide polar and hydrophobic interactions essential for catalysis. Histidine residues have been implicated as iron ligands in 2,4-D/α-ketoglutarate-dependent dioxygenase (Fukumori, 1993b, Hausinger, 1997), flavanone 3β-hydroxylase (Britsch, 1993), isopenicillin N synthase (Borovok, 1996, Jiang, 1991, Ming, 1991, Randall, 1993, Sim, 1994, Tan, 1996), and lysyl hydroxylase and prolyl hydroxylase (Myllylä, 1992), cysteine residues in clavamate synthase (Marsh, 1992), and tyrosine residues in 4-hydroxyphenylpyruvate dioxygenase (Lindstedt, 1992). The only enzyme of the superfamily whose crystal structure has been determined, isopenicillin N synthase, has two histidyl and one asparaginyl iron ligands (Fujishima,1994, Roach, 1995). Substrate binding studies have

been done for for several of the proteins, and positively charged residues essential for α-ketoacid binding have been suggested (Fukumori,1993b, Majammaa, 1985, Pascal, 1985, Ng, 1991).

After initial sequence searches, it was noted that some α-ketoglutarate-dependent dioxygenases show significant protein sequence identity to another group of non-heme iron (II) dioxygenases which bind ascorbate but not α-ketoglutarate (McGarvey, 1992). It was also noted that at least some a-ketoglutarate-dependent enzymes bind both substrates (Fukumori, 1993, Prescott, 1993). In reactions where both compounds may be bound, ascorbate is thought to reduce the oxidized iron metallocenter to the ferrous state after the normal dioxygenation reaction in the case of the ascorbate-dependent group, or after the uncoupled decarboxylation of a-ketoglutarate in the case of the α-ketoglutarate-dependent enzymes (Myllylä, 1984). It has been proposed that the binding sites of these co-substrates overlap (Myllylä, 1984). The mechanism of the ascorbate-dependent dioxygenases is proposed to be similar to that of the α-ketoglutarate-dependent dioxygenases, forming a logical bridge for the inclusion of these enzymes in any biochemical analysis of the superfamily.

In previous biochemical studies of a few members this superfamily of enzymes, there has been some indication of the existence of conserved histidines and possibly binding motifs, that may be involved in iron and/or substrate binding (Borovok,1996, Bradley,1986, Fujishima,1994, Jiang, 1991, Lindsteadt, 1982, Ming, 1991, Myllylä, 1992, Ng, 1991, Randall, 1993, Roach, 1995, Ruetschi, 1993, Tan, 1996, Tiow-Sian, 1994). Since there has been a recent proliferation in uncharacterized sequences due to genome sequencing projects, the structural or phylogenetic relationships of these enzymes to other enzymes that may be isofunctional or may have homologous or analogous regions or domains could be useful. By examining the members of the superfamily, it may be possible to identify shared residues or binding motifs for iron, the α-keto acid used as a cosubstrate, or one or more of the variable regions responsible for binding the second substrates. By examining the sequences using phylogenetic techniques, the superfamily can be assessed for internal clusters of related sequences and compared directly for the identification of universally conserved residues or highly conserved regions. This information could then be used to inform further biochemical analyses.

RESULTS AND DISCUSSION

Since the α-ketoglutarate/ascorbate-dependent dioxygenase superfamily is so diverse in both genetic and protein sequence, direct comparison of the sequences is fairly ineffective in identifying potential evolutionary relationships. A detailed series of sequence alignments and phylogenetic analyses was used to group these enzymes into related clusters for further analysis. The initial step of the analysis of these sequences consisted of a series of multiple sequence alignments that identified groups of sequences within the superfamily that are most closely related (>30% identity), and typically consisted of those enzymes which utilize the same substrates. The sequences within each of the groups were aligned and a preliminary neighbor-joining analysis was used to determine the most divergent members of the highly related cluster (i.e. - the isopenicillin N synthase group with nine members). This most divergent pair was then used in the later analyses to reduce the size of the data set due to computational considerations. A number of sequences did not show significant similarity to any of the other sequences of the superfamily.

Approximately seventy protein sequences (Table 1.1), several of which were representative of a larger highly related group, were then

Table 1.1 Abbreviated names, accession numbers, enzyme names, and references for strains used in this study. 4-hppd is 4-hydroxyphenylpyruvate, ACC is 1-aminocyclopropane-1-carboxylate, and ? denotes that function has not been biochemically confirmed.

Name		Species	Enzyme	Reference
4hppd	S62148	Homo sapiens	4-hppd dioxygenase	Ruetschi, 1993
4hppdb	P80064	Pseudomonas sp.	4-hppd dioxygenase	Reutschi, 1992
4hppdca	gi555806	Streptomyces sp.	4-hppd dioxygenase	unpublished
4hppdh	P32754	Homo sapiens	4-hppd dioxygenase	Reutschi, 1993
4hppdm	P49429	Mus musculus	4-hppd dioxygenase	Endo, 1995
4hppdp	Q02110	Sus scrofa	4-hppd dioxygenase	Endo, 1992
4hppdr	P32755	Rattus norvegicus	4-hppd dioxygenase	Gershwin, 1987
4pgbg	X05130	Homo sapiens	prolyl 4-hydroxylase alpha subunit	Pihlajaniemi, 1987
4phah	X78949	Rattus norvegicus	prolyl 4-hydroxylase alpha subunit	unpublished
a2	P41213	Zea mays	anthrocyanidin synthase	Menssen, 1990
abh	S39897	Bos bovis	asparginyl-beta-hydroxylase	Jia, 1992
accapp	Q00985	Malus domestica	ACC oxidase	Dong, 1992
accara	Q06588	Arabadopsis thaliana	ACC oxidase	Gomez-Lim, 1993
accavo	P19464	Persea americana	ACC oxidase	McGarvey, 1992
accbra	P09052	Brassica juncea	ACC oxidase	Pua, 1992
acccarn	P31528	Dianthus caryophyllus	ACC oxidase	Wang, 1991
acckiw	M97961	Actinida deliciosa	ACC oxidase	MacDiarmid, unpublished
accmel	Q04644	Cucumis melo	ACC oxidase	Balague, 1993
accorc	L07912	Doritaenopsis sp.	ACC oxidase	Nadeau, unpublsihed
accpea	P31239	Pisum sativum	ACC oxidase	Peck, unpublished
accpet1	Q08506	Petunia hybrida	ACC oxidase	Tang, 1993
accpet3	Q08507	Petunia hybrida	ACC oxidase	Tang, 1993
accpet4	Q08508	Petunia hybrida	ACC oxidase	Tang, 1993
accpss	D13182	Pseudomonas syringae	ACC oxidase	Fukuda, 1992
acctom1	P07920	Lycopersicon esculentum	ACC oxidase	Holdsworth, 1987
acctom3	P24157	Lycopersicon esculentum	ACC oxidase	Spanu, 1991
acctom4	P05116	Lycopersicon esculentum	ACC oxidase	Holdsworth, 1987
cs1	L06213	Streptomyces clavuligerus	claviminate synthase	Marsh, 1992
cs2	L06214	Streptomyces clavuligerus	claviminate synthase	Marsh, 1992
cs4	X84101	Streptomyces clavuligerus	claviminate synthase	Hodgeson, unpublished
dacs	M63809	Streptomyces clavuligerus	deacetylcephalosporin synthase	Kovacevic, 1991
daocs	M32324	Streptomyces clavuligerus	deactetoxycephalosporin synthase	Kovacevic, 1989
daocsc	P11935	Cephalosporium acremonium	deacteoxycephalosporin C synthase	Samson, 1987
daocsn	Q03047	Nocardia lactamdurans	deacteoxycephalosporin C synthase	Coque, 1993
daocss	P18548	Streptomyces clavulgerus	deacteoxycephalosporin C synthase	Kovacevic, 1989
e8	X13437	Lycopersicum esculentum	ethylene responsive gene	Deikman, 1988
e8_2	S49975	Arabadopsis thaliana	ethylene responsive gene	Trentman, unpublished
ecox1	D64043	Escherichia coli	taurine dioxygenase	Echelard, 1988
eggdo2	S51766	Solanum melongena	dioxygenase?	Toguri, unpublished

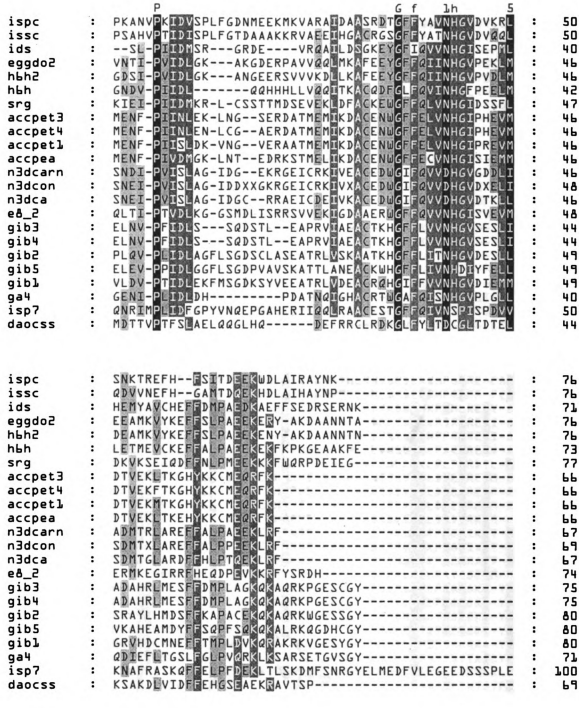
Table 1.1 continued.

Name	ACCN#	Species	Enzyme	Reference
f-ag	S27339	Tetrahymena thermophila	4-hppd dioxygenase?	Hummel, 1992
fpro	A60235	Mus musculus	4-hppd dioxygenase?	Schofield, 1991
ga4	L37126	Arabadopsis thaliana	growth enzyme	Chiang, 1995
gbb	P80193	Pseudomonas sp. AK-1	gamma-butyrobetaine hydroxylase	Ruetschi, 1993
gib1	X73314	Cucurbita maxima	giberellin 20-oxidase	Lange, 1994
gib2	X83381	Arabadopsis thaliana	giberellin 20-oxidase	Phillips, unpublished
gib3	X83380	Arabadopsis thaliana	giberellin 20-oxidase	Phillips, unpublished
gib4	X83382	Arabadopsis thaliana	giberellin 20-oxidase	Phillips, unpublished
gib5	U33330	Spinacia oleracea	giberellin 20-oxidase	Wu, unpublished
h6h	M62719	Hyoscyamus niger	hyoscyamine hydroxylase	Matsuda, 1991
h6h2	U20596	Solanum lycopersicum	hyoscyamine hydroxylase	Milligan, 1995
ids	S47971	Hordeum vulgare	low iron dioxygenase	Okumura, 1994
ids3	D10058	Hordeum vulgare	iron deficiency protein	Nakanishi, 1994
isca	X03148	Cephalosporium acremonium	isopenicillin-N synthase	Samson, 1985
isen	P05326	Emericella nidulans	isopenicillin-N synthase	Wiegel, 1988
isfs	P16020	Flavobacterium sp.	isopenicillin-N synthase	Shiffman, 1990
isnl	P27744	Nocardia lactamdurans	isopenicillin-N synthase	Coque, 1991
isp7	P40902	Schizosaccharomyces pombe	sexual development protein	Sato, 1994
ispc	P08703	Penicillium chrysogenum	isopenicillin-N synthase	Carr, 1986
issc	M19421	Streptomyces clavuligerus	isopenicillin-N synthase	Leskiw, 1988
issj	P18286	Streptomyces jumonjinensis	isopenicillin-N synthase	Shiffman, 1988
issl	P12438	Streptomyces lipmanii	isopenicillin-N synthase	Wiegel, 1988
lh1	M59183	Gallus domestica	lysyl hydroxylase	Myllyla, 1991
lig	S69666	Legionella pneumophila	legiolysin	Winterneyer, 1994
melA	M59289	Shewanella colwelliana	4-hppd dioxygenase	Fuqua, 1991
mtox1	Z77165	Mycobacterium tuberculosis	unknown	MT genome sequencing
mtox2	Z74410	Mycobacterium tuberculosis	unknown	MT genome sequencing
n3dbar	X58138	Hordeum vulgare	flavone 3-dioxygenase	Meldgaard, 1992
n3dca	X69664	Malus sp.	naringenin-3-dioxygenase	Britsch, 1993
n3dcarn	Q05964	Dianthus caryophillus	naringenin-3-dioxygenase	Britsch, 1993
n3dgr	P41090	Vitis vinifera	flavone 3-hydroxylase	Sparvoli, 1994
n3dpet	X60572	Petunia hybrida	flavanone 3-beta-hydroxylase	Britsch, 1992
n3dst	Q05965	Matthiola icana	flavanone 3-hydroxylase	Britsch 1993
scox1	S50963	Saccharomyces cerevisiae	unknown	Wedler, unpublished
srg	S44261	Arabadopsis thaliana	unknown	Callard, unpublished
trc1	JC4220	Coccidioides immitis	T-cell reactive protein	Wyckoff, 1995
trc2	JC4220	Coccidioides immitis	T-cell reactive protein	Wyckoff, 1995
yhc1	P80193	Pseudomonas sp. AK-1	gamma-butyrobetaine hydroxylase	Ruetschi, 1993

analyzed using multiple sequence alignments and neighbor-joining analysis producing five distinct and significant clusters containing most of the sequences of the superfamily.

The a-ketoglutarate/ascorbate-dependent non-heme dioxygenase group. This is the largest cluster and contains all of the ascorbate-dependent dioxygenases, as well as many of the α-ketoglutarate-dependent dioxygenases. Multiple sequence alignments of this group revealed the conservation of a number of residues, including two invariant histidines, an invariant arginine, and invariant aspartic acid, and significant regions of similarity considering the breadth of evolutionary divergence of the organisms and the number of sequences involved. Each of the noted residues is found in one of two regions of high similarity across the group, both occurring in the c-terminal region (Figure 1.1). The invariant histidines are found in two motifs: [histidine-X-aspartic acid-3X-aliphatic residue-X-two aliphatic residues] and [asparagine-7X-histidine-8-10X-arginine-X-serine] (Table 1.2). The [histidine-X-aspartic acid-50-70X-histidine] motif was previously noted in sequences including the isopenicillin N synthase subgroup (Boroviok, 1996). It is known that these histidines are iron ligands in isopenicillin N synthase (Roach,1995). The most divergent sequences in

Figure 1.1 Multiple sequence alignment of the α-keto acid-dependent cluster of the α-ketoglutarate-dependent dioxygenase superfamily. This alignment does not include all of sequences of this cluster, but includes the most divergent amino acid sequences when there are many closely related sequences that form consistent subgroups (i.e. - the ACC oxidases). Two strictly conserved histidines and one aspartic acid residue implicated in ACC oxidase activity and known to be iron ligands in isopenicillin N synthase are located in two regions of high similarity noted under the black bar.


ispc : PKANVPKIDVSPLFGDNMEEKMKVARAHDAASRDTGFFYAVNHGVDVKRL : 50
issc : PSAHVPIIDISPLFGTDAAAKKRVAEEHHGACRGSGFFYATNHGVDVQQL : 50
ids : --SE-PIIDMSR---GRDE----VRQAELDSGKEVGFIQVVNHGISEPLL : 40
eggdo2 : VNTI-PVIDLGK---AKGDERPAVVQQLMKAFEEYGFFQIINHGVPEKLM : 46
h&h2 : GDSI-PVIDLGK---ANGEERSVVVKDLLKAFEEYGFFQIINHGVPVDLM : 46
h&h : GNDV-PIIDL--------QQHHHLLVQQHTKAQDFGLFQVINHGFPEELM : 42
srg : KIEI-PIIDMKR-L-CSSTTMDSEVEKLDFACKEWGFFQLVNHGIDSSFL : 47
accpet3 : MENF-PIINLEK-LNG--SERDATMEMHKDACENWGFFELVNHGIPHEVM : 46
accpet4 : MENF-PIINLEN-LCG--AERDATMEMHKDACENWGFFELVNHGIPHEVM : 46
accpet1 : MENF-PIISLDK-VNG--VERAATMEMHKDACENWGFFELVNHGIPREVM : 46
accpea : MENF-PIVDMGK-LNT--EDRKSTMELHKDACENWGFFEVVNHGISIEMM : 46
n3dcarn : SNDI-PVISLAG-IDG--EKRGEICRKHVEACEDWGIFQVVDHGVGDDLI : 46
n3dcon : SNEI-PVISLAG-IDDXXGKRGEICRKHVXACEDWGIFQVVDHGVDXELI : 48
n3dca : SNEI-PVISLAG-IDGC--RRAEICDEHVKACEDWGIFQVVDHGVDTKLL : 46
eB_2 : QLTI-PTVDLKG-GSMDLISRRSVVEKHGDAAERWGFFQVVNHGISVEVM : 48
gib3 : ELNV-PFIDLS---SQDSTL--EAPRVHAEACTKHGFFLVVNHGVSESLI : 44
gib4 : ELNV-PFIDLS---SQDSTL--EAPRVHAEACTKHGFFLVVNHGVSESLI : 44
gib2 : PLQV-PIIDLAGFLSGDSCLASEATRLVSKAATKHGFFLITNHGVDESLL : 49
gib5 : ELEV-PPIDLGGFLSGDPVAVSKATTLANEACKWHGFFLIVNHDIYFELL : 49
gib1 : VLDV-PTIDLEKFMSGDKSYVEEATRLVDEACRQHGIFFVVNHGVDIEMM : 49
ga4 : GENI-PIIDLDH---------PDATNQHGHACRTWGAFQISNHGVPLGLL : 40
isp7 : QNRIMPLIDFGPYVNQEPGAHERIIQQLRAACESTGFFQIVNSPISPDVV : 50
daocss : MDTTVPTFSLAELQQGLHQ------DEFRRCLRDKGLFYLTDCGLTDTEL : 44

ispc : SNKTREFH--FSITDEEKWDLAIRAYNK---------- : 76
issc : QDVVNEFH--GAMTDQEKHDLAIHAYNP---------- : 76
ids : HEMYAVCHEFFDMPASDKAEFFSEDRSERNK---------- : 71
eggdo2 : EEAMKVYKEFFSLPAEEKEGY-AKDAANNTA---------- : 76
h&h2 : DEAMKVYKEFFSLPAEEKENY-AKDAANNTN---------- : 76
h&h : LETMEVCKEFFALPAEEKEFKPKGEAAKFE---------- : 73
srg : DKVKSEIQDFFNLPMEEKKFWQRPDEIEG---------- : 77
accpet3 : DTVEKLTKGHYKKCMQQRFK---------- : 66
accpet4 : DTVEKFTKGHYKKCMQQRFK---------- : 66
accpet1 : DTVEKMTKGHYKKCMQQRFK---------- : 66
accpea : DTVEKLTKEHYKKCMQQRFK---------- : 66
n3dcarn : ADMTRLAREFFALPAEEKLQF---------- : 67
n3dcon : SDMTXLAREFFALPPEEKLQF---------- : 69
n3dca : SDMTGLARDFFHLPTQEKLQF---------- : 67
eB_2 : ERMKEGIRRFHEQDPEVKKQFYSRDH---------- : 74
gib3 : ADAHRLMESFFDMPLAGKQKAQRKPGESCGY---------- : 75
gib4 : ADAHRLMESFFDMPLAGKQKAQRKPGESCGY---------- : 75
gib2 : SRAYLHMDSFFKAPACEKQKAQRKWGESSGY---------- : 80
gib5 : VKAHEAMDYFFSQPFSQKQKALRKQGDHCGY---------- : 80
gib1 : GRVHDCMNEFFTMPLDVKQRAKRKVGESYGY---------- : 80
ga4 : QDIEFLTGSLFGLPVQRKLKSARSETGVSGY---------- : 71
isp7 : KNAFRASKQFFELPFDEKLTLSKDMFSNRGYELMEDFVLEGEEDSSSPLE : 100
daocss : KSAKDLVIDFFEHGSEAEKRAVTSP---------- : 69

Figure 1.1 continued.

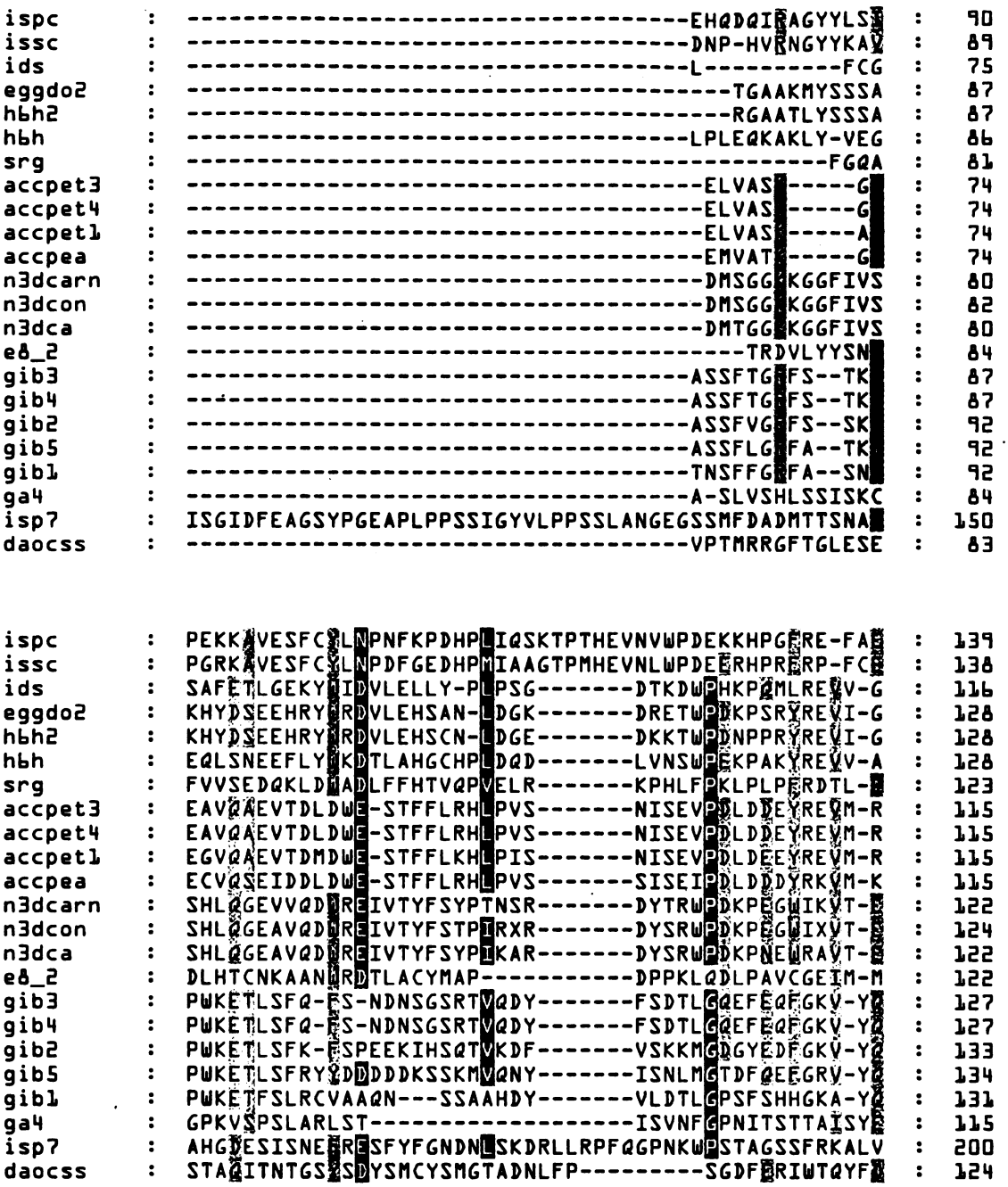


Figure 1.1 continued.

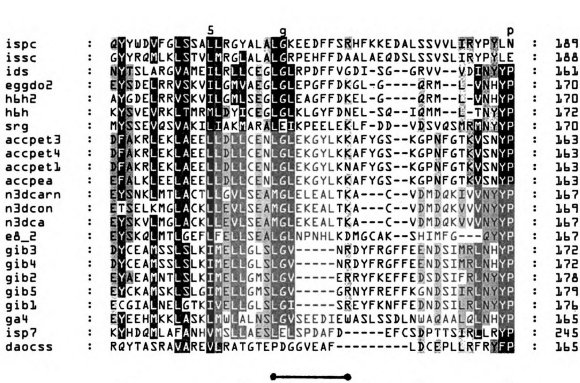


Figure 1.1 continued.

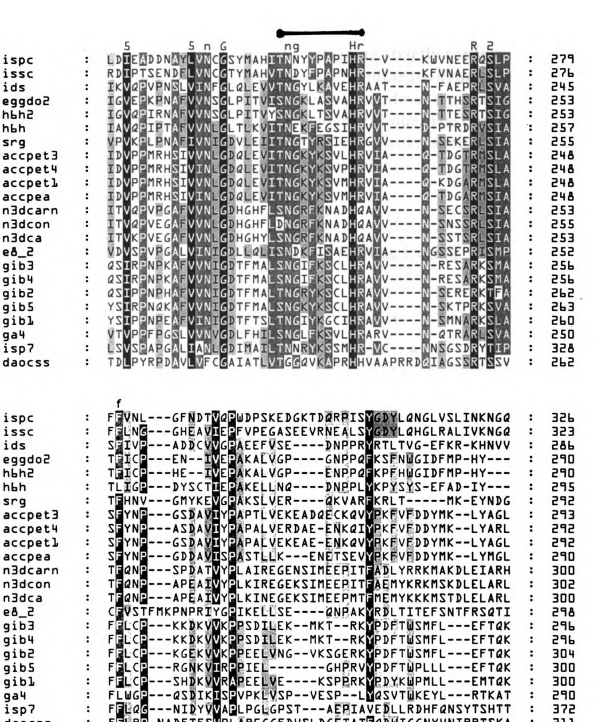


Table 1.2 Putative iron binding regions containing strictly conserved histidine residues from the groups of the α-ketoglutarate-dependent dioxygenase superfamily. Most of the conserved histidines are paired with a conserved aspartic acid, asparagine, or glutamine residue. The regions of the TfdA cluster are very similar to those of the gbb cluster, similar to that of the α-keto acid cluster, and unlike the 4-hppd regions.

Subgroup	Putative Iron Binding Region[a]	
	His-1	His-2
α-ketoacid	xHxDxxx1	Nxxxxxxxx H (8-15x) Rx2 [b]
gbb	1xxHTD	NxR1LH2
TfdA	WHxDxx4	1x1WDNRxxxH (13-14) RxT
TfdA/gbb[c]	xHxD	NxxxxH
4-hppd	1xDHxxxN	GxG1QH1Axx2x31

[a] Letters represent standard IUPAC amino acids, letters represent clusters of similar amino acids (1-LIVM, 2-STA, 3-LF, 4-FY), and x represents a non-conserved residue with the group of enzymes.

[b] The asparagine residue in this motif is conserved in all but one of the proteins in this group.

[c] This row represents the strictly conserved residues in an alignment of the sequence of both the TfdA group and the gbb group (Figure 1.8).

the group that displays this motif have only 15% identity in amino acid sequence.

The multilpe sequence alignment was used for phylogenetic analysis using neighbor-joining (NJ) method (Figure 1.2). Representative sequences were used for the analysis in Figure 1.2. Most of the sequences used are plant enzymes, and, with the exception of ACC oxidase from *Pseudomonas syringe* AK-1, cluster discreetly from the fungal and bacterial antibiotic synthesis dioxygenases (bootstrap value of 93). Though many of the sequences clearly cluster reliably with the other sequences with the same substrate utilization profile (the ACC oxidases, the giberellin oxidases), a number of the sequences show intermediate characteristics (ids,srg,isp7). Interestingly, there is no phylogentic demarcation between those dioxygenases that utilize α-ketoglutarate and those that do not (i.e. - the α-ketoglutarate-dependent flavone hydroxylases (n3dcarn) and the cephalosporin synthases (daocs)) are interspersed with enzymes, like isopenicillin N synthase (inps) and the ACC oxidases, that do not require α-ketoglutarate. This indicated that thwere is very little predictive vslue regarding substrate utilization with intermediate sequences.

The most phylogentically interesting enzyme of this cluster is the ACC oxidase from *Pseudomonas syringe* AK-1, It does not cluster with the

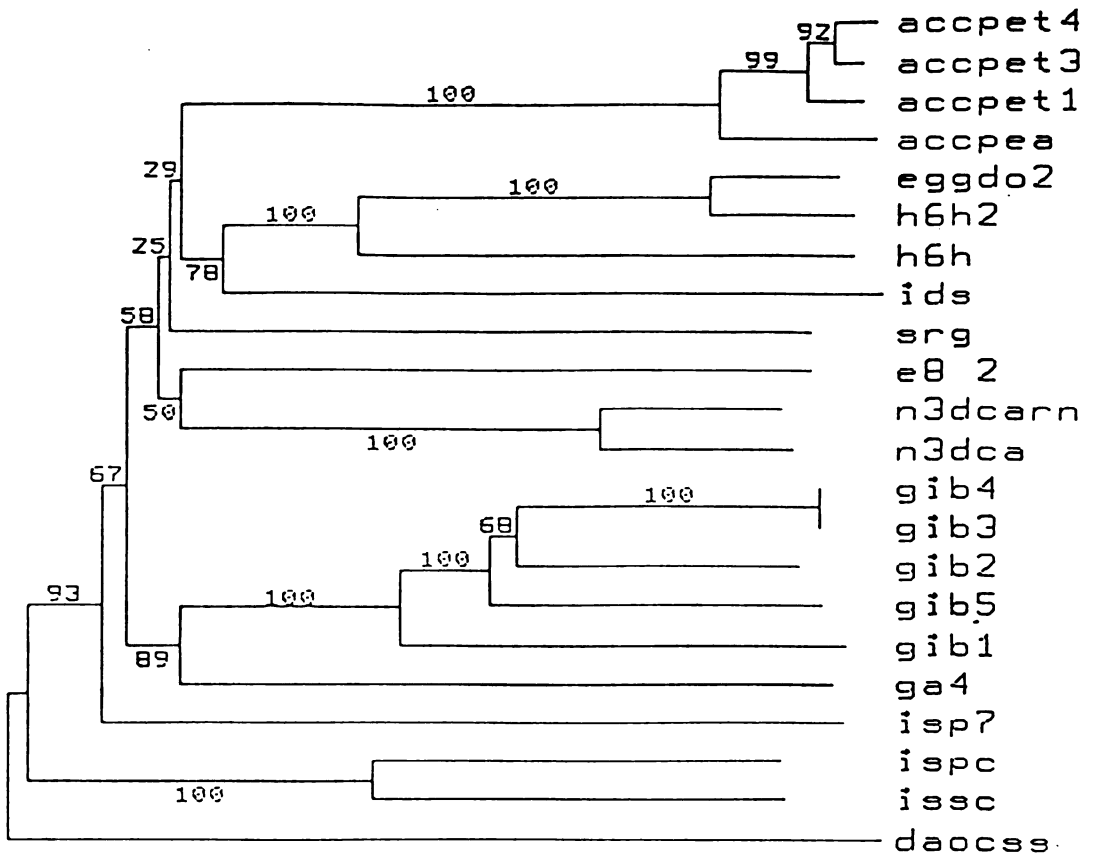


Figure 1.2 Neighbor-joining tree of the multiple sequence aligned α-keto acid cluster of the α-ketoglutarate-dependent dioxygenase superfamily. This tree shows the analysis of representative sequences of this cluster. The numbers noted are the bootstrap values for this tree, a number representing the number of times per hundred that this tree is recreated during a random reanalysis process. The closer the bootstrap is to 100, the higher the statistical significance.

other ACC oxidases, nor with the prokaryotic enzymes. This enymes is atypically large and has several regions of very high homology to the ACC oxidase sequences from tomato. This could indicate that this enzyme is the result of genetic exchange with its plant host or may have more than one biochemical function.

Of the enzymes of this group, detailed biochemical data was available only for isopenicillin N synthase. Studies of the metallocenter of this enzyme revealed that two histidines and an aspartic acid residue are four of the six ligands with the ferrous iron. These residues are among the strictly conserved residues predicted in this work. All of these residues appear in predicted regions of similarity. It is very likely that the conserved residues that serve as iron ligands in isopenicillin N synthase also have the same function in other members of this group.

The hydroxyphenylpyruvate dioxygenase group, which contains a bacterial melanin synthase (MelA), bacterial and eucaryotic 4-hydroxyphenylpyruvate dioxygenases (4-HPPD), and several other membrane associated enzymes of unknown function. The 4-hydroxyphenylpyruvate dioxygenase group enzymes are membrane associated enzymes that are likely to function as the bacterial and eucaryotic

4-hydroxyphenyl pyruvate dioxygenases. Again, these sequences exhibit two C-terminal invariant histidines, but only one resides in a region that is significantly similar to those found in the other groups. These enzymes have been referred to as a second functional subgroup of this superfamily (Bradley, 1986), as they catalyze a reaction where the α-keto group as well as the aromatic ring of the substrate is oxygenated using the same mechanism as the two substrate dioxygenases (Pascal, 1985).

Database searches for sequence that are similar to human 4-hydroxyphenylpyruvate dioxygenase noted distant but significant similarity (27.9% identity, 51.7% similarity) with a reported melanin synthesis enzyme of *Shewanella colwelliana* (Fuqua, 1991), and two other prokaryotic sequences. All statistically significant (as determined by BLASTP (Altshul, 1990)) were examined for conserved resides that were potential iron ligands. These sequences were then aligned (Figure 1.3) and found to be highly similar, containing many conserved residues even though the group contains eukaryotic and prokaryotic sequences. The alignment noted several potential iron ligands: three conserved histidines, three conserved tyrosines, two aspartic acids, and two glutamic acids. Phylogenetic analysis using the NJ method indicates that the prokaryotic

and eukaryotic sequences are significantly related, but form discreet subclusters within the group (bootstrap of 100, Figure 1.4).

Human 4-hppd and the *Pseudomonas* MelA show strong c-terminal similarity (52% identity over 71 amino acids). This c-terminal region is the same region containing the conserved histidines and regions of similarity implicated in iron binding in the α-keto acid dependent group. These regions do have some similarity, leaving open the possibility of the presence of specific binding motifs for these enzymes.

The gamma-butyrobetaine hydroxylase group, has only one well characterized seqeunce (gbb), and contains several new sequences of unknown function. Gamma-butryobetaine hydroxylase is an α-ketoglutarate-dependent enzyme, and has been partially biochemically characterized (Blanchard, 1983, Englard, 1985, Ng, 1991). Multiple sequence alignment of these sequences indicated the presence of several conserved residues, including four aspartic acid residues, three arginines, two histidines, and an asparagine (Figure 1.5). This group also displays a c-terminal histidine-X-aspartic acid motif and a glutamate-4X-histidine motif within regions of high similarity. Binding studies using gbb indicate that

Figure 1.3 Multiple sequence alignment of the 4-hydroxyphenyl-pyruvate dioxygenase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. This cluster displays a considerable number of conserved residues.

```
                                        vg akqaa  5    Gf   ay g
4hppdb  :  --ADLYENPMGLMGFEFIELASPTPNTLEPIF-EIMGFTKVATH-  :  41
lig     :  ---MQNNNPCGLDGFAFLEFSGPQRNKLHQQF-SEMGFQAVAHH-  :  40
mela    :  --MASEQNPLGLLGIEFTEFATPDLDFMHKVF-IDFGFSKLKKH-  :  41
trc1    :  LQPAQPSDLNQYRGYDHVHWYVGNAKQAATYYVTRMGFERVAYKG  :  45
trc2    :  LQPAQPSDLNQYRGYDHVHWYVGNAKQAATYYVTRMGFERVAYKG  :  45
4hppd   :  K--GAKPERGRFLHPHSVTFLVGNAKQAASFYCSKMGFEPLAYRG  :  43
4hppdh  :  K--GAKPERGRFLHPHSVTFLVGNAKQAASFYCSKMGFEPLAYRG  :  43
4hppdm  :  K--GPKPERGRFLHPHSVTFLVGNAKQAASFYCNKMGFEPLAYRG  :  43
fpro    :  K--GPKPERGRFLHPHSVTFLVGNAKQAASFYCNKMGFEPLAYRG  :  43
4hppdr  :  K--GPKPERGRFLHPHSVTFLVGNAKQAASFYCNKMGFEPLAYRG  :  43
4hppdp  :  K--GEKPERGRFLHPHSVTFLVGNAKQAASYYCSKIGFEPFAYRG  :  43
4hppdce :  K--GAKPDIGTFVAPDHVRFVVGNAKQAAYLYCANFGFEPFAYRG  :  43
f-ag    :  KPVGERPTGGKFLGVDHLHFWVGNAKQAAGVYTSRFGFEYYAYRG  :  45
4hppdsa :  PDTARQADPFPVKGMDAVFAVGNAKQAA-HYSIAFGMQLVAYSG  :  44

                e gs4    h       i fb   s
4hppdb  :  -----RSKDVHLYRQGAINLIL---------------NNEPHSVA  :  66
lig     :  -----KNQDITLFRQGEMQFIV---------------NAASHCQA  :  65
mela    :  -----KQKDIVYYSQNDINFLL---------------NNEKQGFS  :  66
trc1    :  LETGSKAVASHVVRNGNITFILTSPLRSVEQASRFPEDEALLKEI  :  90
trc2    :  LETGSKAVASHVVRNGNITFILTSPLRSVEQASRFPEDEALLKEI  :  90
4hppd   :  LETGSREVVSHVIKQGKIVFVLSSAL--------NPNLK----EM  :  76
4hppdh  :  LETGSREVVSHVIKQGKIVFVLSSAL--------NPNLK----EM  :  76
4hppdm  :  LETGSREVVSHVIKRGKIVFVLCSAL--------NPNLK----EM  :  76
fpro    :  LETGSREVVSHVIKQGKIVFVLCSAL--------NPNLK----EM  :  76
4hppdr  :  LETGSREVVSHVIKQGKIVFVLCSAL--------NPNLK----EM  :  76
4hppdp  :  LETGSREVVSHVVKQGKIVFVFSSAL--------NPNLK----EM  :  76
4hppdce :  LETGSRITAQHAIRQGKIVFIFESAL--------LPDNS----EL  :  76
f-ag    :  LETGSREVATHVVRNKQGVTLAFSTP--------YGNDKDNQREM  :  82
4hppdsa :  PENGSRETASYVLTNGSARFVLTSVIKPATPWGHF---------L  :  80

                h   HGdg  dba  V d          A   Ga        d
4hppdb  :  SYFAAEHGPSVCGVAFRVKDSQKAYKRALELGAQ----PIHIETG  :  107
lig     :  EAHASTHGPGACAWGFKVKDAKAAFQHAIAHGGI----AEQ-DAP  :  105
mela    :  AQFAKTHGPATSSMGVRVEDANFAFEGAVARGAK----BAADEV-  :  106
trc1    :  HAHLERHGDGVKDVAFEVDCVESVFSAAVRNGAEVVSDVRTVEDE  :  135
trc2    :  HAHLERHGDGVKDVAFEVDCVESVFSAAVRNGAEVVSDVRTVEDE  :  135
4hppd   :  GDHLVKHGDGVKDIAFEVEDCDYIVQKARERGAKIMREPWVEQDK  :  121
4hppdh  :  GDHLVKHGDGVKDIAFEVEDCDYIVQKARERGAKIMREPWVEQDK  :  121
4hppdm  :  GDHLVKHGDGVKDIAFEVEDCDHIVQKARERGAKIVREPWVEQDK  :  121
fpro    :  GDHLVKHGDGVKDIAFEVEDCNHIVQKARERGAKIVRDAWVEQDK  :  121
4hppdr  :  GDHLVKHGDGVKDIAFEVEDCEHIVQKARERGAKIVREPWVEEDK  :  121
4hppdp  :  GDHLVKHGDGVKDIAFEVEDCDYIVQKARERGAIIVREPWIEQDK  :  121
4hppdce :  GNHLVQHGDGVKDVCFEVEDLDSIIAHAKAAGATIVHDITEESDA  :  121
f-ag    :  NQHQSLHGDGVKDVAFAVEDCHSIYNKAIQRGAKCAYPPQDLKDE  :  127
4hppdsa :  ADHVAEHGDGVVDLAHEVPDARAAHAYAIEHGARSVAEPYELKDE  :  125
```

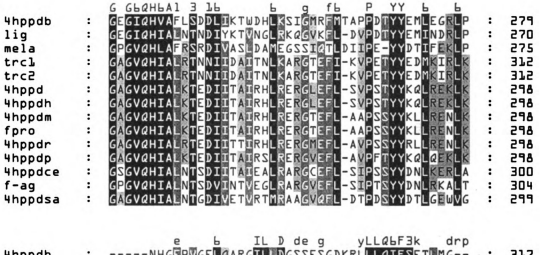
<u>Figure 1.3</u> continued.

```
                  g     a 6 t G    htl6     y g    pg
4hppdb    :  PMELNLPAIKGIGGAPLYLIDRFGEGSSIYD--IDFVFLEGVDRH  :  150
lig       :  HANHGLPAIQAIGGSVIYFVDEEHQPFS-HE--WNITSPEPV---  :  144
mela      :  -KDLPYPAIYGIGDSLIYFIDTFGDDNNIYT--SDFEALDEPIIT  :  148
trc1      :  DGQIKMATIRTYGETTHTLIERSGYRGGFMPGYRMESNADATSKF  :  180
trc2      :  DGQIKMATIRTYGETTHTLIERSGYRGGFMPGYRMESNADATSKF  :  180
4hppd     :  FGKVKFAVLQTYGDTTHTLVEKMNYIGQFLPGYEAPAFMDPLLPK  :  166
4hppdh    :  FGKVKFAVLQTYGDTTHTLVEKMNYIGQFLPGYEAPAFMDPLLPK  :  166
4hppdm    :  FGKVKFAVLQTYGDTTHTLVEKINYTGRFLPGFEAPTYKDTLLPK  :  166
fpro      :  FGKVKFAVLQTYGDTTHTLVEKINYTGRFLPGFEAPTYKDTLLPK  :  166
4hppdr    :  FGKVKFAVLQTYGDTTHTLVEKINYTGRFLPGFEAPTYKDTLLPK  :  166
4hppdp    :  FGKVKFAVLQTFGDTTHTLVEKMNYTGCFLPGFEAPTFTDPLLSK  :  166
4hppdce   :  DGSIRYATLRTYGETDHTLLERKNYRGAFLPGFKAHPMPATFFKT  :  166
f-ag      :  HGSVTIAAVHTYGEVIHTFIQRNDYKGFFMPGFVAHPLKDPLNNV  :  172
4hppdsa   :  HGTVVLAAIATYGKTRHTLVDRTGYDGPYLPGYVAAA---PIVEP  :  167

                    l  6DH vgN      M       Y    F          6
4hppdb    :  PVGAGLKIIDHLTHNVYRGRMAYWANFYEKLFNFREIRYFD---I  :  192
lig       :  -VGNGLTAIDHLTHNVYRGNMDKWASFYASIFNFQEIRFFN---I  :  185
mela      :  Q-EKGFIEVDHLTNNVHKGTMEYWSNFYKDIFGFTEVRYFD---I  :  189
trc1      :  LPKVVLERIDHCVGNQDWDEMERVCDYYEKILGFHRFWSVDDKDI  :  225
trc2      :  LPKVVLERIDHCVGNQDWDEMERVCDSYEKILGFHRFWSVDDKDI  :  225
4hppd     :  LPKCSLEMIDHIVGNQPDQEMVSASEWYLKNLQFHRFWSVDDTQV  :  211
4hppdh    :  LPKCSLEMIDHIVGNQPDQEMVSASEWYLKNLQFHRFWSVDDTQV  :  211
4hppdm    :  LPRCNLEIIDHIVGNQPDQEMQSASEWYLKNLQFHRFWSVDDTQV  :  211
fpro      :  LPRCNLEIIDHIVGNQPDQEMQSASEWYLKNLQFHRFWSVDDTQV  :  211
4hppdr    :  LPSCNLEIIDHIVGNQPDQEMESASEWYLKNLQFHRFWSVDDTQV  :  211
4hppdp    :  LPKCGLEIIDHIVGNQPDQEMESASQWYMRNLQFHRFWSVDDTQI  :  211
4hppdce   :  LPRVGLNFLDHCVGNQPDLQMDSAVQWYEKVLKFHRFWSVDDSMI  :  211
f-ag      :  LPDISYNYVDHIVGNQPDNMMTSAADWYEKTLDFHRFWSVDDSMI  :  217
4hppdsa   :  PAHRTFQAIDHCVGNVELGRMNEWVGFYNKVMGFTNMKEFVGDDI  :  212

                e 3 L S v6           P6Nepa        sQI E56
4hppdb    :  KGEYTGLTSKAMTAPDGMIRIPLNE---ESSKGAGQIEEFLMQFN  :  234
lig       :  KGKMTGLVSRALGSPCGKIKIPLN----ESKDDLSQIEEFLHEYH  :  226
mela      :  KGSQTALISYALRSPDGSFCIPINE---GKGDDRNQIDEYLKEYD  :  231
trc1      :  CTEFSALKSIVMASPNDIVKMPINEPAKGKKQ--SQIEEYVDFYN  :  268
trc2      :  CSEFSALKSIVMASPNDIVKMPINEPAKGKKQ--SQIEEYVDFYN  :  268
4hppd     :  HTEYSSLRSIVVANYEESIKMPINEPAPGKKK--SQIQEYVDYNG  :  254
4hppdh    :  HTEYSSLRSIVVANYEESIKMPINEPAPGKKK--SQIQEYVDYNG  :  254
4hppdm    :  HTEYSSLRSIVVTNYEESIKMPINEPAPGRKK--SQIQEYVDYNG  :  254
fpro      :  HTEYSSLRSIVVTNYEESIKMPINEPAPGRKK--SQIQEYVDYNG  :  254
4hppdr    :  HTEYSSLRSIVVANYEESIKMPINEPAPGRKK--SQIQEYVDYNG  :  254
4hppdp    :  HTEYSALRSVVMANYEESIKMPINEPAPGKKK--SQIQEYVDYNG  :  254
4hppdce   :  HTEYSALRSIVVTNFEETIKMPINEPATSDKKAISQIQEYVDYYG  :  256
f-ag      :  HTEFSSLRSIVMTDYDQKIKMPINEPADGKRK--SQIQEYIDFYA  :  260
4hppdsa   :  ATEYSALMSKVVADGTLKVKFPINEPALAKKK--SQIDEYLEFYG  :  255
```

Figure 1.3 continued.



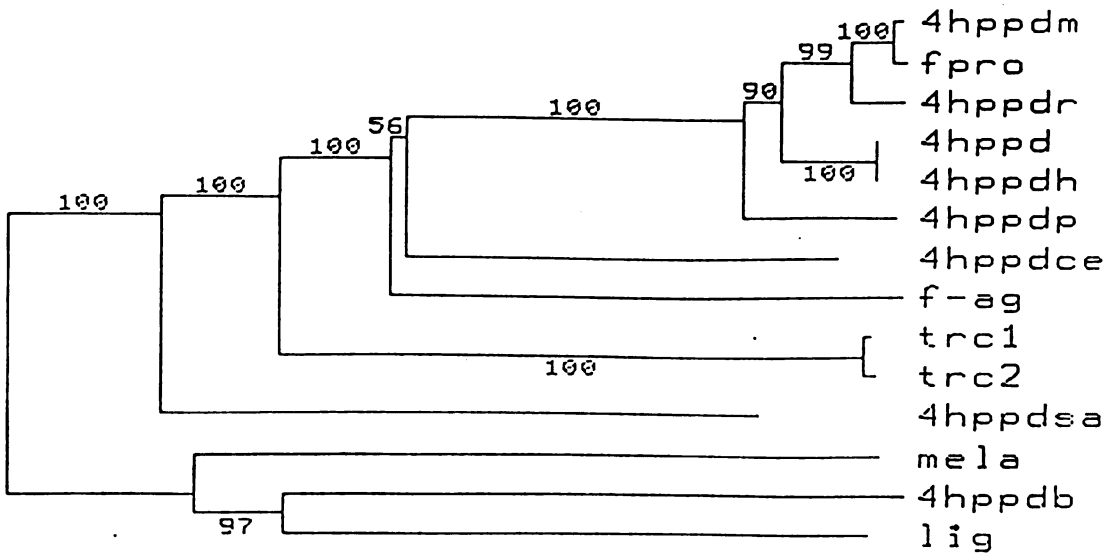


Figure 1.4 Neighbor-joining tree of the multiple sequence aligned 4-hydroxyphenylpyruvate (4-hppd) dioxygenase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. The numbers noted are the bootstrap values for this tree, a number representing the number of times per hundred that this tree is recreated during a random reanalysis process. The closer the bootstrap is to 100, the higher the statistical significance. This tree shows highly significant clustering of eucaryotic and procaryotic 4-hppd-like sequences.

Figure 1.5 Multiple sequence alignment of the gamma-butryobetaine hydroxylase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. This cluster contains a single well characterized enzyme, gbb, and several other recent sequences that probably have a similar function. This cluster has a number of strictly conserved residues, including four aspartic acid residues, three arginines, two histidines, and an asparagine.

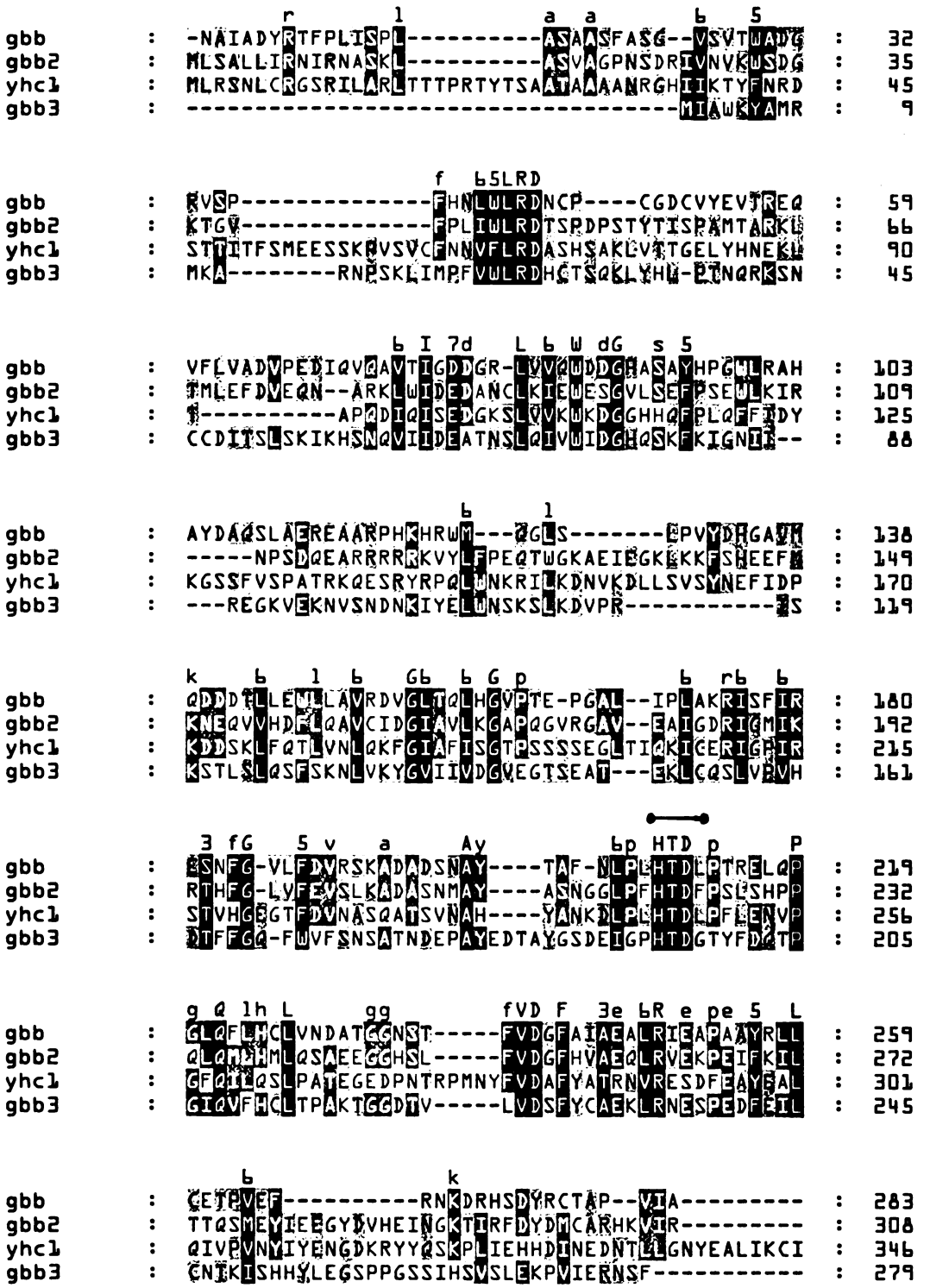
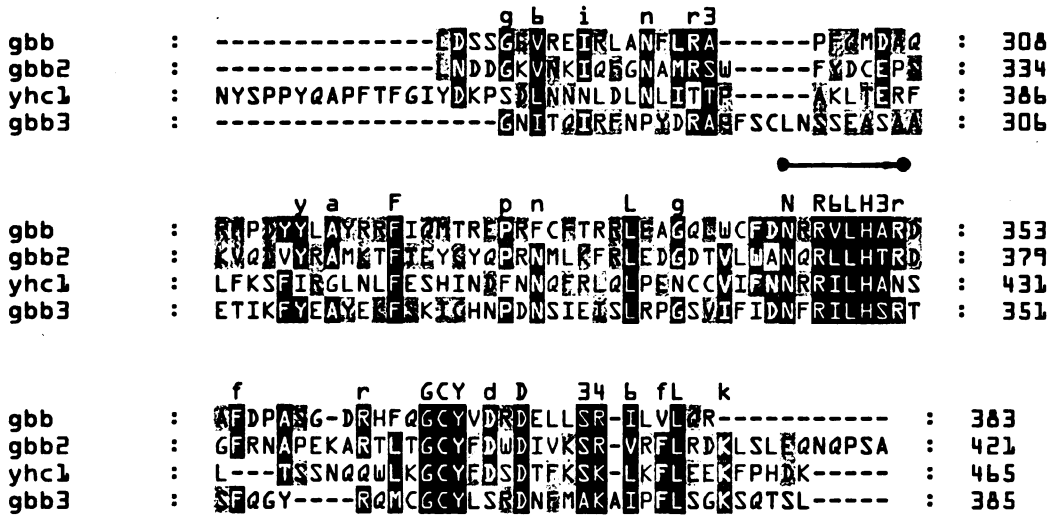

gbb
gbb2
yhc1
gbb3

Figure 1.5 continued.



one of the positively charged residues is involved in α-ketoglutarate binding in this strain (Ng, 1991).

The 2,4-dichlorophenoxyacetic acid/a-ketoglutarate-dependent dioxygenase (TfdA) group, which contains TfdA, close relatives (>90% identity at the protein sequence level), and distant relatives (>27% identity at the protein sequence level). Only one of the distantly related sequences has been characterized to any extent, and is know as taurine dioxygenase (TauD, noted in this work as ecox1)(van der Ploeg,1996). Multiple sequence alignment of this group indicated that there are a number of conserved residues, including three histidines, three threonines, two arginines, two aspartic acid residues, and an asparaginine residue (Figure 1.6). There is a c-terminal pair of histidine motifs: [tryptophan-histidine-X-aspartic acid] and [tryptophan-aspartic acid-asparaginie-3x-histidine]. The forat motif somewhat resembles the first c-terminal histidine region of the α-keto acid group. NJ analysis of these sequences indicated that the more distantly related sequences have diverged significantly from TfdA and from each other, as indicated by a lack of reliable branches (Figure 1.7). Chemical inactivation studies of TfdA indicated that histidines are very likely to be the essential iron lignads (Fukumori, 1993a).

Figure 1.6 Multiple sequence alignment of the 2,4-D/α-ketoglutarate-dependent dioxygenase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. This cluster contains two highly similar TfdA sequences, TfdA and $TfdA_{rasc}$, and several other recent sequences that probably have a similar mechanism of action. Ecox1 encodes taurine dioxygenase and id known as TauD. The other enzymes of this subgroup are not biochemically characterized. The members of this subgroup share a number of strictly conserved residues, including three histidines, three threonines, two arginines, two aspartic acids, and an asparagine residue.

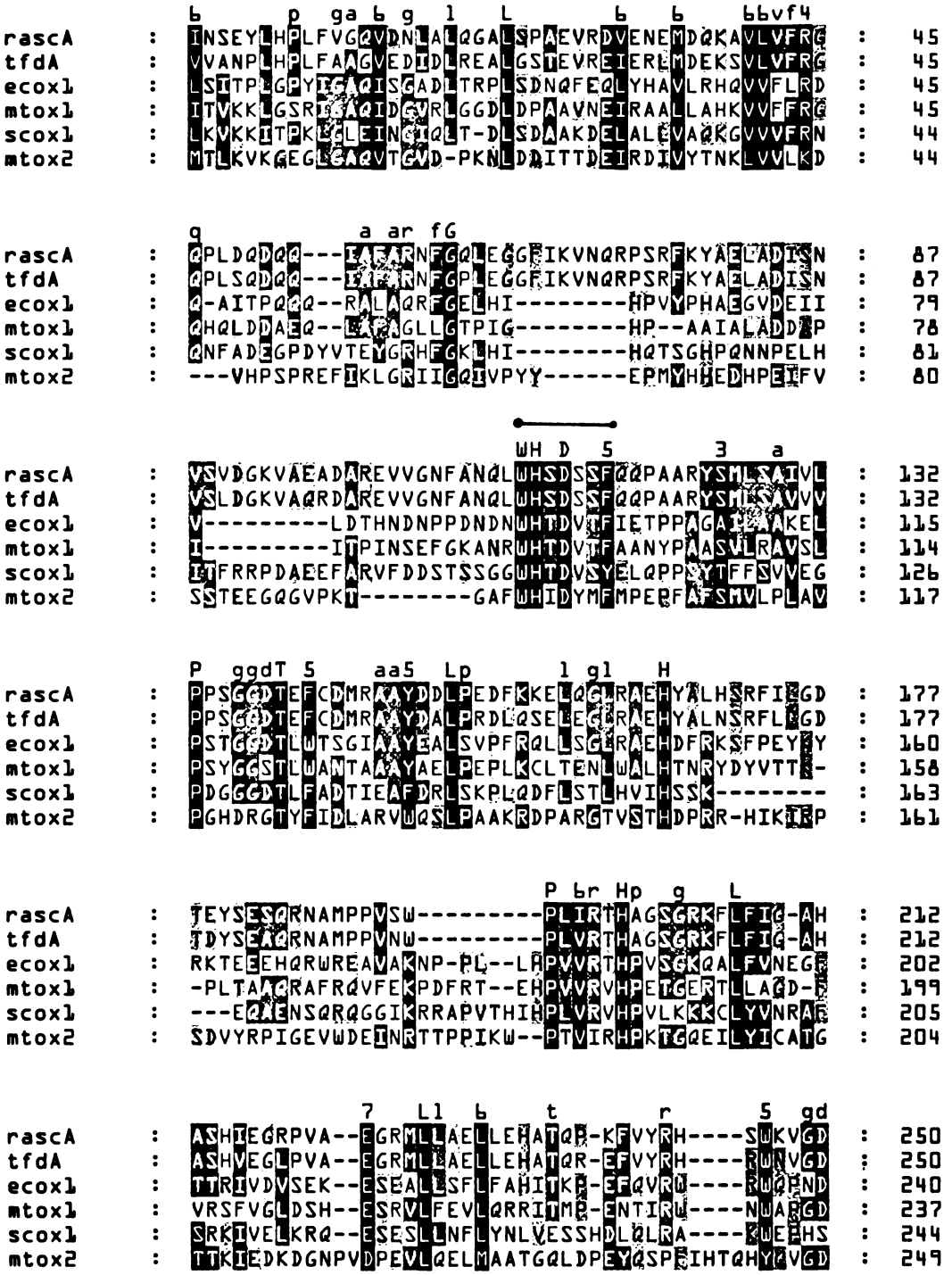
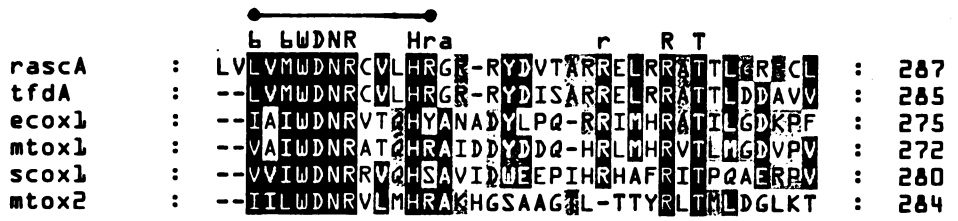

rascA
tfdA
ecox1
mtox1
scox1
mtox2

Figure 1.6 continued.



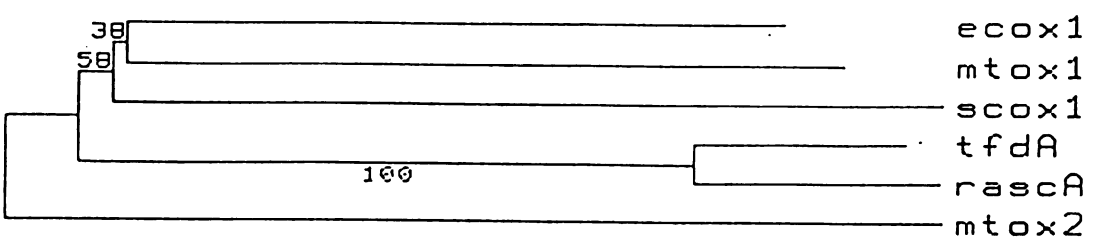


Figure 1.7 Neighbor-joining tree of the multiple sequence aligned 2,4-dichlorophenoxyacetic acid/α-ketoglutarate-dependent (TfdA) dioxygenase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. The numbers noted are the bootstrap values for this tree, a number representing the number of times per hundred that this tree is recreated during a random reanalysis process. The closer the bootstrap is to 100, the higher the statistical significance. The very closely related (>90% identity) TfdA and RascA cluster reliably, with the other sequences being nearly equally evolutionary distant from one another.

Figure 1.8 Multiple sequence alignment of the 2,4-D/α-ketoglutarate-dependent dioxygenase cluster and the gamma-butyrobetaine hydroxylase cluster of the α-ketoglutarate-dependent dioxygenase superfamily. These clusters show the most similarity within the superfamily, sharing two regions of high similarity containing histidine and aspartic acid or asparagine residues.

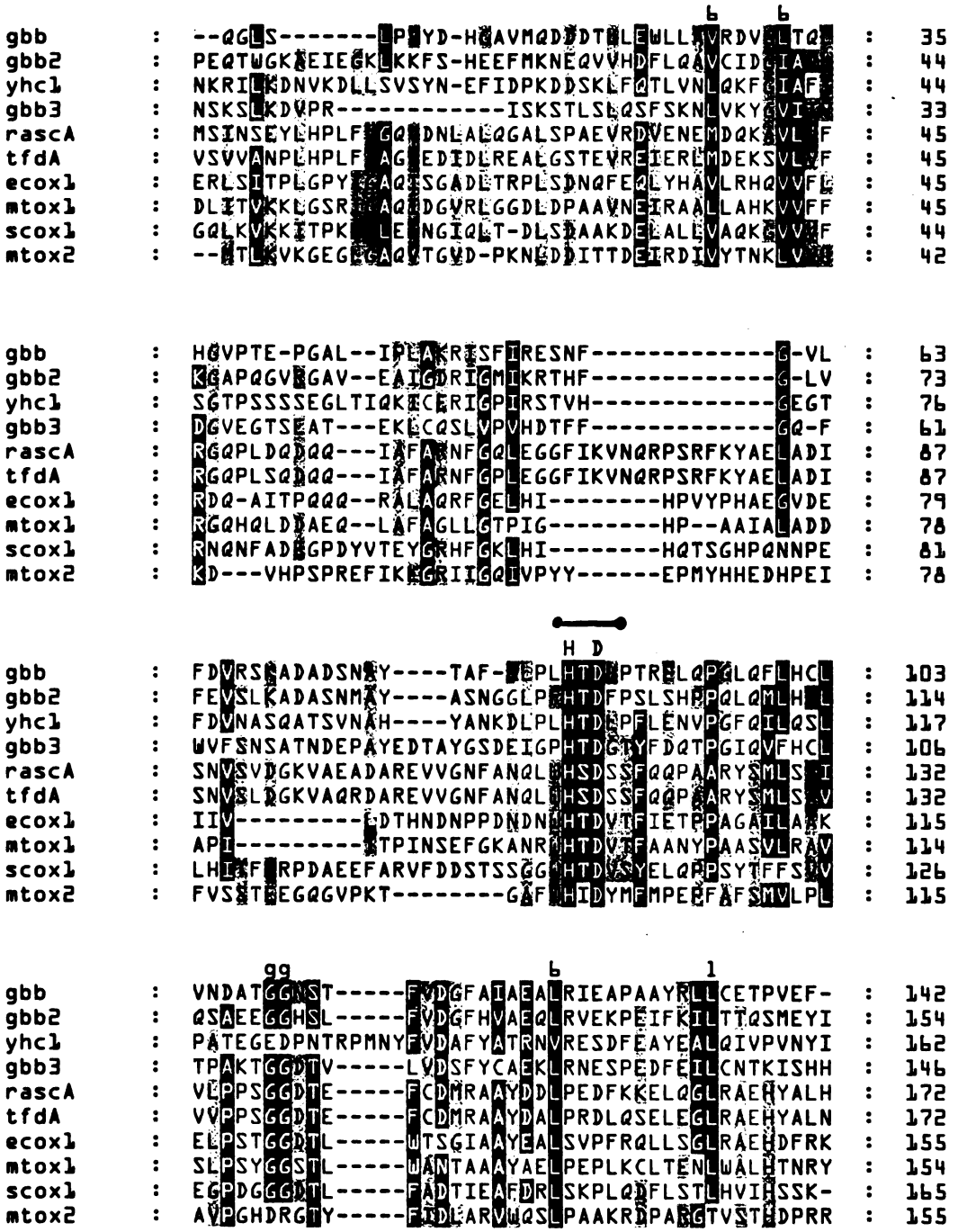
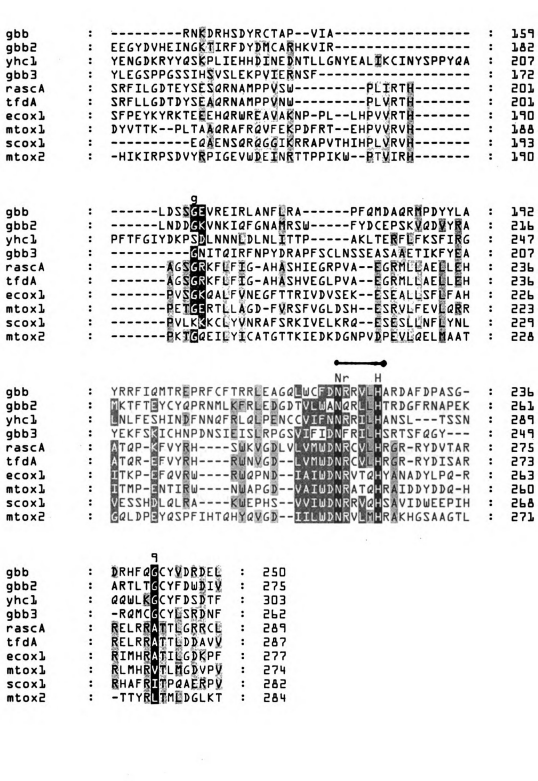

b b
gbb : --QGLS-------LPEYD-HGAVMQDDDTGLEWLLAVRDVELTQL : 35
gbb2 : PEQTWGKAEIEGKLKKFS-HEEFMKNEQVVHDFLQAVCIDGIA : 44
yhc1 : NKRILKDNVKDLLSVSYN-EFIDPKDDSKLFQTLVNLQKFGIAF : 44
gbb3 : NSKSLKDVPR------------ISKSTLSLQSFSKNLVKYGVI : 33
rascA : MSINSEYLHPLFGQIDNLALQGALSPAEVRDVENEMDQKAVLGF : 45
tfdA : VSVVANPLHPLFAGLEDIDLREALGSTEVREIERLMDEKSVLAF : 45
ecox1 : ERLSITPLGPYGAQTSGADLTRPLSDNQFEQLYHAVLRHQVVFL : 45
mtox1 : DLITVKKLGSRGAQIDGVRLGGDLDPAAVNEIRAALLAHKVVFF : 45
scox1 : GQLKVKKITPKGLENGIQLT-DLSDAAKDELALLVAQKGVVGF : 44
mtox2 : --ETLKVKGEGEGAQVTGVD-PKNEDDITTDEIRDIVYTNKLVG : 42
gbb : HGVPTE-PGAL--IPLAKRISFIRESNF----------G-VL : 63
gbb2 : KGAPQGVRGAV--EAIGDRIGMIKRTHF----------G-LV : 73
yhc1 : SGTPSSSSEGLTIQKICERIGPIRSTVH----------GEGT : 76
gbb3 : DGVEGTSEAT---EKLCQSLVPVHDTFF----------GQ-F : 61
rascA : RGQPLDQDQQ---IAFARNFGQLEGGFIKVNQRPSRFKYAELADI : 87
tfdA : RGQPLSQDQQ---IAFARNFGPLEGGFIKVNQRPSRFKYAELADI : 87
ecox1 : RDQ-AITPQQQ--RALAQRFGELHI--------HPVYPHAEGVDE : 79
mtox1 : RGQHQLDDAEQ--LAFAGLLGTPIG--------HP--AAIALADD : 78
scox1 : RNQNFADEGPDYVTEYGRHFGKLHI--------HQTSGHPQNNPE : 81
mtox2 : KD---VHPSPREFIKEGRIIGQIVPYY------EPMYHHEDHPEI : 78
H D
gbb : FDVRSAADADSNAY----TAF-VEPLHTDVPTRELQPGLQFLHCL : 103
gbb2 : FEVSLKADASNMAY----ASNGGLPEHTDFPSLSHPPQLQMLHGL : 114
yhc1 : FDVNASQATSVNAH----YANKDLPLHTDEPFLENVPGFQILQSL : 117
gbb3 : WVFSNSATNDEPAYEDTAYGSDEIGPHTDGTYFDQTPGIQVFHCL : 106
rascA : SNVSVDGKVAEADAREVVGNFANQLGHSDSSFQQPAARYSMLSGI : 132
tfdA : SNVSLDGKVAQRDAREVVGNFANQLGHSDSSFQQPAARYSMLSGV : 132
ecox1 : IIV---------LDTHNDNPPDNDNLHTDVTFIETPPAGAILAAK : 115
mtox1 : API---------STPINSEFGKANRLHTDVTFAANYPAASVLRAV : 114
scox1 : LHIFFGRPDAEEFARVFDDSTSSGGLHTDVAYELQPPSYTFFSGV : 126
mtox2 : FVSSTGEGQGVPKT--------GAFLHIDYMFMPEPFAFSMVLPL : 115
gg b l
gbb : VNDATGGNST-----FVDGFAIAEALRIEAPAAYRLLCETPVEF- : 142
gbb2 : QSAEEGGHSL-----FVDGFHVAEQLRVEKPEIFKILTTQSMEYI : 154
yhc1 : PATEGEDPNTRPMNYFVDAFYATRNVRESDFEAYEALQIVPVNYI : 162
gbb3 : TPAKTGGDTV-----LVDSFYCAEKLRNESPEDFEILCNTKISHH : 146
rascA : VLPPSGGDTE-----FCDMRAAYDDLPEDFKKELQGLRAEHYALH : 172
tfdA : VVPPSGGDTE-----FCDMRAAYDALPRDLQSELEGLRAEHYALN : 172
ecox1 : ELPSTGGDTL-----WTSGIAAYEALSVPFRQLLSGLRAEHDFRK : 155
mtox1 : SLPSYGGSTL-----WANTAAAYEALPEPLKCLTENLWALHTNRY : 154
scox1 : EGPDGGGDTL-----FADTIEAFDRLSKPLQDFLSTLHVIHSSK- : 165
mtox2 : AVPGHDRGTY-----FIDLARVWQSLPAAKRDPARGTVSTHDPRR : 155

Figure 1.8 continued.



Interestingly, the c-terminal motifs of this cluster and the gamma-butyrobetaine hydroxylase group are quite similar. The amino acid sequences of TfdA and Gbb have display low identity but a high (53%) similarity. When optimally aligned, only a limited version of the two c-terminal histidine motifs emerge as conserved between the two groups (Figure 1.8). These motifs - [histidine-X-aspartic acid] and [asparagine-4X-hisidine] - occur in the region of greatest similarity between the two groups. These results indicate that TfdA and Gbb may not be evolutionarily related, but may share an active site with a similar three dimensional configuration.

The clavamate synthase group, with several very closely related members. This group appears to form a unique group in this superfamily, and bears little resemblance to any other sequence in the sequence databases. Some initial biochemical work indicates that cysteine residues are essential ligands in the metallocenter of this enzyme. These sequences are so similar, performing alignment and phylogenetic analysis do not generate any additional information.

The remaining sequences. Several of the other sequences (lysyl hydroxylase, prolyl hydroxylase, aspartyl hydroxylase) did not share

sufficient similarity to any of the other sequences to be included in the groups defined above. These sequences are also thought to use histidine residues to bind iron, and several short regions of similarity that contain histidine residues have been noted (Myllylä,1992). Thses sequences may be distantly related, but this relationship is not revealed using the techniques used in this study.

SUMMARY

It was possible to generate a phylogeny of the α-ketoglutarate-dependent dioxygenase superfamily that identified distinct clusters of divergent enzymes that retain patterns of conserved residues. This analysis provides some insight into the evolution of these enzymes, illustrating that the superfamily is likely to be polyphyletic, though there are several clusters of enzymes within the superfamily that are probably related or share domains or regions which are related. The clustering of the α-ketoglutarate-dependent dioxygenases with the ascorbate-dependent dioxygenases is intriguing. Further, the presence of regions of similarity containing invariant histidines among severalk of the groups could indicate convergent evolution of the metallocenters from disparate ancestral enzymes or the divergent evolution of ancient motifs capable of binding metals. Several of the enzymes of this superfamily do not exhibit similarity to nay of the other sequences, though do contain short regions of similarity. These enzymes may be unrelated to any of the outline subgroups outlined in this work, or may simply be to divergent to be identified with these methods. Since the initial analysis of this superfamily, new sequences have become available, including several bacterial genomes, providing many new but as yet

unstudied members of the outlined subgroups of the α-ketoglutarate-dependent superfamily. It is hoped that this kind of analysis will aid the ongoing biochemical analysis of these enzymes.

REFERENCES

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

Balague, C., Watson, C. F., Turner, A. J., Rouge, P., Picton, S., Pech, J. C.,and D. Grierson. 1993. Isolation of a ripening and wound-induced cDNA from *Cucumis melo* L. encoding a protein with homology to the ethylene-forming enzyme. Eur. J. Biochem. **212**: 27-34.

Borovok, I., Landman, O., Kriesberg-Zakarin, R., Aharonowitz, Y., and G. Cohen. 1996. Ferrous active site of isopenicillin N synthase: genetic and sequence analysis of the endogenous lignads. Biochemistry **35**: 1981-1987.

Blanchard, J. S., and S. Englard. 1983. Gamma-butyrobetaine hydroxylase: primary and secondary tritium effects. Biochemistry **22**: 5922-5929.

Bradley, F. C., Lindsteadt, S., Lipscomb, J. D., Que, L. Jr., Roe, A. L., and M. Rundgren. 1986. 4-hydroxyphenylpyruvate hydroxylase is an iron-tyrosinate protein. J. Biol. Chem. **261**: 11693-11696.

Britsch, L., Ruhnau-Brich, B., and G. Forkman. 1992. Molecular cloning, sequence analysis, and in vitro expression of flavanone 3 beta-hydroxylase from *Petunia hybrida*.J. Biol. Chem. **267**: 5380-5387.

Britsch, L., Dedio, J., Saedler, H., and G. Forkman. 1993. Molecular characterization of flavanone 3 beta-hydroxylases. Consensus sequence, comparison with related enzymes and the role of conserved histidine residues. Eur. J. Biochem. **217**: 745-754.

Callard, D., Axelos, M., and L. Mazzolini. 1994. Submitted to the EMBL Data Library.

Carr, L. G., Skatrud, P. L., Scheetz, M. E. II, Queener, S. W., and T. D. Ingolia. 1986. Cloning and expression of the isopenicillin N synthase from *Penicillium chrysogenum*. Gene **48**: 257-259.

Chiang, H. H., Hwang, I., and H. M. Goodman. 1995. Isolation of the Arabidopsis GA4 locus. Plant Cell **7**: 195-201.

Coque, J. J. R., Martin, J. F., Calzada, J. G., and P. Liras. 1991. The cephamycin biosynthetic genes *pcbAB*, encoding a large multi-domain peptide synthetase, and *pcbC* of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Streptomyces clavuligerus*. Mol. Microbiol. **5**: 1125-1133.

Coque, J. J. R., Martin, J. F., and P. Liras. 1993. Characterization and expression in *Streptomyces lividans* of *cefD* and *cefE* from *Nocardia lactamdurans*: the organization of the cephamycin gene cluster

differes from that in *Streptomyces clavuligerus*. Mol. Gen. Genet. **236**: 453-458.

Deikman, J. and R. L. Fischer. 1988. Interaction of a DNA binding factor with the 5'-flanking region of an ethylene-responsive fruit ripening gene from tomato. EMBO J. **7**: 3315-3320.

Echelard, Y., Dymetryszyn, J., Drolet, M., A. Sasarman. 1988. Nucleotide sequence of the *hemB* gene of *Escherichia coli* K12. Mol. Gen. Genet. **214**: 503-508.

Endo, F., Awata, H., Tanque, A., Ishiguro, M., Eda, Y., Titani, K., and I. **267**: 24235-24240.

Endo, F., Awata, H., Katoh, H., and I. Matsuda. 1995. Genomics. **25**: 164-169.

Englard, S., Blanchard, J. S., and C. F. Midlefort. 1985. Gamma-butyrobetaine hydroxylase: stereochemical course of the hydroxylation reaction. Biochemistry **24**: 1110-1116.

Fukuda, H., Ogawa, T., Ishihara, K., Fujii, T., Nagahama, Omata, T., Inoue, Y., Tanase, S., and Y. Morino. 1992. Biochem. Biophys. Res. Commun. **188**: 826-832.

Fukumori, F., and R.P. Hausinger. 1993a. *Alcaligenes eutrophus* JMP134

"2,4-dichlorophenoxyacetic acid monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175**:2083-2086.

Fukumori, F., and R. P. Hausinger. 1993b. Purification and characterization of 2,4-dichlorophenoxyacetate/ α-ketoglutarate dioxygenase. J. Biol. Chem. **268**:24311-24317.

Fuqua, W. C., Coyne, V. E., Stein,D. C., Lin, C. M., and R. M. Weiner. 1991. Characterization of melA: a Gene encoding melanin biosynthesis from the marine bacterium Shewanella colwelliana. Gene **109**: 131-136.

Gershwin, M. E., Coppel, R. L., Bearer, E., Peterson, M. G., Sturgess, A., and I. R. McKay. 1987. Molecular cloning of the liver-specific rat F antigen. J. Immunol. **139**: 3828-3833.

Gomez-Lim, M. A., Valdez-Lopez, V. M., Cruz-Hernandez, A., and J. L. Saucedo-Arias. 1993. Isolation and characterization of a gene involved in ethylene biosynthesis from *Arabadopsis thaliana*. Gene **134**: 217-221.

Hodgson, J. E., Fosberry, A., Rawlinson, N. S., Ross, H. N. M., Neal, R. J., Arnell, J. C., Earl, A. J., and E. J. Lawlor. (GenBank, unpublished).

Holdsworth, M. J., Bird, G. R., Ray, J., Schuch, W., and D. Grierson. 1987a.

Structure and expression of an ethylene -related gene from tomato. Nucleic Acids Res. **15**: 731-739.

Holdsworth, M. J., Schuch, W., and D. Grierson. 1987b. Nucleotide sequence of an ethylene-related gene from tomato. Nucleic Acids Res. **15**: 10600-10600.

Holloway, B. W., 1965. Variations in restriction and modification of bacteriophage following the increase in growth temperature of *Pseudomonas aeruginosa*. Virology **25**: 634-642.

Hopkinson, I., Smith, S. A., Donne, A., Gregory, H., Franklin, T .J., Grant, M. E., J. Rosamond. (unpublished). The complete cDNA sequence of rat prolyl 4-hydroxylase alpha subunit.

Hummel, R., Norgaard, P., Andreasen, P. H., Neve, S., Skjodt, K., Jornehave, D., and K. Kristiansen. 1992. Tetrahymena gene encodes a protein that is homologous with the liver-specific F-antigen and associated with membranes of the Golgi apparatus and transport vesicles. J. Mol. Biol. **228**: 850-861.

Jia, S., Van Dusen, W. J., Diehl, R. E., Kohl, N. E., Dixon, R. A. F., Elliston, K. O. Stern, A. M., and P. A. Friedman. 1992. cDNA cloning and expression of bovine aspartyl (asparaginyl) b-hydroxylase. J. Biol. Chem. **267(20)**: 14322-14327.

Johnson, M., Andrerws, S., Brinkman, R., Cooper, J., Ding, H., Dover, J., Du, Z., Favello, A., Fulton, L., Gattung, S., Geisel, C., Kirsten, J., Kucaba, T., Hiller, L., Jier, M., Johnson, L., Langston, Y., Latreille, P., Louis, E.J., Macri, C., Mardis, E., Menezes, S., Mouser, L., Nhan, M., Rifkin, L., Riles. L., St. Peter, H., Trevaskis, E., Vaughan, K., Vignati, D., Wilcox, L., Wohldman, P., Waterston, R., Wilson, R., and M. Vaudin. 1994. Science. **265**: 2077-2082.

Kim, W. T., Silverstone, A., Yip, W. K., Dong, J. G., and S/ F. Yang. Induction of 1-aminocyclopropane-3-carboxylate synthase mRNA by auxin in mung bean hypocotyls and cultured apple shoots. Plant Physiol. **98**: 1530-1534.

Kovacevic, S., Weigel, B. J., Tobin, M. B., Ingolia, T. D., and J. R. Miller. 1989. Cloning, characterization, and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. J. Bacteriol. **171**: 754-760.

Kovacevic, S., and J. R. Miller. 1991. Cloning and sequencing of the beta-lactam hydroxylase gene (cefF) from Streptomyces clavuligerus: Gene duplication may have led to separate hydroxylase and expandase activities in the Actinomycetes. J. Bacteriol. **173**: 398-400.

Lange, T., Hedden, P., and J. E. Graebe. 1994. Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. Proc. Natl. Acad. Sci. U.S.A. **91**: 8552-8556.

Leskiw, B. K., Aharonowitz, Y., Mevarech, M., Wolfe, S., Vining, L. C., Westlake, D. W., S. E. Jensen. 1988. Cloning and nucleotide sequence determination of the isopenicillin N synthetase gene from Streptomyces clavuligerus. Gene **62**: 187-196.

Lindsteadt, S., and M. Rundgren. 1982. Blue color, metal content, and substrate binding in 4-hydroxyphenylpyruvate dioxygenase from *Pseudomonas sp.* P.J.874. J. Biol. Chem. **257**: 11922-11931.

MacDiarmid, C. W., and R. C. Gardner. 1992. Direct submission of sequence to EMBL/GENBANK/DDBJ Data Banks.

Majamma, R. K., Gunzler, V., Hanauske-Abel, H. M., Myllylä, R., and K. I. Kivirikko. 1986. Partial identity of the 2-oxoglutarate and ascorbate binding sites of prolyl 4-hydroxylase. J. Biol. Chem. **261**:7819-7823.

Marsh, E. N., Chang, M. D. T.,and Townsend,C.A. 1992. Two isozymes of clavaminate synthase central to clavulanic acid formation: cloning and sequencing of both genes from Streptomyces clavuligerus. Biochemistry **31**: 12648-12657.

Matsuda, J., Okabe, S., Hashimoto, T., and Y. Yamada. 1991. Molecular cloning of Hyoscyamine 6-beta Hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of Hyoscyamus niger. J. Biol. Chem. **266**: 9460-9464.

McGarvey, D.J., Yu, H., and R. E. Christoffersen. 1990. Nucleotide sequence of a ripening-related cDNA from avocado. Plant Mol. Physiol. **15**: 165-167.

McGarvey, D. J. and R. E. Christofferson. 1992. Characterization and kinetic parameters of ethylene-forming enzyme from avocado fruit. J. Mol. Biol. **267**: 5964-7.

Meldgaard, M. 1992. Expression of chalcone synthase, dihydroflavanol reductase and flavanone-3-hydroxylase in mutants of barley deficient in anthrocyanidin biosynthesis. Theor. Appl. Genet. **83**: 6 95-706.

Menssen, A., Hoehmann, S., Martin, W., Schnable, P. S., Peterson P. S., Saedler, H., and A. Gierl. 1990. The En/Spm transposible element of *Zea mays* contains a splce at the termini generating a novel intron from a dSpm element in the A2 gene. EMBO J. **9**: 3051-3056.

Milligan, S. B., and C. S. Gasser. 1995. Nature and regulation of pistil-expressed genes in tomato. Plant Mol. Biol. **28(4)**: 691-711.

Myllylä, R., Majammaa, R. K., Gunzler, V., Hanauske-Abel, H. M., and K. Kivirikko. Ascorbate is consumed stoichiometrically in the uncoupled reaction catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. J. Biol. Chem. **259**:5403-5405.

Myllylä, R., Pihlajaniemi, T., Pajunen, L., Turpeenniemi-Hujamen, T., and K. I. Kivirikko. 1991. Molecular cloning of chick lysyl hydroxylase. J. Biol. Chem. **266(5)**: 2805-2810.

Myllylä, R., Gunzler, V., Kivirikko, K., and D. D. Kaska. 1992. Modification of vertabrate prolyl 4-hydroxylases and lysyl hydroxylase by diethyl pyrocarbonate. **286**: 923-927.

Nadeau, J. A., Zhang, X., Nair, H.,and S. D. O'Neill. 1992. Submitted directly to EMBL/GENBANK/DDBJ Data Banks.

Nakanishi, H., Okumura, N., Umehara, Y., Nishizawa, N. K., Chino, M. and S. Mori. 1994. Expression of a gene specific for iron deficiency (Ids3) in the roots of Hordeum vulgare. Plant Cell Physiol. **34**: 401-410.

Ng, S., Hanauske-Abel, H. M., and S. Englard. 1991. Cosubstrate binding site of *Pseudomonas sp.* AK-1 gamma-butyrobetaine hydroxylase. J. Biol. Chem. **266**: 1526-1533.

Okumura, N., Nishizawa, N.K., Umehara, Y., Ohata, T., Nakanishi, H., Yamaguchi, T., Chino, M., and S. Mori. 1994. A dioxygenase gene (Ids2) expressed under iron deficiency conditions in the roots of Hordeum vulgare.Plant Mol. Biol. **25**: 705-719.

Pascal, R. A., Oliver, M. A., and Y. C. Chen. 1985. Alternate substrates and inhibitors of bacterial 4-hydroxyphenylpyruvate dioxygenase. Biochemistry **24**: 3158-3165.

Peck, S. C., Olson, D. C., and H. Kende. 1992. Submitted directly to EMBL/GENBANK/DDBJ Data Banks.

Phillips, A. L., Ward, D. A., Uknes, S., Appleford, N. E. J., Lange, T., Huttly, A. K., Gaskin, P., Graebe, J. E., and P. Hedden. (unpublished). Isolation and expression of three gibberellin 20-oxidase cDNA clones from Arabidopsis.

Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllyla, R., Huhtala, M. L., Koivu, J., and K. I. Kivirikko. 1987. Molecular cloning of the beta-subunit of human prolyl 4-hydroxylase. EMBO J. **6 (3)**: 643-649.

Prescott, A. G. 1993. A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). J. Exp. Bot. **44**:849-861.

Pua, E. C., Sim, G. E., and M L. Chye. 1992. Isolation and sequence

analysis of a cDNA clone encoding ethylene-forming enzyme in *Brassica juncea* (L.) Czern and Coss. Plant Mol. Biol. **19**: 541-544.

Randall, C. R., Zang, Y., True, A. E., Que, L. Jr., Charnock, J. M., Garner, D., Fujishima, Y., Schofield, C. J., and J. E. Baldwin. 1993. X-ray absorption studies of the ferrous active site of isopenicillin N synthase and related model compounds. Biochemistry **32**: 6664-6673.

Roach, P. L., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. L., Hajdu, J., Andersson, I., Schofield, C. J., and J. E. Baldwin. 1995. Crystal structure of isopenicillin N synthase is the forst from a new structural family of enzymes. Nature **375**: 700-704.

Ruetschi, U., Odelhoeg, B., Lindsteadt, S., Barros-Soederling, J., Persson, B., and H. Joernvall. 1992. Characterization of 4-hydroxyphenyl-pyruvate dioxygenase. Primary structure of the *Pseudomonas* enzyme. Eur. J. Biochem. **205**: 459-466.

Ruetschi, U., Dellsen, A., Sahlin, P., Stenman, G., Rymo, L., and S. Lindsteadt. 1993a. Human 4-hydroxyphenylpyruvate dioxygenase: primary structure and chromosomal localization of the gene. Eur. J. Biochem. **213**: 1081-1089.

Ruetschi, U., Nordin, I., Odelhoeg, B., Joernvall, H., and S. Lindsteadt.

1993. Gamma-butyrobetaine hydroxylase Structural characterization of the *Pseudomonas* enzyme. Eur. J. Biochem. **213**: 1075-1080.

Ross, G. S., Knighton, M. L., and M. Lay-Yee. 1992. The ethylene-related cDNA from ripening apples. Plant Mol. Biol. **19**: 231-238.

Samson, S. N., Belagaje, R., Blankenship, D. T., Chapman, J. L., Perry, D., Skatrud, P. L., VanFrank, R. M., Abraham, E. P., Baldwin, J. E., Queener, S. W., and T. D. Ingolia. 1985. Isolation, sequence determination and expression in Escherichia coli of the isopenicillin N synthetase gene from Cephalosporium acremonium. Nature **318**: 191-194.

Samson, S. M., Dotzlaf, J. E., Slisz, M. L., Becker, G. W., Van Frank, R. M., Veal, L. E., Yeh, W. K., Miller, J. R., Queener, S. W., and T. D. Ingolia. 1987. BIO/TECHNOLOGY **5**: 1207-1214.

Sato, S., Suzuki, H., Widyastuti, U., Hotta, Y., and S. Tabata. 1994. Identification and characterization of genes induced during sexual differentiation in *Schizosaccharomyces pombe*. Curr. Genet. **26**: 31-37.

Schofield, J. P., Vijayakumar, R. K., And D. B. G. Oliveira. 1991. Sequences of the mouse F protein alleles and identification of a T-cell epitope. Eur. J. Immunol. **21**: 1235-1240.

Shiffman, D., Mevarech, M., Jensen, S.E., Cohen, G., and Y. Aharanowitz. 1988. Cloning and comparative sequence analysis of the gene coding for isopenicillin N synthase in *Streptomyces*. Mol. Gen. Genet. **214**: 562-569.

Shiffman, D., Cohen, G., Aharonowitz, Y., Von Dohren, H., Kleinkauf, H., and M. Mevarech. 1990. Nucleotide sequence of isopenicillin N synthase (*pcbC*) of the gram negative *Flavobacterium sp.* SC12,154. Nucleic Acids Res. **18**: 660-660.

Sparvoli, F., Martin, C., Sceinza, A., Gavazzi, G., and C. Tonelli. 1994. Cloning and molecular analysis of structural genes involved in flavanoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol, Biol. **24**: 743-755.

Spanau, P., Reinhardt, D., and T. Bollert. 1991. Analysis and cloning of the ethylene-forming enzyme from tomato by functional expression of its mRNA in *Xenopus laevis* oocytes. EMBO J. **10**: 2007-2013.

Tan, D. S. H., and T. Sim. 1996. Functional analysis of conserved histidine residues in *Cephalosporium acremonium* isopenicillin N synthase by site-directed mutagenesis. J. Biol. Chem. **271**: 889-894.

Tang, X., Wang, H., Brandt, A. S., and W. R. Woodson. 1993. Organization

and structure of the 1-aminocyclopropane-3-carboxylate dioxygenase gene family from *Petunia hybrida*. Plant Mol. Biol. **23**: 1151-1164.

Taio-Suan, S., and D. S. H. Tan 1994. Histidine-272 if isopenicillin N synthase of *Cephalosporium acremonium*, which is possibly involved in iron binding, is essential for catalytic activity. FEMS Microbiol. Lett. **120**: 241-248.

Trentmann, S. M., and H. Kende. 1994. Submitted to the EMBL Data Library.

Toguri, T. 1994. Submitted to the EMBL Data Library.

Van der Ploeg, J. R., Weiss, M. A., Saller, A., Nashimoto, H., Saito, N., Kertesz, M. A., and T. Liesinger. 1996. Identification of sulfite starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. J. Bacteriol. **178**: 5438-46.

Wang, H., and W. R. Woodson. 1991. Reversible inhibition of ethylene action and interruption of petal senescence in carnation flowers by nobornadiene. Plant Physiol. **96**: 1000-1001.

Weidler, H., and W. R. Woodson. 1995. Submitted to the EMBL Data Library.

Weigel, B. J., Burgett, S. G., Chen, V. J., Skatrud, P. L., Frolik, C. A., Queener, S. W., and T. D. Ingolia. 1988. Cloning and expression in *Escherichia coli* of isopenicillin N synthase from *Streptomyces lipmanii* and *Aspergillus nidulans*. J. Bacteriol. **170**: 3817-3826.

Wintermeyer, E., Flugel, M., Ott, M., Steinert, M., Rdest, U., Mann,K. H., and J. Hacker. 1994. Sequence determination and mutational analysis of the lly locus of Legionella pneumophila. Infect. Immun. **62(3)**: 1109-1117.

Wu, K., Li, L., Gage, D. A., and J. A. D. Zeevaart. (unpublished). Cloning and photoperiod-regulated expression of gibberellin 20-oxidase from the long-day plant spinach.

Wyckoff, E. E., Pishko, E. J., Kirkland, T. N., and G. T. Cole. 1995. Cloning and expression of a gene encoding a T-cell reactive protein from Coccidioides immitis: homology to 4-hydroxyphenylpyruvate dioxygenase and the mammalian F antigen.Gene **161**: 107-111.

CHAPTER 2.

Use of 4-Nitrophenoxyacetic Acid for the Detection and Quantification of 2,4-Dichlorophenoxyacetic Acid (2,4-D)/α-ketoglutarate Dioxygenase Activity in Natural and Engineered Microorganisms

ABSTRACT

TfdA is the initial enzyme in the canonical 2,4-D degradation pathway, one of the best studied of the chloro-substituted aromatic degradation pathways in bacteria. Purified 2,4-D/α-ketoglutarate dioxygenase (TfdA) was shown to use 4-nitrophenoxyacetic acid (4-NPAA) (K_m = 0.89 ± 0.04 mM, k_{cat} = 540 ± 10 min^{-1}), producing intensely yellow 4-nitrophenol. The generation of this intensely yellow chromophore from 4-NPAA was used to develop a rapid, continuous, colorimetric assay for the detection of TfdA and analogous activities in 2,4-D degrading bacterial cells and extracts. The 4-NPAA assay was found to be suitable for large scale colony screening and direct, quantitative activity measurements, and these methods were shown to offer significant advantages over previous approaches. A diverse collection of environmental 2,4-D degrading strains was screened, revealing significant diversity among 2,4-D degrading isolates.

INTRODUCTION

The canonical 2,4-dicholrophenoxyacetic acid (2,4-D) degradation pathway is that of *Alcaligenes eutrophus* JMP134 (pJP4), which uses six enzymes to convert 2,4-D degradation to TCA cycle intermediates (Don, 1985; Perkins, 1990; Streber, 1987). The first enzyme in the pathway, 2,4-D/α-ketoglutarate dioxygenase (TfdA), converts 2,4-D, α-ketoglutarate, and oxygen to 2,4-dichlorophenol, glyoxylate, carbon dioxide, and succinate (Fukumori, 1993a). Analysis of TfdA-like activities in 2,4-D degrading microorganisms requires the availability of a functional assay for the enzyme. Several methods to assay TfdA-like activities have been described, but all have limitations. HPLC (Fukumori, 1993a) and GC (Perkins, 1988) methods can be used to monitor 2,4-D disappearance or 2,4-dichlorophenol production, but these methods are time consuming and require sophisticated equipment. Oxygen electrodes can be used for continuous assay of the activity (Streber, 1987), but this approach is confounded by the presence of other oxygenases and requires specialized equipment. Radioactive methods to measure the release of $^{14}CO_2$ from labeled substrate (Fukumori, 1993b; Fulthorpe, 1996) are known, but these approaches require special handling. A discontinuous spectrophotometric assay for substituted phenols that uses 4-aminoantipyrene to form a highly colored complex has been used for the *in vitro* detection of substituted phenols (Fukumori, 1993b; King, 1991), and was used in a petri plate assay to detect cells with overexpressed or deregulated TfdA systems (King, 1991). Unfortunately, the required high

pH is lethal to the cells, and reagent does not work well for regulated pathways in the plate assay (King, 1991). Here, we evaluated 4-nitrophenoxyacetic acid (4-NPAA, Tokyo Kasai, Tokyo, Japan) as a sensitive, continuous spectrophotometric assay for TfdA.

RESULTS AND DISCUSSION

Using purified *Alcaligenes eutrophus* JMP134 TfdA and the kinetics methods as previously reported (Fukumori, 1993b), 4-NPAA was shown to be a substrate (K_m = 0.89 ± 0.04 mM, k_{cat} 540 ± 10 min^{-1} , k_{cat}/K_m = 610 mM^{-1} min^{-1}). When compared to published values for the decomposition of 2,4-D by this enzyme (K_m = 17.5 ± 1.0 μM, k_{cat} = 529 min^{-1} (Fukumori, 1993b)), it is clear that the affinity for 4-NPAA is 50-fold less, the turnover number is unchanged, and the catalytic efficiency is approximately 2.0% of that for 2,4-D. One product of 4-NPAA decomposition was demonstrated to be 4-nitrophenol (4-NP) on the basis of UV-visible spectroscopy and HPLC analysis using a Hewlett Packard Series 1050 HPLC unit with a Hibar Lichrosorb RP-18 column and a 60% methanol and 40% 20 mM KH_2PO_4 buffer at pH 3.

We sought to use the intensely yellow (18.4 mM^{-1} cm^{-1}) 4-nitrophenol chromophore as a method to detect whole cell TfdA-like activities in a set of control and engineered strains (Table 2.1) and in a collection of 2,4-D degrading strains (Table 2.2) using a plate assay. Colonies were carefully wetted with freshly prepared assay solution (10 mM 4-NPAA, 50 μM $Fe(NH_4)_2(SO_4)_2$, and 50 μM ascorbate in 20 mM Tris-HCl, pH 7.4), allowed to dry, and rewetted using an atomizer. Plates were incubated at 30°C with intermittent observation for 5 minutes to 6 hours. Inclusion of α-ketoglutarate in the assay solution did not improve the effectiveness of the assay with most strains. Colonies of *A. eutrophus*, *Burkholderia cepacia* DB01 (pRO101), and numerous environmental isolates capable of

Table 2.1 Control and engineered strains and their response to the 4-NPAA assays.

Host	Strain	Plasmid[a]	Reference[b]	Growth Media[c] / Relevant Genotype	4-NPAA plate[d]	conversion lysate[e]
Alcaligenes eutrophus	JMP134	pJP4	D	M2 / *tfdA*$^+$	+	0.98 ± 0.0420
Alcaligenes eutrophus	JMP228r	none	D	MC2 / nal^r	-	-
Burkholderia cepacia	DB01	none	H	MC2	-	-
Burkholderia cepacia	DB01	pRO101	H	M2 / Tc^r, *tfdA*$^+$	+	3.5 ± 0.176
Escherichia coli	CB101	none	B	MC2	-	-
Escherichia coli	CB1811	pKJS31 (Km^r, Sm^r)	S	MC2 / *tfdA*$^+$, tfdS$^+$	-	2.6 ± 0.119
Escherichia coli	CB1832	pMMB510 (Sm^r)	This work.	MCI / *tfdA*$^+$, unexpressed	-	-
Escherichia coli	CB1833	pMMB511 (Sm^r)	This work.	MCI / *tfdA*$^+$, overexpressed	-	1.8 ± 0.517

[a] Plasmid pKJS31 is pKT235 (3) with a 3.1 kb *Sac*I fragment of pJP4 (13, containing a functional *tfdA* gene and a complete and probably functional *tfdS* gene) inserted into the *Sac*I site. Plasmid pMMB510 is pMMB503EH(Sm^r) (10) with the *SacI*/*Xba*I fragment of pKJS31 (in an orientation opposite to the plasmid promoter) inserted into the *SacI*/*Xba*I site of the polylinker, with pMMB511 containing the same insertion in pMMB503HE(Sm^r) (10) in the correct orientation to be expressed from the plasmid promoter. Plasmid pRO101 is pJP4:Tn1721 (Tc^r), with no effect on the known genes for 2,4-D degradation (8).

[b] B is Bachmann, 1972; D is Don, 1985; H is Harker, 1989; and S is Streber, 1987.

[c] Minimal salts medium (7) was supplemented with 1.5 mM 2,4-D (termed M2), with 0.2% vitamin free Casamino acids (termed MC), or both (termed MC2). MCI medium is MC with 100 μM IPTG. Kanamycin (Km) and streptomycin (Sm) were used at 100 μg/μl, naladixic acid (Nx) at 50 μg/μl, and tetracycline (Tc) at 25 μg/μl.

[d] Visible production of 4-NP from 1.5 mM 4-NPAA by 2,4-D induced whole cells on agar plate.

[e] Spectrophotometrically quantifiable 4-NP production using the cytosolic fraction of cell lysates in the lysate assay, expressed in μM ml $^{-1}$min^{-1} mg^{-1} protein.

Table 2.2 4-NPAA conversion to 4-NP by a diverse set of natural isolates that degrade 2,4-D[a]

Strains	Species Identification	Hybridization pattern group	Reference[b]	*tfdA* hybridization[c]	4-NPAA conversion	
					plate[d]	lysates[e]
JMP134	*Alcaligenes eutrophus*	1	D	+	+	+
EML159	*Burkholderia* sp.	1	T	+	+	+
TFD39	*Burkholderia* sp.	1	T	+	slow	weak
TFD33*, 38, 41	*Alcaligenes eutrophus*	3	T	+	+	+
B6-9	*Rhodoferax fermentans*	6	F	+	+	+
TFD6, 13*, 14, 15, 16*, 17, 18, 20, 21, 22*, 25*, 27, 34*	*Burkholderia mallei*	7	T	+	+	+
TFD2, 4, 5*, 7, 8, 36	*Burkholderia* sp.	9	T	+	+	+
RASC	*Burkholderia* sp.	10	F,T	+	+	+
M1	*Nitrobacter winogradskii*	11	F	-	-	+[f]
EML146	*Sphingomonas* sp.	12	F	-	slow	-
TFD44	*Sphingomonas* sp.	13	F	-	+	-
K1443	*Sphingomonas* sp.	14	F	-	+	+[f]
B6-5, B6-7a*	*Sphingomonas* sp.[g]	15	F	-	-	-
TFD19, 40	ND	ND	T	ND	slow	weak
TFD12, 32, 35, 45	ND	ND	T	ND	slow	-
TFD28	ND	ND	T	ND	-	-
TFD1, 11	ND[h]	ND	T	-	-	-

Table 2.2 continued.

[a] Closest species identifications were previously determined in many cases by partial 16S rRNA sequence alignment to GENBANK sequences (8), with strain M1 identity clarified in this work by subsequent alignment to additional sequences and Neighbor -Joining phylogenetic analysis. Hybridization groupings, where defined, were previously determined by hybridization band pattern using *tfdA*, *tfdB*, and *tfdC* gene probes, with conditions set to exhibit hybridization at >60% identity (8). For strains marked with an * inclusion was based on REP patterns (8) or on responses to the 4-NPAA assays (this work).

[b] Reference codes: D is Don, 1985; F is Fulthorpe, 1995; and T is Tonso, 1995.

[c] Southern hybridization under low stringency conditions using an *Alcaligenes eutrophus tfdA* as a probe (6). (+) denotes positive hybridization, ND denotes not determined.

[d] Visibly detectable conversion of 4-NPAA to 4-NP in liquid cultures containing minimal medium, 0.5 mM 2,4-D and 0.5 mM 4-NPAA. (+) denotes color detection within 2 hours, (slow) denotes color detected within 6 hours, (-) denotes no detectable reaction after 24 hours.

[e] Strains that degraded 2,4-D were tested for 4-NPAA conversion in cell lysates. Qualitative data was provided for clarity using the general survey parameters noted. Precise growth conditions for maximum induction of TfdA-like enzymes varied widely among natural isolates. (+) denotes quantifiable conversion within 2-5 minutes, (weak) denotes reaction only detectable after several hours, (-) denotes no reaction within 24 hours. Lysates which did not convert 4-NPAA were retested under varying growth conditions.

[f] 4-NPAA conversion was at least partially present in pellet fraction of cell lysates.

[g] Denotes probable co-metabolizers.

[h] Denotes Gram positive isolates.

degrading 2,4-D were found to turn yellow and produced a yellow halo of diffusing 4-NP around the colony (Fig. 2.1). The halo was particularly useful for identification of activity with pigmented colonies. Importantly, however, not all 2,4-D degraders presented a yellow halo and selected engineered strains known to express *tfdA* remained as colorless as the controls lacking *tfdA* expression (Table 2.1). Similar results were obtained in batch culture or microtiter plate assays, where cultures were spiked with 1.5 mM 4-NPAA when the cultures were in late log to early stationary phase and incubated for 5 minutes to six hours at 30°C. Samples of the cultures, or entire microtiter plates, were centrifuged to pellet the cells, supernatants were transferred to cuvettes or fresh microtiter plates where 1/10 volumes of 0.1 N NaOH (final pH 11) were added, and the absorbances at 401 nm were measured. As with the plate assay, not all of the cells known to express *tfdA* or degrade 2,4-D produced a yellow color.

The lack of color production in selected whole cells is likely due to the inability of these cultures to transport 4-NPAA through the cell envelope. For example, when *E. coli* CB1833 or *E. coli* CB1811 cells (negative in whole cell assays, yet known to express *tfdA* under the *tac* promoter) were sonicated, centrifuged for 30 minutes at 100,000 x *g*, and assayed spectrophotometrically at 401 nm, the production of 4-NP was observed (Table 2.1). Lysates of *Alcaligenes eutrophus* JMP134 that were treated similarly also converted 4-NPAA to 4-NP (Figure 2.2), and converted this substrate in a quantitative manner (Figure 2.3). Lysate reaction mixtures include all of the known co-factors of TfdA - a final concentration of 1 mM α-ketoglutarate, 50 μM ascorbate, 50 μM

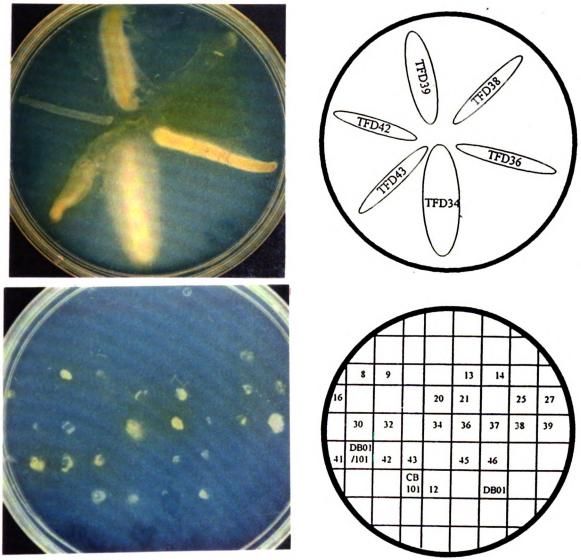


Figure 2.1. The 4-NPAA petri plate assay. Panels on the left are color photos of assay plates after treatment and incubation. The right panes denote the strains depicted in the photos. The top pair of panels shows environmental strains that have different levels of growth and different rates of 4-NP production after an incubation time of four hours. The strains shown in the bottom pair of panels are engineered strains (positive control *Burkholderia cepacia* DB01 containing pRO101) and environmental isolates (numbers indicate TFD strain designations). The negative controls, *E. coli* CB101 and *Burkholderia cepacia* DB01, are in the bottom row. Incubation time was six hours.

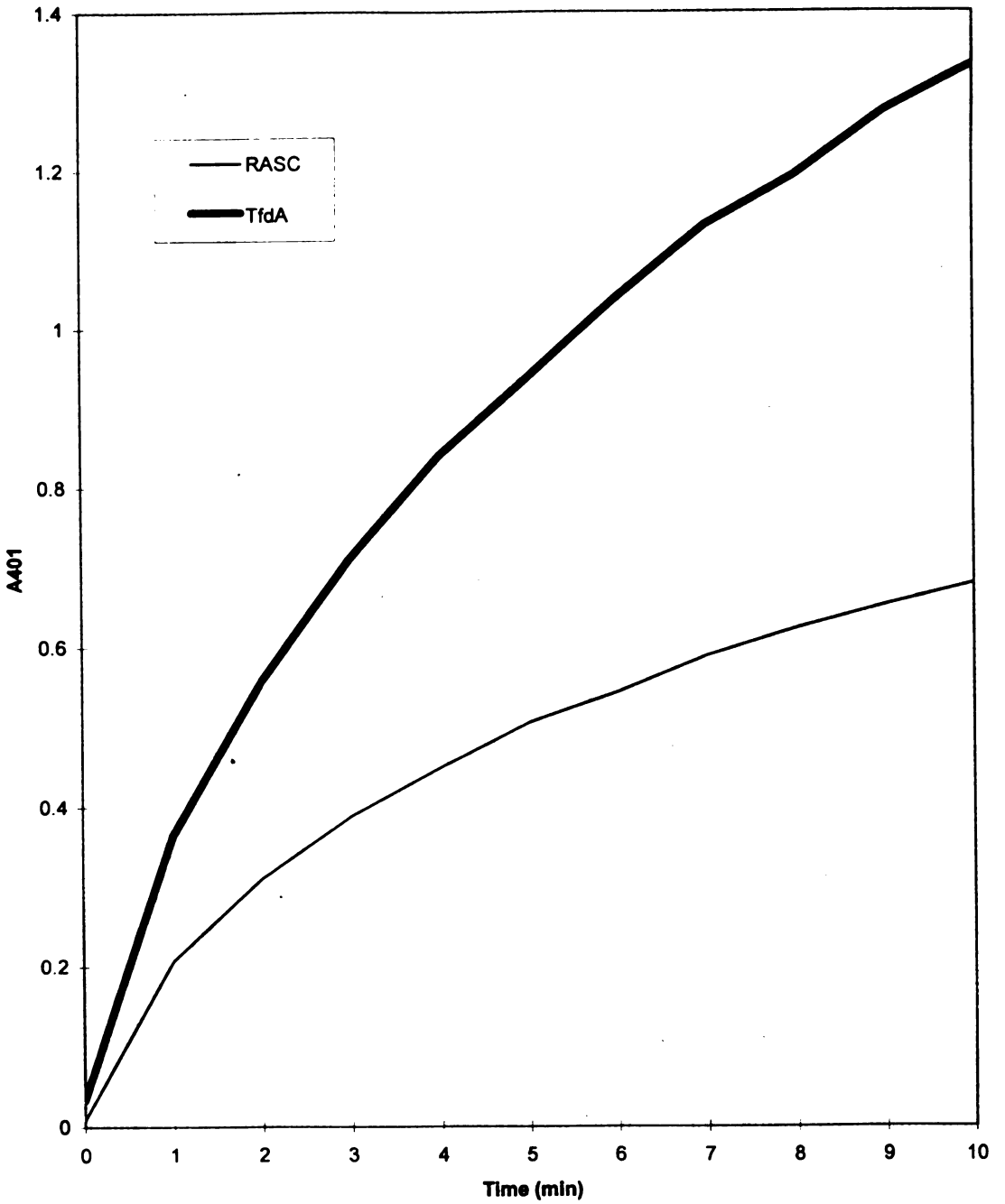
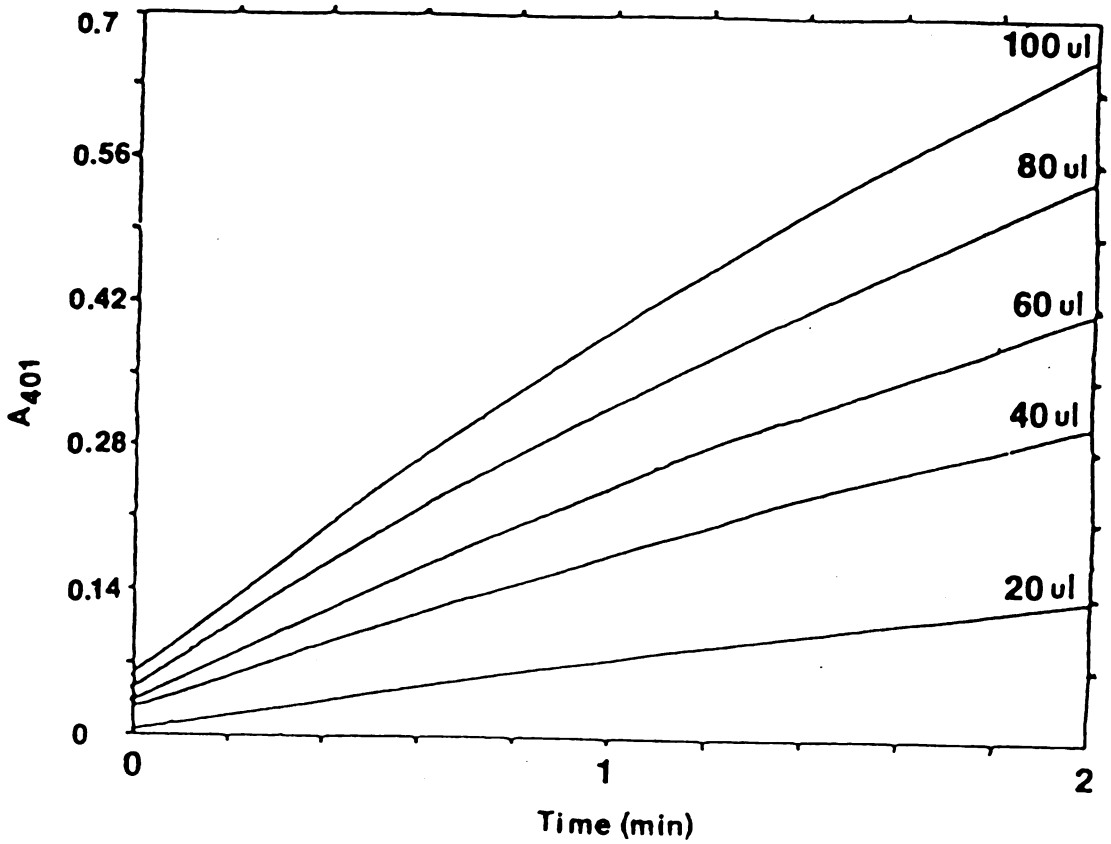


Figure 2.2 4-NPAA cell lysate assay using *Alcaligenes eutrophus* JMP134 and *Burkholderia cepacia* RASC. Continuous assay of 4-nitrophenol produced by TfdA was measured spectrophotometrically by absorbance at 401 nm. TfdA from *Alcaligenes* and $TfdA_{rasc}$ from *Burkholderia* are 91% identical at the amino acid level.



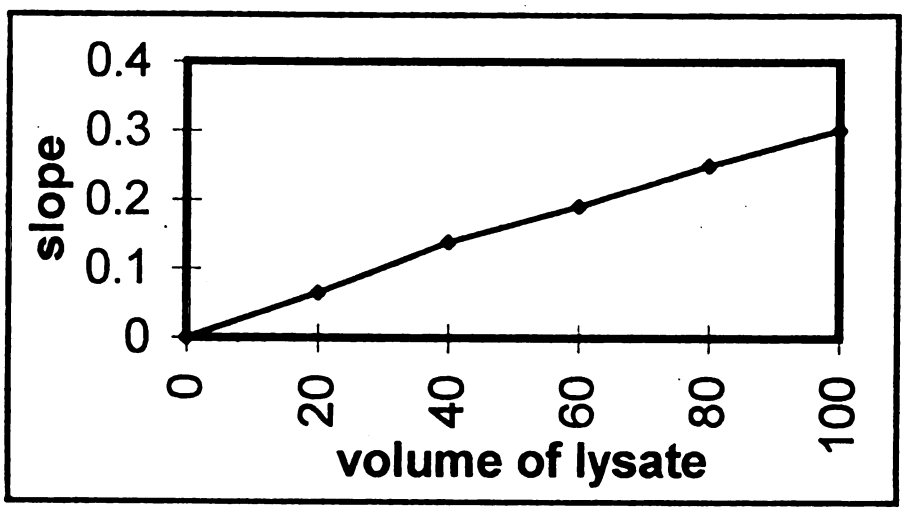


Figure 2.3. 4-NPAA cell lysate assay using *Alcaligenes eutrophus* JMP134 indicating 4-NPAA assay is quantitative. Top figure: continuous assay of 4-nitrophenol produced by TfdA was measured spectrophotometrically by absorbance at 401 nm. Bottom figure shows the linearity of the increase of 4-NP production over time versus the amount of enzyme used in the assay.

$Fe(NH_4)_2(SO_4)_2$ - and 1.5 mM 4-NPAA in 20 mM Tris-HCl, pH 7.4. Reaction mixtures varied in protein content, and the resulting 4-NP production was normalized to cell lysate protein content using Lowry assay determination (Lowry, 1951). Inclusion of a 0.1 volume 0.1 N NaOH quench step in the assay increased its sensitivity by converting all of the released 4-NP to its deprotonated form ($pK_a = 7.15$); however, the assay can not be used in a continuous manner at high pH values (Fukumori, 1993b). Among the six environmental isolates that were negative in whole cell assays, only one, *Nitrobacter winogradskii* strain M1 (Fulthorpe, 1995), was positive in the lysate assay (Table 2.2). Of note, much of the activity in this strain was found in the particulate fraction. The lack of activity in whole cells and presence of activity in lysates is taken as evidence of a transport deficiency. Attempts to permeabilize *E. coli* cells using chloroform or toluene vapor, or by treatment with Triton X-100 at various concentrations in conjunction with this assay, were ineffective or completely eliminated TfdA activity.

4-NPAA can also be used effectively with many isolates to assay for TfdA-like activity in batch cultures. Cultures of 2,4-D degrading isolates can be tested for the induction of *tfdA* by including 4-NPAA in the media, using 2,4-D as the growth substrate (Appendix B). Since induction of the 2,4-D pathway varies widely among isolates, testing a sample of a culture for induction by adding 100 mM 4-NPAA is an excellent way to ensure the presence of pathway enzymes in cell lysates. Spiking a sample of the large culture is advisable, since the presence of 4-NPAA or 4-NP in the media has widely varying effects on the growth of different species. This technique also allows the effective use of supplementary carbon sources to increase

the cell mass of strains which do not grow to sufficient cell density for cell lysate preparation. Casamino acids were the best general carbon supplementation, with sugars and yeast extracts having an inhibitory effect on the induction of the 2,4-D pathway of several strains including *Alcaligenes eutrophus* JMP134. 4-NPAA was not a gratuitous inducer of the 2,4-D pathway in any of the strains tested, and none of the strains in this study could utilize 4-NP as a carbon source (Appendix B).

The 4-NPAA whole cell and cell lysate assay results for the environmental isolates demonstrate that caution must be used in making interpretations based on this substrate. Five 2,4-D degrading isolates (B6-5, B6-7 and TFD1, TFD11, and TFD28 (Fulthorpe, 1995; Tonso, 1995)) were negative for 4-NPAA conversion both as whole cells and in extracts (Table 2). These results are consistent with the presence of a 2,4-D degrading enzyme that fails to recognize 4-NPAA or fails to convert 4-NPAA under conditions used in this assay. Lysates of these strains also failed to convert 2,4-D to 2,4-dichlorophenol. Several other strains (EML146, and TFD12, TFD32, TFD35, TFD44, and TFD45 (Fulthorpe, 1995; Tonso, 1995)) were able to degrade 4-NPAA as whole cells, but exhibited no activity in extracts. This outcome could suggest the presence of a highly labile enzyme or one that metabolizes 2,4-D by an alternate mechanism not requiring α-ketoglutarate. Nevertheless, for most environmental isolates, 4-NPAA can serve as a useful substrate to allow the detection and continuous spectrophotometric assay of TfdA activity in whole cell enriched fractions or cell lysates.

SUMMARY

Purified 2,4-D/α-ketoglutarate dioxygenase (TfdA) was shown to use 4-NPAA at about 0.2% the rate of 2,4-D, producing intensely yellow 4-nitrophenol. This reagent was used to develop a rapid, continuous, colorimetric assay for the detection of TfdA and analogous activities in 2,4-D degrading bacterial cells and extracts. The assay was developed for whole cells in liquid, cell lysates, and petri plates, can be used in a qualitative or quantitative manner, and may be used in a continuous manner unlike many of the other assays. This study indicates that there is a wide diversity among environmental isolates in 4-NPAA conversion. All strains that hybridize to a *tfdA* probe convert 4-NPAA to 4-NP readily, though the assay can also detect TfdA-like activity in strains that do not hybridize to *tfdA*. None of the strains tested can utilize 4-NPAA in the absence of 2,4-D, nor can they utilize 4-NP as a carbon source in the presence or absence of 2,4-D. Growth in a media that contains both 2,4-D and 4-NPAA has widely varying deleterious effects on the growth of various environmental strains. The assay also provides indication that there is a wide diversity of ability to take up 2,4-D and 4-NPAA among 2,4-D degrading environmental isolates.

REFERENCES

Bachmann, B. J. 1972. Pedigrees of some mutant strains of *Escherichia coli* K12. Bacteriol. Rev. **36**:525-557.

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

Frey, J., M. Bagdasarian, and K. N. Timmis. 1984. *Escherichia coli* cloning vectors that auto-amplify in the stationary phase of bacterial growth, p.377-383. *In* S. Mitsuhashi and V. Krčméry (ed.), Transferable Antibiotic Resistance: Plasmids and Gene Manipulation. Avericum Praha and Springer, Heidelberg.

Fukumori, F., and R.P. Hausinger. 1993a. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetic acid monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175**:2083-2086.

Fukumori, F., and R. P. Hausinger. 1993b. Purification and characterization of 2,4-dichlorophenoxyacetate/ α-ketoglutarate dioxygenase. J. Biol. Chem. **268**:24311-24317.

Fulthorpe, R.R., C. McGowan, O.V. Maltseva, W.E. Holben, and J.M. Tiedje. 1995. 2,4-dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. **61**: 3274-3281.

Fulthorpe, R.R., A. N. Rhodes, and J. M. Tiedje. 1996. Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via

different microbial populations. Appl. Environ. Microbiol. **62**:1159-1166.

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

King, R. J., K. A. Short, and R. J. Seidler. 1991. Assay for detection and enumeration of genetically engineered microorganisms which is based on the activity of a deregulated 2,4-dichlorophenoxyacetate monooxygenase. Appl. Environ. Microbiol. **57**:1790-1792.

Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**: 265-275.

Overbye Michel, L. M. Sandkvist, and M. Bagdasarian. 1995. Specificity of the protein secretory apparatus: secretion of the heat-labile enterotoxin B subunit pentamers by different species of Gram$^-$ bacteria. **152**:41-45.

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

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

Streber, W.R., K.N. Timmis, and M.H. Zenk. 1987. Analysis, cloning, and high level expression of 2,4-dichlorophenoxyacetic acid

monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169**:2950-2955.

Tonso, N.L., V.G. Matheson, and W.E. Holben. 1995. Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microbial Ecol. **30**:3-24.

Chapter 3:

2,4-D Degrading Strain Characterization and the Conjugal Capture of a Plasmid from *Nitrobacter winogradskii* M1 That Rescues 2,4-Dichlorophenoxyacetic Acid Degradation in a TfdA$^-$ *Alcaligenes eutrophus* Strain.

INTRODUCTION

As part of the effort to explore the evolution of catabolic pathways by closely examining 2,4-D degradation, a phylogenetically diverse collection of Gram-negative isolates that could degrade 2,4-D as a sole source of carbon and energy was examined. This work focused on ascertaining whether the collection contained alternate degradation pathways or enzymes to the canonical 2,4-D degradation pathway of pJP4, examining the diversity of genes among the pathways found, and perhaps identifying non-homologous isofunctional enzymes. Since chlorocatechol cleavage pathways are common in soil bacteria (Fulthorpe, 1996) and phenol hydroxylases are also fairly common, the search for a non-homologous isofunctional enzyme became restricted to the initial enzymatic reaction of 2,4-D degradation. TfdA, the initial enzyme of the canonical pathway from *Alcaligenes eutrophus* JMP134 (pJP4), was reclassified as 2,4-D/α-ketoglutarate dioxygenase, rather than a monooxygenase as originally reported (Fukumori, 1993). This enzyme converts 2,4-D and α-ketoglutarate to 2,4-dichlorophenol, glyoxylate, carbon dioxide, and succinate (Fukumori,

1993). In nature, *tfdA* is thought to be comparatively rare (Fulthorpe, 1996), and has been found on plasmids independent from the other genes of the 2,4-D pathway (Top, 1996).

Extensive strain characterization at the genomic and biochemical level - including FAME and BIOLOG analysis (Tonso, 1995), REP pattern, hybridization to probes to the *tfdA*, *tfdB*, and *tfdC* genes (Fulthorpe, 1995), and tentative identification based on partial 16S rDNA analysis (McGowan, 1994) - was performed by a number of workers (Fulthorpe, 1996; Suwa, 1996). Some of the 2,4-D degrading isolates failed to hybridize with a *tfdA* specific probe (Table 2.1 and 2.2) and failed to amplify during PCR using *tfdA* specific primers, indicating that there may be significant diversity among the genes that have not been examined. Four α-proteobacterial strains showed the least similarity to the canonical pathway and were studied in hopes of finding at least one strain that uses an evolutionarily distinct TfdA-like enzyme. Preliminary work that suggests that among the four strains selected, there is at least one strain, *Nitrobacter winogradskii* M1, that carries a plasmid borne TfdA-like enzyme that is functionally similar but has distinct biochemical properties from the well studied TfdA from *Alcaligenes eutrophus* JMP134. Further work will be required to determine if this enzyme is evolutionarily unrelated to TfdA.

RESULTS AND DISCUSSION:

General comments. The strains and plasmids used in this chapter are found in Table 3.1. Detailed methods are provided in Appendix C.

Phylogenetic identification of the a-proteobacterial strains. Most of the strains in the diverse collection of 2,4-D degrading strains compiled for study by the NSF Center for Microbial Ecology are members of the β-proteobacterial subdivision (McGowan, 1995) and many contain genes that are highly homologous to *tfdA* (Fulthorpe, 1995; Tonso, 1995). None of the four isolates that were classified as α-proteobacteria hybridized with the *tfdA* probe, making them candidates for possible alternate TfdA-like enzymes. Three of the strains (TFD44, EML146, K1443) were identified as *Sphingomonas* species. The genus *Sphingomonas* is a recent reclassification division of the α–proteobacteria that may still be polyphyletic (Takeuchi, 1994), a concept supported by the variable characteristics of the three strains in this study (Table 3.2 and 3.3). The initial identification of the fourth strain, M1, indicated it was *Rhodopseudomonas palustris*, but was identified with the qualification that it was atypically non-pigmented and therefore may not be correct (McGowan, 1995). A subsequent examination of the partial 16S rDNA sequence of strain M1 using multiple sequence alignment and neighbor-joining (NJ) phylogenetic analysis resulted in the identification of M1 as *Nitrobacter winogradskii* (Figure 3.1). M1 clustered with a bootstrap of 84 in a clade containing *N. winogradskii, N. hamburgensis, Bradyrhizobium japonicum, Blastobacter denitrificans, and R. palustris.* Within that clade, *Nitrobacter winogradskii* and strain M1 clustered with a bootstrap value of 94, indicating a clustering with relative

Table 3.1. Strains and plasmids used to investigate novel *tfdA*-like activities.

Strain	Plasmid/Relevant Genotype[a]	Growth Media[b]	Reference
Alcaligenes eutrophus BH501	pBH501 (TfdA$^-$, Km^r)	LB + Km	Top, 1995
A. eutrophus JMP134	pJP4 (Tfd$^+$)	M2	Don, 1981
A. eutrophus JMP228r	none (nal^r)	LB + Nx	Don, 1981.
A. eutrophus P1 - P6	pBH501 and pM1.1 (Tfd$^+$)	M2	This work
Burkholderia cepacia DB01	none	LB	Kukor, 1989
B. cepacia DB01	pRO101 (Tc^r, *tfdA*$^+$)	M2	Kukor, 1989
B. cepacia RASC	none, (Tfd$^+$)	M2	Suwa, 1996
B. cepacia TFD6	pTFD6 (*tfdA*$^+$), (Tfd$^+$)	M2	Matheson, 1996
Nitrobacter winogradskyi M1	pM1.1 (*tfdA*-like$^+$), pM2.2	M2	Fulthorpe, 1996
P. aeruginosa PB2036	none	LB	Holloway, B., 1981
P. putida KT2442	none	LB	Bagdasarian, 1991
Sphingomonas sp. K1443	multiple, (Tfd$^+$)	M2	Ka, 1994
Sphingomonas sp. TFD44	multiple, (Tfd$^+$)	M2	Tonso, 1995
Sphingomonas sp. EML146	multiple, (Tfd$^+$)	M2	Amy, 1985
Escherichia coli CB1833	pMMB511 (Sm^r, *tfdA*$^+$)	LB + Sm	Sassanella, 1997
E. coli DH5α	none, pRL1062a (Cm^r)	LB	Cohen, in press
E. coli HB101	none	LB	Bachmann, 1972
E. coli HB101	none, ColE1:Tn3	LB	Bagdasarian, 1997
E. coli HB101	none, ColE1:Tn7	LB	Bagdasarian, 1997

Table 3.1. continued.

Strain	Plasmid/Relevant Genotype[a]	Growth Media[b]	Reference
E. coli HB101	none, pMMB207 (Cm^r)	LB + Cm	Morales, 1991
E. coli HB101	none, pMMB503	LB + Sm	Overbye, 1996
E. coli HB101	none, pRK2013(Km^r), Sm^r	LB	Figurski, 1979
E. coli HB101	none, pTJS75a (Tc^r)	LB + Tc	Schmidhauser, 1985
E. coli S-17	*tra*$^+$, pSUP101[Tn1]	LB + various	Simon, 1983
E. coli S-17	*tra*$^+$, pSUP202[Tn5]	LB + various	Simon, 1983
E. coli S-17	*tra*$^+$, pSUP203[Tn501]	LB + various	Simon, 1983
E. coli S-17	*tra*$^+$, pSUP2017[Tn7]	LB + various	Simon, 1983
E. coli SM10	λ *pir*, pUT(mini-Tn5(*lacZ1*))	LB	deLorenzo, 1990
E. coli SM10	λ *pir*, pUT(mini-Tn5(*lacZ2*))	LB	deLorenzo, 1990

[a] Plasmids are listed first, followed by plasmid characteristics in parenthesis. A comma separates plasmid characteristics from chromosomal or strain characteristics.

[b] Media abbreviations are described in the abbreviations section. Kanamycin and streptomycin are used at 100 μg/ml, naladixic acid at 50 μg/ml, chloramphenicol at 25 μg/ml, and tetracycline at 25 μg/ml.

Table 3.2 Carbon source utilization of four β-proteobacterial 2,4-D degrading strains and several strains used for genetic manipulation.

Carbon Source[a]	DB01	EML146	JMP228	K1443	M1	TFD44
Acetate	+[b]	+ (y)[c]	-	-	Slow	+ (w)
α-ketoglutarate	+	+ (c)	+	-	+	+ (w)
Citrate	+	-	+	-	-	+ (w)
Glucose, 1.25%	+	+ (o-y)	Slow	+ (w-y)	+ (y)	+ (y)
Maltose, 1%	+	+ (c)	Slow	Slow (c)	Slow (c)	+ (w)
Sucrose, 1.25%	+	+	Slow	-	-	+ (w)
Glycerol, 1.25%	+	+	Slow	-	-	+ (w)
Benzoate, 40 mM	+	+	+	-	-	+
2-chlorobenzoate	-	-	-	-	-	-
3-chlorobenzoate	-	Slow (c)	-	-	-	+ (c)
4-chlorobenzoate	-	+	-	-	-	-
Phenoxyacetic acid	-	-	-	-	-	+ (c)
2-CPAA	-	-	-	-	-	+ (w)
4-CPAA	-	+ (c)	-	-	-	+ (c)
2,4-D	-	+ (o-y)	-	+ (w-y)	+ (w)	+ (w)
4-NPAA	-	-	-	-	-	-
2,4-DCP	-	+ (c-w)	-	+ (c-w)	+ (c-w)	+ (c-w)

Table 3.2. Contiued.

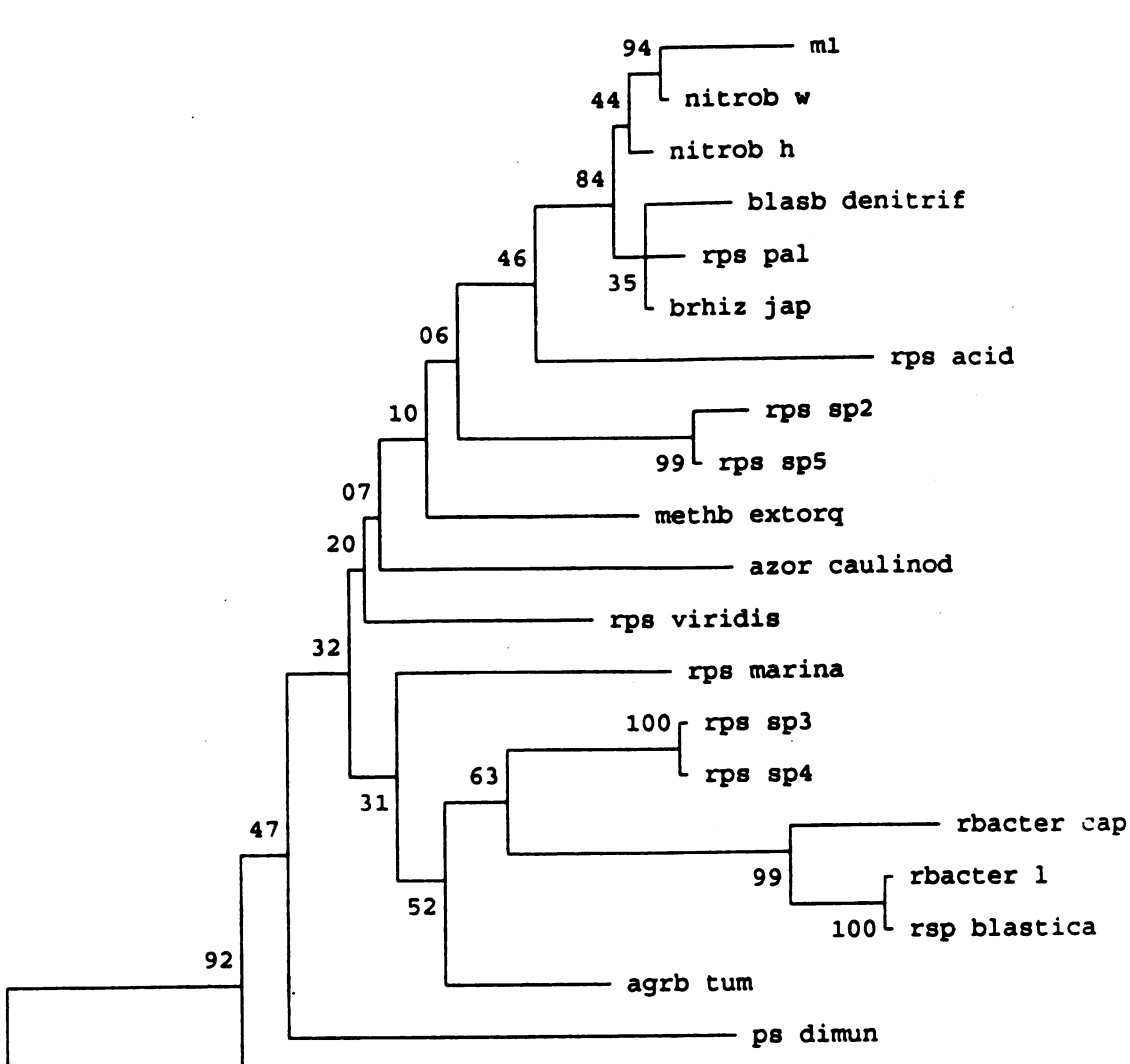
4-NP	-	-	-	-	-	-
Casamino acids	+	+	+	+	+	+

[a] Unless otherwise noted, the media used was minimal salts (Fulthorpe, 1995) and 1.6% noble agar supplemented with 1 mM of the noted carbon source.

[b] Growth was recorded as similar to that on 1/10 LB (+), distinguishably less than that on 1/10 LB, and absence of growth (-) after a 96 hour incubation at 30° C.

[c] Letters in parenthesis note a change in color of the colony on alternate types of media, where (c) represents a change to clear colonies, (o) orange colonies, (w) white colonies, and (y) yellow colonies. On diluted complex media (1/10 Luria-Bertani), all of the strains are cream colored except EML146, which is slightly orange, and TFD44, which is slightly yellow.

Figure 3.1. Phylogenetic tree representing the neighbor-joining analysis of the 16S rDNA sequence of strain M1 and close relatives. This analysis supports the identification of M1 as *Nitrobacter winogradskii* (bootstrap = 94). This analysis was completed using MEGA (Kamar, 1993), with the sequence identification and alignment as noted in Figure 3.3. Numbers at the nodes of the tree represent bootstrap values using Tamura-Nei distance estimation, complete removal of identical sites, and 500 replications. Sequence abbreviations are as follows: m1- strain M1; nitrob w - *N. winogradskii*; nitrob h - *N. hamburgensis*; blasb denitrif - *Blastobacter denitrificans*; rps pal - *Rhodopseudomonas palustris*; brhiz jap - *Bradyrhizobium japonicum*; rps acid - *R. acidophila*; rps sp2 and rps sp5 - *Rhodopseudomonas* sp. SP2 and SP3; methb extorq - *Methylobacterium extorquens*; azor caulinod - *Azorhizobium caulinodans*; rps viridis - *R. viridis*; rps marina - *R. marina*; rps sp3and rps sp4 - *Rhodopseudomonas* sp. SP3 and SP4; rbacter cap - *Rhodobacter capsulatus*; rbacter 1 - *Rhodobacter euryhalinus*; rps blastica - *R. blastica*; agrb tum - *Agrobacterium tumefaciens*; ps dim - *Pseudomonas diminuta*; rspir rubrum - *Rhodospirillum rubrum*; rdphila globif - *Rhodophila globiformis*; and e coli - *Escherichia coli.*



Scale: each – is approximately equal to the distance of 0.002205

confidence that clearly distinguishes between the closely related sequences. The NJ tree was supported by subsequent testing using the same alignment and maximum parsimony (MP) analysis (Felsenstein, 1981). MP analysis produced 10 most parsimonius trees, all of which supported the clustering of M1 and *N. winogradskii*.

Strain characterizations. *Whole cell assays.* The α-proteobacterial 2,4-D-degrading strains were tested for growth on various carbon sources: acetate, α-ketoglutarate, citrate, glucose, maltose, sucrose, glycerol, benzoate, 2-chlorobenzoate, 3-chlorobenzoate, 4-chlorobenzoate, phenoxyacetic acid (PAA), 2-chlorophenoxyacetic acid (2-CPAA), 4-chlorophenoxyacetic acid (4-CPAA), 2,4-D, 4-NPAA, 2,4-DCP, 4-NP, and casamino acids (Table 3.2). All of the strains were able to utilize casamino acids as a growth substrate and grew well in diluted complex media (1/10 LB). Most were capable of utilizing TCA cycle intermediates as carbon sources. Growth on the other carbon sources varied. Among the α-proteobacteria, strain M1 and strain K1443 had the most similarity in carbon source utilization, particularly noting their inability to degrade all phenoxyacetic acid compounds other than 2,4-D. The other two strains of this group, TFD44 and EML146, have a distinct pattern from one another as well as from M1 and K1443. TFD44 was unique in its ability to utilize all chlorophenoxy acids tested, suggesting that the TfdA-like enzyme in these strains may be distinct from one another and from TfdA.

Cell lysate assays. Assays using cell lysates of the α-proteobacterial 2,4-D degrading strains indicate that all of the strains possess 2,4-dichlorophenol hydroxylase (data not shown, Table 3.3; T.Sassanella and H. Takemi, personal observation) and 3,5-dichlorocatechol dioxygenase

Table 3.3 Hybridization and enzyme activity characteristics of four β-proteobacterial 2,4-D degrading strains. Hybridization was performed by Fulthorpe, 1995. Assays for 2,4-dichlorophenol hydroxylase (TfdB) activity were done in conjunction with Hideto Takemi and are described in Appendix C. The assay for 3.5-dichlorocatechol dioxygenase and some of the results were previously described (Fulthorpe, 1995).

	TFD44	EML146	K1443	M1
Genus	*Sphingomonas*	*Sphingomonas*	*Sphingomonas*	*Nitrobacter*
Growth on 2,4-D	Yes	Slow	Yes	Slow
4-NPAA conversion (whole cell)	Yes	Yes, Slow	Yes	No*
4-NPAA conversion (lysate)	Yes, cytosol	No	Yes, both	Yes, membrane
Hybridize to *tfdA*[2]	No	No	No	No
TfdA-like activity in lysates	Yes, cytosol	No	Yes, both	Yes, membrane
Hybridize to *tfdB*[2]	Yes, weak	Yes, weak	Yes, weak	Yes, weak
TfdB-like activity in lysates[3]	Yes	Yes	Yes	Yes
Hybridize to *tfdC*[2]	No	No	No	No
TfdC-like activity in lysates[4]	Yes	Yes	Yes	Yes
Presence of plasmids	Yes	Unknown	Yes	Yes

activities (Fulthorpe, 1995). Additionally, all strains that hybridize to the *tfdA* probe (as well as some that do not hybridize) possess activity when using 4-nitrophenoxyacetic acid (4-NPAA) in a chromogenic assay for TfdA (Sassanella, 1997; summarized in Table 2.2). Strains that do not convert 4-NPAA in cell lysates also lost their TfdA-like activity, but retained their 2,4-dichlorophenol hydroxylase and 3,5-dichlorocatechol dioxygenase activities (Fulthorpe, 1995; Sassanella, 1997). These studies indicate that all the strains in the collection follow an *ortho*-cleavage pathway similar to that found on the catabolic plasmid pJP4, though the hybridization studies indicate a significant diversity among individual genes in different pathways (i.e - isofunctional genes of pathways from different isolates vary in their level of identity to the probe, forming a mosaic pattern of genes with differing relatedness to the canonical pathway). Further, some of the genes were not detectable by hybridization in a number of strains, though their enzymatic activity was present (Table 2.2 and 3.3). These results indicate that there may be non-homologous isofunctional genes involved in the degradation pathway of some of the collection strains.

The lysates of these strains were also tested in a series of reactions to determine if the α-proteobacterial TfdA-like enzymes used the same co-factors as TfdA. All lysates exhibited a requirement for additional iron, but neither M1 nor K1443 required α-ketoglutarate, and M1 did not seem to require ascorbate (Figure 3.2).

Enzymatic activity partitioning. The alternate substrate 4-NPAA was used in place of 2,4-D in many of the subsequent reactions because 4-NP is not utilized by these strains and accumulates (Sassanella, 1997). Though the TfdA-like activity in fractionated lysates of TFD44 are similar to JMP134 in

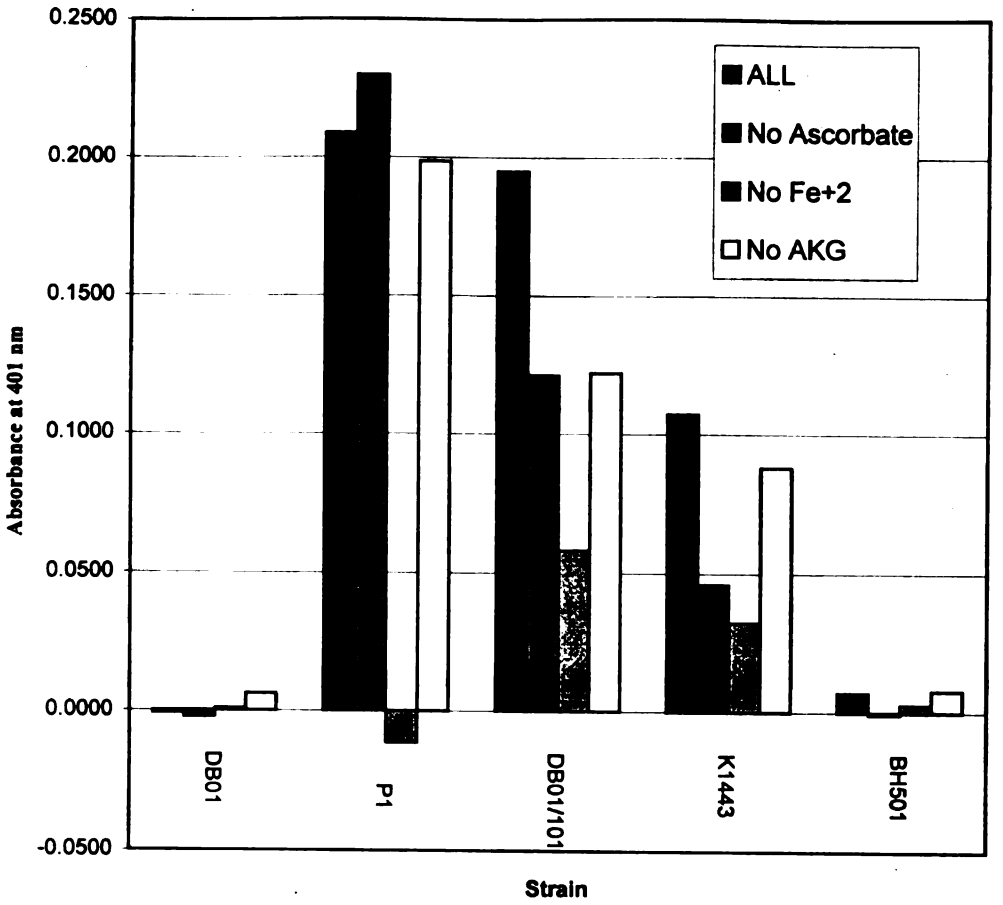


Figure 3.2. Effect of known TfdA co-factors on the conversion of 4-NPAA to 4-NP by the crude lysates of 2,4-D degrading bacteria. The production of 4-NP by the cytosolic fraction (DB01, DB01/101, BH501) or pellet fraction (P1, K1443) of crude lysates was measured spectrophotometrically after a 20 min incubation. Co-factor dependence was determined by including all co-factors in the initial reaction (f.c. - 1 mM Tris (pH 6.8), 1.0 mM 4-NPAA, 50 μM Fe^{+2}, 1 mM ascorbate, 1 mM α-ketoglutarate), then removing a single co-factor (α-ketoglutarate, ascorbate, and iron) in subsequent reactions. The results were normalized to the absorption of the reaction mix containing none of the co-factors.

being nearly completely cytosolic, the lysates of strains M1 and K1443 indicate that the TfdA-like activity in these lysates is at least partially membrane associated (Figure 3.3). TfdA is a cytosolic enzyme, with approximately 5% of the lysate activity remaining in the pellet fraction in the standard assay. K1443 has approximately 23% of its TfdA-like activity present in the pellet fraction and M1 had approximately 67% (results identical to P1, Figure 3.3).

Genetic testing of the α-proteobacterial strains. Attempts were made by several researchers to clone the TfdA-like enzyme in all of these strains, but the *Sphingomonas* strains (EML146, TFD44, K1443) proved intractable to attempts at genetic manipulation. All three of the *Sphingomonas* strains conjugated poorly with *Escherichia coli* and displayed a strong resistance to transposon mutagenesis. All attempts at electroporating these strains under various conditions failed. Plasmid DNA from these strains was easy to obtain, but performed poorly in standard digestion and ligation reactions. *Nitrobacter winogradskii* M1 also proved to be difficult, resisting mutagenesis by 8 different transposons using four delivery systems (S-17/pSUP203 [Tn501], pSUP101[Tn1], pSUP2021[Tn5], pSUP2017 [Tn7] (Simon, 1983); SM10(λ*pir*)/pUT[mini-Tn5 lacZ1] and (λ*pir*)/pUT[mini-Tn5 lacZ2] (deLorenzo, 1990); HB101/ColE1[Tn3] and ColE1[Tn7] (Bagdasarian, 1997), DH5α/pPRL1062a[Km^r] (Wolk, 1991; Cohen, in press)). M1 also conjugated poorly, with conjugal matings producing no transconjugants when using *Escherichia coli* strains (HB101, DH5α, XL-1 Blue), *Pseudomonas putida* KT2442, or *Pseudomonas aeruginosa* PB2076 as donors or recipients. M1 would conjugate at very low frequencies with

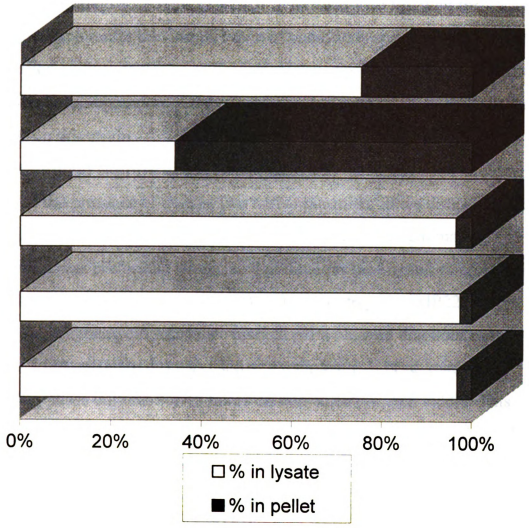


Figure 3.3. Compartmentalization of TfdA-like activity in several 2,4-D degrading strains. The strains shown, from top to bottom: K1443, M1, JMP134, AE228, and BH501. The bars represent 100% of the TfdA-like activity found in each of the lysates, are not representative of total units of activity, and are divided into the activity found in the cytosolic and pellet fractions of crude lysates. In the case of DB01, the level of 4-NP production is 0.05% that of DB01 carrying pJP4, yet the compartmentalization of that activity is the same. TfdA is a cytosolic enzyme, typically showing about 5% of the cellular conversion in the pellet fraction. M1 and K1443 are notable among the strains tested, having 23% and 67% of their activity in the pellet fraction. Strain P1 produced results identical to that of M1. This may indicate membrane association of their TfdA-like activity.

Alcaligenes eutrophus strains ($3.12x10^{-9}$). M1 was also resistant to electroporation under the normal range of conditions and after cell preparation as *Pseudomonas* or *Escherichia* species. Since M1 was the most amenable to manipulation of the α-proteobacteria strains, further work was focused on this strain.

Conjugal capture of M1 plasmids. CHEF gel analysis of strain M1 indicated the presence of at least two native plasmids. Since direct genetic manipulation of M1 was difficult, a plasmid capture conjugal mating experiment was preformed similar to that done for *tfdA* in soil microcosms (Top, 1996). *Alcaligenes eutrophus* BH501 (*A. eutrophus* JMP228, containing a kanamycin resistance marked pJP4 plasmid that does not express TfdA- pBH501dy) was used as a recipient in a broth mating of liquid log phase cultures for 18 hours on non-selective, rich media. This mating procedure was successful in producing colonies capable of growing on minimal salts noble agar with 100μM kanamycin and 100 mM 2,4-D as primary carbon source. Six colonies were isolated, and verified as *A. eutrophus* by carbon source utilization, colony color, colony morphology, and growth rate. These strains are given the *Alcaligenes eutrophus* strain names P1 through P6. A parallel experiment using M1 as a donor and JMP222 as the recipient produced no colonies capable of growing on 2,4-D. This may indicate that an incomplete pathway is being transferred on the M1 plasmid since the plasmid alone does not allow for growth in *Alcaligenes*. In the complemented strains, the frequency of complementation was more indicative of conjugation rather than transposition.

Growth and genetic characteristics of M1 and P strains. *Nitrobacter* species are typically slow growing bacteria that are facultative

lithoautotrophs. Mixotrophic growth is possible, and in the presence of both nitrate and organic substances, a biphasic growth pattern is observed (Holt, 1994). Strain M1 grew slowly on 2,4-D after a considerable lag phase in a minimal medium that used ammonium as the primary nitrogen source. This lag phase was repeated whenever the strain was allowed to enter stationary phase, or whenever the media type was changed.

Chromosomal plugs were prepared from *A. eutrophus* strains P1 through P6 and their parental strains, M1 and BH501. CHEF gel analysis of these strains indicated that an approximately 60 KB plasmid from strain M1 was conjugally transferred to *Alcaligenes eutrophus* BH501 (Figure 3.4). This plasmid was present in all of the P strains. The size of pBH501dy was unaffected, suggesting that the TfdA-like activity was associated with the transfer of the M1 plasmid rather than from any form of recombination.

The strains were confirmed to be *A. eutrophus* by carbon utilization, and it was noted that growth on 2,4-D was substantially slower than observed for *A. eutrophus* JMP134. P1, however, does not suffer the long lag phase of strain M1 grown on 2,4-D. Enzymatic assays using 4-NPAA indicate membrane association of the TfdA-like activity and a lack of α-ketoglutarate-dependence, similar to that found in strain M1 and unlike that of JMP134 (Figure 3.2).

Attempts to isolate and subclone the genes encoding the TfdA-like activity. Several attempts to isolate the 60 kb plasmid from pJP4 failed, perhaps indicating that the plasmid lacks its own transfer genes and is merely mobilized by pJP4. The 60 kb plasmid was able to replicate in *Pseudomonas putida* 2442 and *A. eutrophus* JMP222, but was not found to replicate in *Escherichia coli* after a series of CHEF gel analyses. Cloning of

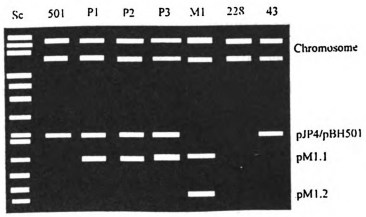


Figure 3.4. Schematic diagram of the CHEF gel analysis of the intact genetic elements found in strains used in the conjugal capture of a TfdA-like gene from strain M1. This schematic represents an original CHEF gel, traced directly from a scanned photograph. Strains included: 501, *A. eutrophus* BH501 (containing pBH501, *tfdA*$^-$; Top, 1996); P1, P2, and P3, *A. eutrophus* strains (containing pBH501 and a 60 kb plasmid from strain M1) that may utilize 2,4-D as a sole carbon source; M1, *N. winogradskyi* strain M1 (containing at least two uncharacterized plasmids); 228, *A. eutrophus* JMP228 (containing no plasmids, (Don, 1985)); 43, *A. eutrophus* JMP134 (TFD43) (containing pJP4). Sc represents the *Saccharomyces cereviseae* chromosomal molecular weight marker, with bands representing 2200, 1600, 825, 785, 750, 680, 610, 565, 450, 365,450, 365, 285, and 225 kb of linear DNA. Plasmid DNA was undigested and migrated more slowly. This analysis clearly shows the presence of pJP4 (or the *tfdA*$^-$ version, pBH501) in *A. eutrophus* strains BH501, P1, P2, P3, and JMP134. A second plasmid of approximately 60 kb was present in M1, P1, P2, and P3. A third, smaller plasmid was present in strain M1 alone. There may be additional smaller plasmids present in these strains that would not be seen well under these conditions.

the *tfdA*-like gene was attempted by shotgun cloning methods. Plasmid DNA for the cloning attempts was prepared using Quiagen maxi-prep columns and the large plasmid procedural modifications or by rapid plasmid DNA preparation (Holmes, 1981). Mixed plasmid DNA from strain M1 or strain P1 was used for cloning. Shotgun cloning was attempted using various partial digestions (*Bam*HI, *Hin*DIII, *Eco*RI, *Eco*RV, *Kpn*I, *Not*I, *Sal*I, *Sac*I, *Sau*3A, *Sph*I) and broad host range plasmid vectors (pTJS75a [Tc^r] (Schmidhauser, 1985); pMMB207 [Cm^r] (Morales, 1991); pMMB503 [Sm^r] (Overbye, 1996)). Competent *E. coli* XL1-Blue cells were transformed with the ligation mixtures and plated on selective media. Transformants were conjugally transferred to *A. eutrophus* BH501 and screened for growth on 2,4-D as a sole carbon source. No plasmids that complemented BH501 were found, perhaps indicating that the genes responsible for complementation in the intact plasmid were arranged in an operon or in such a way that the shotgun clones were not of suitable size. Additionally, deregulated or multiple copies of this gene or genes may be toxic to the host cells.

SUMMARY

This study was designed to look at the TfdA-like enzymes of a diverse set of 2,4-D degrading bacteria, concentrating on those which may be least like the canonical enzyme, in hope of finding an alternate 2,4-D degradation pathway or evolutionarily distinct enzymes. After extensive genetic and biochemical testing, four α-proteobacterial strains were selected. Three of these strains proved intractable to genetic manipulation. In a conjugal capture experiment, strains containing a plasmid that complemented a *tfdA*$^-$ canonical pathway were isolated from *Nitrobacter winogradskii* strain M1. These complemented strains had the same distinct membrane association and indifference to the known co-factors of TfdA as strain M1, indicating that this enzyme may be evolutionarily distinct. CHEF gel analysis confirmed the presence of a 60KB plasmid in the complemented strains. All attempts to clone the TfdA-like activity from the complemented strain or from M1 by shotgun cloning have failed. This enzyme may be encoded on a fragment too large for this method, or the enzyme may be toxic in multiple copies. Further work should provide the DNA sequence of the M1 *tfdA*-like gene(s) and allow more complete biochemical elucidation.

REFERENCES

Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and D. J. Lipman. 1990. Basic Local Alignment Search Tool. J. Mol. Biol. **215**: 403-410.

Amy, P.S., Schulke, J. W., Frazierand, L. M., and R.J. Seidler. 1985. Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. **49**:1237-1245.

Bachmann, B. J. 1972. Pedigrees of some mutant strains of *Escherichia coli* K12. Bacteriol. Rev. **36**:525-557.

Bagdasarian, M. 1997. Strain collection. Mailing address: S-112 Plant Biology Building, Michigan State University, E. Lansing, MI 48224.

Cohen, M. F., Meeks, J.C., Cai, Y. A., and C. P. Wolk. In press. Transposon mutagenesis of heterocyst-forming filamentous cyanobacteria.

DeLorenzo, V., Herrero, M., Jakubzik, U., and K. Timmis. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative Eubacteria. 1990. J. Bacteriol. **172**: 6568-6572.

Devereaux, J. 1989. *The GCG Sequence Analysis Software Package Versions 6.0 - 8.1*, Genetics Computer Group, Inc., University Research Park, 575 Science Dr., Suite B, Madison, WI, 53711, USA.

Don, R.H., A.J. Weightman, H.J. Knackmuss, and K.N. Timmis. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate

in *Alcaligenes eutrophus* JMP134 (pJP4). 1985. J. Bacteriol. **161**:85-90.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **31(6)**: 368-76.

Figurski, D. H., and D. R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76**: 1648-1652.

Fukumori, F. and R. P. Hausinger. 1993. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J. Biol. Chem. **268**:24311-24317.

Fulthorpe, R.R., C. McGowan, O.V. Maltseva, W.E. Holben, and J.M. Tiedje. 1995. 2,4-dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. **61**: 3274-3281.

Fulthorpe, R.R., A. N. Rhodes, and J. M. Tiedje. 1996. Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetate via different microbial populations. Appl. Environ. Microbiol. **62**:1159-1166.

Herrero, M,. deLorenzo, V., and K. N. Timmis. 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. 1990. J. Bacteriol. **172**: 6557-6567.

Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids.Anal. Biochem. **114**: 193-197.

Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T., and S. T. Williams,

eds. Bergey's Manual of Determinative Bacteriology, Ninth Edition, Williams and Williams, Philadelphia, 1994.

Ka, J.O., Holben, W.E., and J.M. Tiedje. 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. Appl. Environ. Microbiol. **60**:1106-1115.

Kamar, S., Tamura, K., and M. Nei. 1993. *MEGA: Molecular Evolutionary Genetics Analysis, version 1.0*, The Pennsylvania State University, University Park, PA, 16802, USA.

Matheson, V. G., Forney, L. J., Suwa, Y., Nakatsu, C. H., Sexstone, A. J., and W. E. Holben. 1996. Evidence for acquisition in nature of a chromosomal 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase gene by different *Burkholderia* spp. Appl. Environ. Microbiol. **62(7)**: 2457-2463.

McGowan, C. 1995. Interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetate degrading bacteria. Ph.D. thesis, Department of Microbiology, Michigan State University.

Morales, V. M., Bakman, R., and M. Bagdasarian. 1991. A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. Gene **97**: 39-47.

Overbye Michel, L. M. Sandkvist, and M. Bagdasarian. 1995. Specificity of the protein secretory apparatus: secretion of the heat-labile enterotoxin B subunit pentamers by different species of Gram$^-$ bacteria. Gene **152**: 41-45.

Peck, S. C., Reinhardt, D., Olson, D. C., Boller, T., and H. Kende. 1992.

Localization of the ethylene-forming enzyme from tomatoes, 1-aminocyclopropane-1-carboxylate oxidase, in transgenic yeast. J. Plant Physiol. **140(6)**: 681-686.

Sassanella, T., F. Fukumori, M. M. Bagdasarian, and R. P. Hausinger. Use of 4-Nitrophenoxyacetic Acid for the Detection and Quantification of 2,4-D/α-Ketoglutarate Dioxygenase Activity in 2,4-D Degrading Microorganisms. Appl. Environ. Microbiol. **63**(3):1189-1191.

Schmidhauser, T. J. and D. R. Helinski. 1985. Regions of broad-host-range plasmid RK2 involved in replication and stable maintenance in nine species of Gram-negative bacteria. J. Bacteriol. **164**: 446-455.

Simon, R., Priefer, U., and A. Pühler. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. 1983. BioTechnology. **1**: 784-790.

Suwa, Y., Wright, A. D., Fukumori, F., Nummy, K. A., Hausinger, R. P., Holben, W. E., Forney, L. J. 1996. Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase from *Burkholderia* sp. Strain RASC. Appl. Environ. Microbiol. **67(2)**: 2464-2469.

Takeuchi, M., Sawada, H., Oyaizu, H., and A. Yokota. 1994. Phylogenetic evidence for *Sphingomonas* and *Rhizomonas* as nonphotosynthetic members of the α–4 subclass of the Proteobacteria. Int. J. Syst. Bacteriol. **44(2)**: 308-14.

Tonso, N.L., V.G. Matheson, and W.E. Holben. 1995. Polyphasic characterization of a suite of bacterial isolates capable of degrading 2,4-D. Microbial Ecol. **30**: 3-24.

Top, E. M., Holben, W. E., and L. J. Forney. 1995. Characterization of diverse 2,4-D degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. **61**: 1691-1698.

Top, E. M., O. M. Maltseva, and L. J. Forney. 1996. Capture of a catabolic plasmid that encodes only 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase (TfdA) by genetic complementation. Appl. Environ. Microbiol. **62(7)**: 2470-2476.

Wolk , C. P., Cai, Y., and J. M. Panoff. 1991. Use of a transposon with luciferase as a reporter to identify environmentally responsive genes in a cyanobacterium. Proc. Natl. Acad. Sci. USA **88**: 5355-5359.

Chapter 4:

Variable 2,4-D Permeability in Bacteria

ABSTRACT

During the investigation of the evolution and assembly of 2,4-dichlorophenoxyacetic acid (2,4-D) catabolic pathways, a colorimetric assay was developed for the detection and quantification of 2,4-dichlorophenoxyacetic acid (2,4-D)/α-ketoglutarate dioxygenase (TfdA) activity in extracts and whole cells of environmental and genetically engineered bacteria (Chapter 2). During the construction of many of the engineered strains used in these experiments, results were observed that suggested that cell permeability to 4-NPAA (and thereby 2,4-D) varied by species, and that this permeability was affected in some strains by the introduction of the low copy number pJP4 plasmid, but not affected by the higher copy number expression vector containing an overexpressible *tfdA* gene. In addition, radiolabeled 2,4-D uptake assays indicate substantially different levels of permeability to 2,4-D among 2,4-D degrading isolates and engineered strains. Permeation is very likely to be a significant factor in successful 2,4-D degradation. The canonical 2,4-D degradation plasmid pJP4 influences 2,4-D uptake in some cells and this activity can be eliminated by transposon mutagenesis. These findings could have significant biochemical and evolutionary implications.

INTRODUCTION

The selective transport of solutes across cellular membranes is a significant part of cellular metabolism. Gram-negative bacteria have a complex cell surface, consisting of an outer membrane comprised of phospholipids, proteins and lipopolysaccharides, a cell wall comprised of peptidoglycan, and an inner, or cytoplasmic, membrane composed of phospholipids and proteins (Nikaido, 1985). Solutes have to pass through considerable obstacles before they may be metabolized in the cytoplasm. The outer membrane provides limited permeability to small solutes via proteinaceous channels that may be non-specific (porins)(Sukhan, 1995), or substrate specific (Nikaido, 1985). The cell wall is thought to be completely permeable to most solutes while providing rigidity to the cell membrane complex. The cytoplasmic membrane, however, is impermeable to most solutes unless a transport system is provided. Gram-positive bacteria have only a single membrane, and its permeability characteristics are like those of the Gram-negative cytoplasmic membrane.

It has been noted that Gram-negative bacteria have more resistance to lipophilic and amphiphilic compounds (dyes, detergents, antibiotics) than Gram-positive bacteria (Nikaido, 1996). This intrinsic resistance of these bacteria was once entirely attributed to the outer membrane, as narrow porin channels slow the penetration of even small hydrophobic solutes and the low fluidity of the lipopolysaccharide decreases the diffusion of lipophilic solutes (Nikaido, 1985, Plésiat, 1992). Even with these characteristics, however, solute equilibrium across the outer membrane is achieved rapidly. Periplasmic concentrations of many antibiotics can reach 50% of the

external concentration in 10 to 30 seconds in *Pseudomonas aeruginosa*, known as a highly impermeable strain, and in a much shorter in *E. coli* (Nikaido, 1994). Solutes like 2,4-D should have no difficulty in rapidly penetrating to the periplasmic space of most Gram-negative bacteria, where transport across the inner membrane is typically facilitated in some manner. The general impermeability of the cytoplasmic membrane is likely to have greater significance in the case of xenobiotic compounds due the lack of specific transporters. Many of these compounds may have to rely on fortuitous transport by permeases specialized for the transport of structurally or chemically related substrates.

Bacterial permeases can be broadly grouped according to their mechanism of energy coupling: those driven by electrochemical gradients (synport, antiport, uniport), and those driven by substrate-level phosphorylation (including the ABC import permeases) (Ames, 1990). Electrochemical gradient transporters are typically osmotic shock resistant, are a single, highly hydrophobic transmembrane protein, transport substrates by synport or antiport mechanisms driven by an ion or proton gradient, have a lower affinity for the substrate, and generate low to moderate concentration gradients. Substrate-level phosphorylation permeases are typically osmotic shock sensitive, are complex in structure, have a higher affinity for their substrates, and can generate very large concentration gradients (up to 10^5-fold).

Among the import permeases, ABC transporters are members of the second group, forming an evolutionarily related superfamily of proteins that transport a wide variety of substances (sugars, amino acids, peptides, ions, and vitamins). They are present in Bacteria, Eucarya (Ames, 1986), and

Archaea (Jovell, 1996). In Gram-negative bacteria, a periplasmic substrate-binding protein will bind the substrate once it diffuses to the periplasm. The binding protein then presents the substrate in a concentrated form to a cytoplasmic membrane bound complex. This concentration is effective due to the high binding affinity of the periplasmic protein and the high concentration of the protein in the periplasm, which can be in the mM range. The binding protein interacts with the hydrophobic membrane spanning subunits of the complex, releasing the substrate and allowing it to be transported across the membrane in an ATP driven mechanism. Each transport requires the hydrolysis of one or two ATP molecules.

Due to the chemical structure of 2,4-D, it is likely that 2,4-D is transported into some cells fortuitously by one or more native permeases, like the aromatic amino acid permeases for phenylalanine (*pheP*), tryptophan (*mtr* and *tnaB*), or tyrosine (tyrP) (reviewed in Sarsero, 1991). Additionally there are permeases, like *aroP* the general aromatic amino acid permease (Sarsero, 1991), that have a greater substrate range and may transport 2,4-D across the cytoplasmic membrane. The permeases that are most likely to be involved in 2,4-D transport are the substrate-level phosphorylation driven ABC importers, due to the high substrate affinity required to transport 2,4-D in the environment and the degree to which 2,4-D is removed from culture media and contaminated soils.

In nature, conjugative plasmids are known to carry many types of permease genes, some provide resistance to anti-microbial agents by increasing efflux (arsenic, cadmium, chromate, tetracycline, ethidium) or by increasing influx or a detoxifying reaction (mercury). Others assist bacteria in co-factor assimilation (iron), or by extending potential carbon sources

(sucrose and citrate in *E. coli*) (reviewed in Tisa, 1990). Though many catabolic plasmids for synthetic chemicals have been studied, few have examined the permeation of the xenobiotic compounds into bacterial cells that degrade them. The bioavailability and uptake of these compounds is likely to be a significant factor in the process of biodegradation, and has direct biochemical and evolutionary implications. In the case of 2,4,5-T degradation, there has been some indication that permeation is affected by a subcloned region of DNA containing some of the degradation pathway genes (Haughland, 1991). This work attempts to examine permeation factors that may affect the utilization of 2,4-D.

RESULTS AND DISCUSSION

Identification of Putative Permeation Effects.

After examining the response of various host/vector constructs to the 4-NPAA assay, the results demostrated that native cell permeability to 2,4-D varied widely by species, from naturally permeable to impermeable, and that permeability was affected in some strains by the addition of the 90 kb plasmid pJP4 (containing the entire known 2,4-D degradation pathway and several other genes), but was not affected by a high copy number expression vector containing an overexpressed *tfdA* gene (Table 2.1 and 4.1). *Escherichia coli* CB101 carrying a fully expressed *tfdA* gene (pMMB511) produced no 4-NP from 4-NPAA in whole cells, though when the cells were sonicated, the lysates rapidly converted 4-NPAA to 4-NP. The same strain containing pJP4 reacted similarly, though tfdA was expressed at a much lower level. *Burkholderia cepacia* DB01 did not convert 4-NPAA, but

Table 4.1. Response of various engineered host/plasmid combinations using 4-NPAA as an alternate substrate to assay for TfdA activity. (+) indicates production of 4-nitrophenol from 4-NPAA under conditions noted in Appendix B.

Strain	no plasmid	$TfdA^+$	$TfdA^+$	pJP4	
		whole cells	lysates	whole cells	lysates
Escherichia coli CB101	-	-	+	-	-
Pseudomonas putida PB2442	-	-	+	+	+
Alcaligenes eutrophus JMP228	-	+	+	+	+
Burkholderia cepacia DB01	-	+	+	+	+

produced 4-NP from 4-NPAA when either pJP4 or pMMB511 were present, indicating that 2,4-D and 4-NPAA were transported into the cell. However, in *Pseudomonas putida* PB2442, the introduction of plasmid pJP4 allowed the uptake and conversion of 4-NPAA, whereas the introduction of pMMB511 alone did not. These results support the variability of permeation of 4-NPAA (and likely 2,4-D) in different environmental strains, and suggest that there may be a factor on the plasmid pJP4 that influences the permeation of substituted phenoxy compounds.

In order to clarify that the permeation effects noted for 4-NPAA were also applicable to 2,4-D, a ^{14}C-2,4-D uptake assay was developed (Figure 4.2). *E. coli* DH5α was shown to be impermeable to 2,4-D and 4-NPAA using the 4-NPAA assay and the radiolabeled uptake assay (Figure 4.2). The addition of the test plasmids (pJP4 and pMMB511) had no effect on the permeability of this strain to 2,4-D. *Burkholderia cepacia* DB01 and *Alcaligenes eutrophus* JMP228 were shown to retain ^{14}C-2,4-D at a low level (presumably cell adsorption or cellular transport equilibrium). When carrying pMMB511 (expression vector with *tfdA* alone), both strains readily transported and converted 4-NPAA (Table 2.1), and when carrying pJP4, both showed significant uptake and conversion of ^{14}C-2,4-D over a 5 minute period (Figure 4.2). However, in the case of *Pseudomonas putida* PB2442, ^{14}C-2,4-D was retained very poorly in the absence of a plasmid or when carrying pMMB511 (in a similar manner as for 4-NPAA; Fig. 4.1), yet transported 2,4-D into the cell at a much higher rate in the presence of pJP4. These findings indicate that the 4-NPAA results are a good reflection of the transport of 2,4-D, that there is substantial variability among bacterial

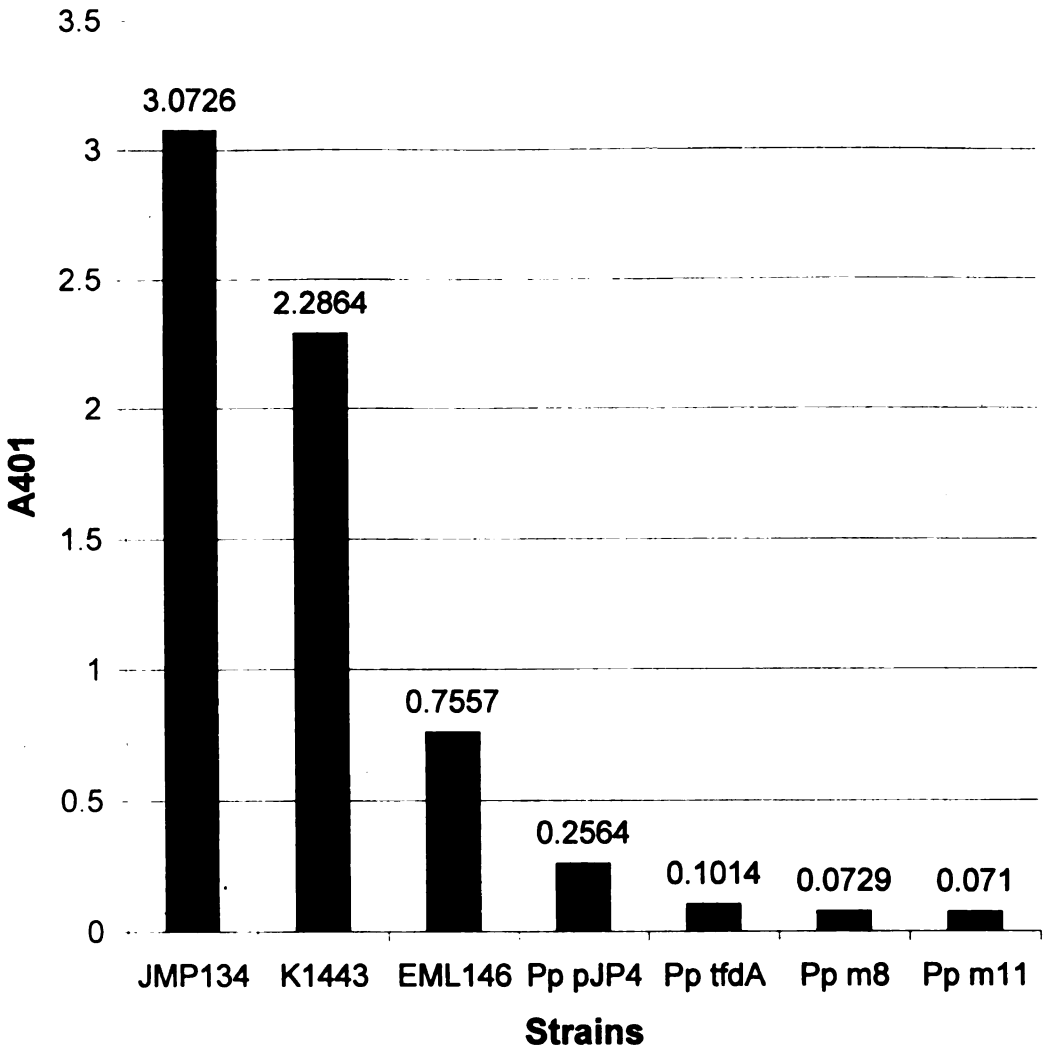


Figure 4.1. Variability in the permeation or utilization of 2,4-D of environmental strains and putative permease mutants is indicated by 4-NPAA whole cell assay. Among the native 2,4-D degrading isolates shown (*A. eutrophus* JMP134, *Sphingomonas* sp. K1443 and EML146), there is a substantial difference in the amount of 4-NPAA that is converted to 4-NP over 24 hours in liquid cultures. Preconditioned log-phase cultures grown on M2 medium were used to inoculate secondary cultures of M2 media supplemented with 1 mM 4-NPAA (M24 media). Pp denotes engineered *Pseudomonas putida* PB2442 strains: pJP4 denotes the presence of that plasmid, tfdA denotes the presence of the expressible *tfdA* construct pMMB511 (Table 2.1), and m8 and M11 are Tn5 insertion mutants of pJP4. *P. putida* strains were preconditioned in M2 media supplemented with 0.1% casamino acids and treated as noted. All cultures were well shaken at 30°C for 24 hours. One ml cultures samples were chilled to 4°C, pelleted, and the absorbance of the supernatant measured at 401 nm.

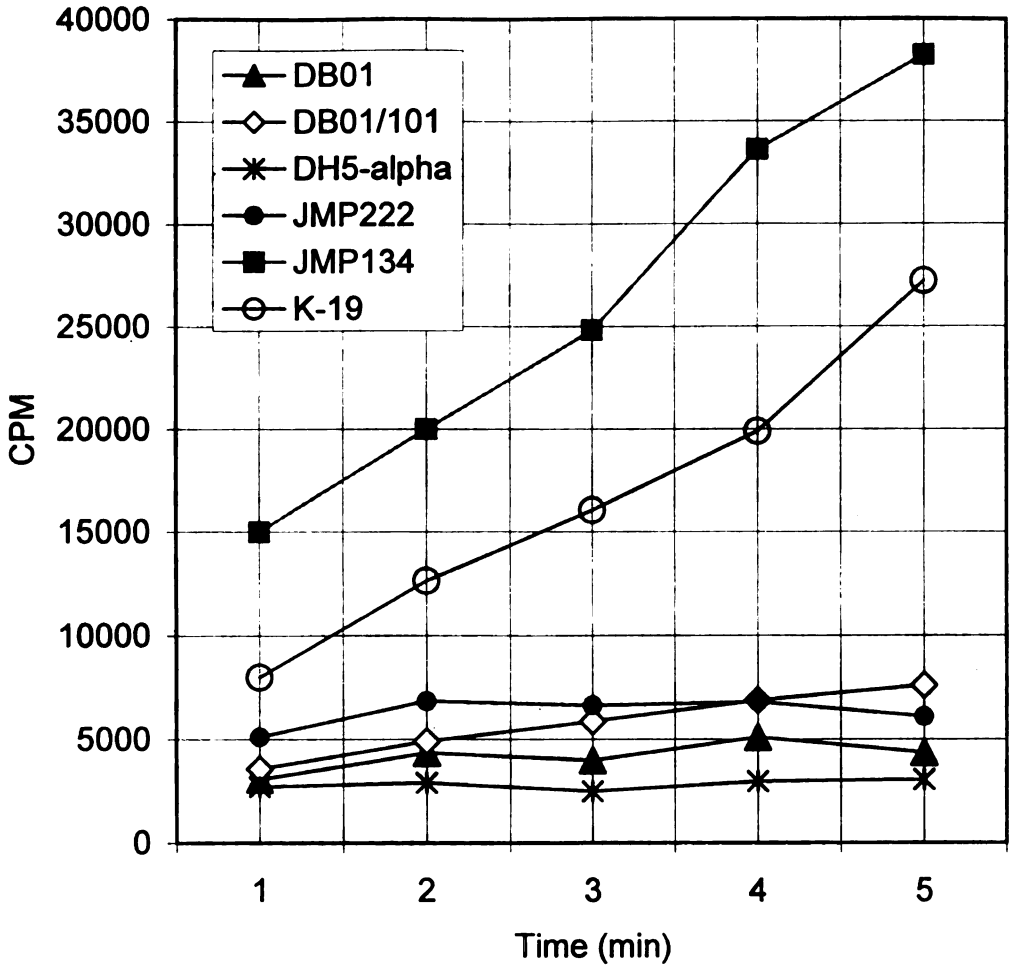


Figure 4.2. Determination of 2,4-D permeation in various environmental and engineered strains using a 2,4-D uptake assay. Environmental strains *Burkholderia cepacia* DB01, *Alcaligenes autrophus* JMP228 and JMP134, and *Alcaligenes* sp. K-19, and the engineered strains *E. coli* DH5α and *B. cepacia* DB01/101 were tested for permeation of 2,4-D. These strains were grown in the presence of 2,4-D, washed in minimal medium without 2,4-D, and incubated in a large volume of the same medium. The culture was then spiked with 1 mM 2,4-D with a ^{14}C-2,4-D tracer, the suspension was mixed thoroughly, and samples drawn at one minute intervals. The samples were rapidly filtered through a 0.45 μm Millipore nitrocellulose filter, and washed three times with minimal salts medium containing 1 mM 2,4-D. The filters were immediately paced into a toluene based scintillation fluid. Uptake of the radiolabeled 2,4-D was measured by scintillation counter, and adjusted for background and absorbance of 2,4-D to the nitrocellulose filters.

isolates in the ability to take up 2,4-D, and that the plasmid pJP4 allows *P. putida* PB2442 to take up 2,4-D at a higher rate than the native strain.

The Mutagenesis of pJP4, and the Subsequent Screening for Permease Mutants.

To test the idea that pJP4 may harbor an unknown gene that affects the permeability of bacterial cells like *P. putida* to 2,4-D and 4-NPAA, random transposon mutagenesis of pJP4 was performed in *E. coli*, and the mutant plasmids were mated into *P. putida* PB2442. The transconjugants were then screened using the 4-NPAA colony assay (Appendix B), and 22 colonies that failed to convert 4-NP were chosen. These strains were purified, then tested for 4-NPAA conversion again, using the whole cell liquid culture assay (Appendix B). The same strains were tested as cell lysates, and two strains, denoted m8 and m22, were found to convert 4-NPAA in lysates, but not in whole cell assay.

Hybridization Experiments.

Southern hybridization experiments using the pJP4 mutant plasmids m8 and m22 and an internal fragment of Tn5 as a probe indicated that the transposon insertions did not occur within the known 2,4-D genes, but rather in a 33 kb fragment containing the replication and maintenance genes. Production of 4-NP using the whole cell 4-NPAA assay to show uptake and conversion by TfdA in these strains is shown (Fig. 4.1). The mutant strains have about 28% of the transport rate of the wild type pJP4.

2,4-D Incorporation Assay.

The deleterious effect of the pJP4:Tn5 transposon insertions m8 and m22 on permeation of the host strain was tested more rigorously using a ^{14}C-2,4-D incorporation assay. *Alcaligenes eutrophus* JMP134 (wild-type pJP4) incorporated radiolabeled 2,4-D, whereas *A. eutrophus* JMP228 (no plasmid) and *A. eutrophus* carrying the mutant pJP4 plasmids m8 and m22 did not appreciably incorporate radiolabeled 2,4-D within 150 minutes (Figure 4.3). The increase in incorporation of carbon from 2,4-D by JMP134 was significant, and could not be attributed to simple increase in cell mass (or cell absorbance) (Figure 4.3). All of the strains except JMP228 had a detectable level of TfdA activity when tested in lysates. These findings were further supported by a simple growth study where the strains were inoculated into minimal 2,4-D medium and incubated for 5 days. JMP134 grew to maximum density in less than 48 hours and m22 produced a small increase in cell mass - perhaps indicating that the transposon insertion hasn't completely eliminated the function of the permeation factor - whereas none of the other strains grew appreciably.

Partial Subcloning of the Transposon Insertion Site.

Several methods of cloning the putative permeation factor have been attempted unsuccessfully due to what is thought to be toxic effects of multiple copies of the region. Asymmetric subclones of the transposon insertion site have been isolated by partially digesting plasmid DNA from m8 and m22 using *Bam*H1. This enzyme cuts in the center of Tn5, yet leaves the *neoR* gene intact, allowing the tracking of the fragment by screening for kanamycin resistance. A subclone from m22 was isolated

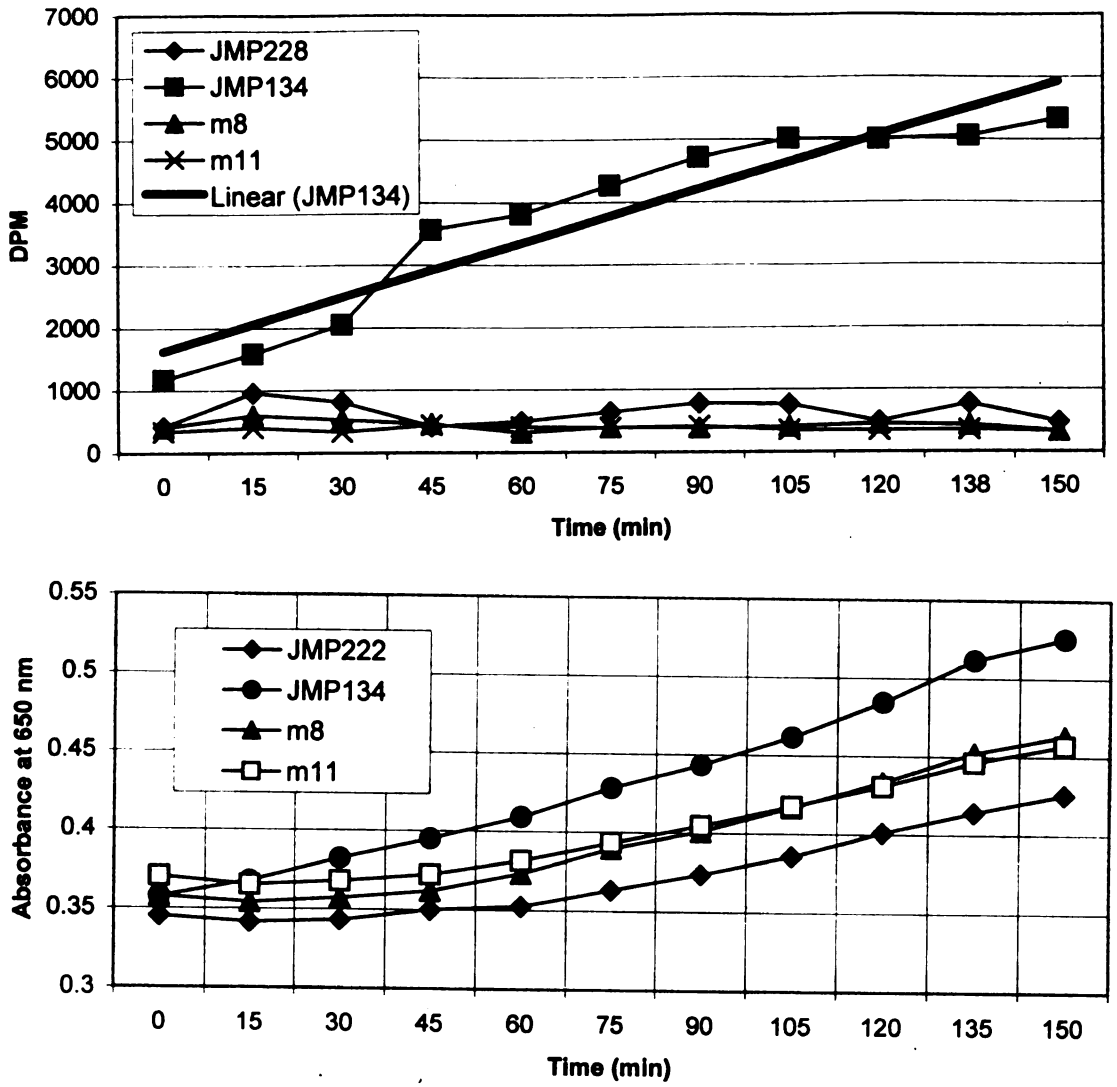
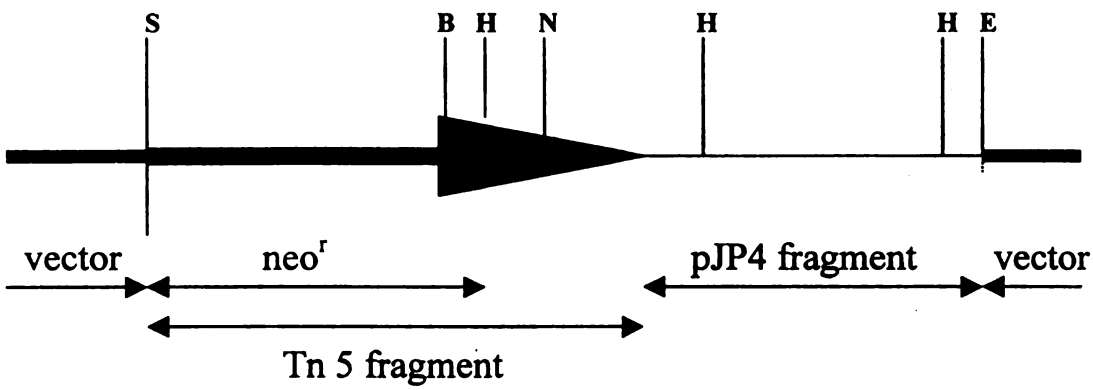


Figure 4.3. 2,4-D incorporation assay. The permeation of 2,4-D in *Alcaligenes eutrophus* strains was tested by measuring incorporation of ^{14}C from ring-labeled ^{14}C-2,4-D. JMP222 contains no plasmids, JMP134 carries pJP4 (the canonical 2,4-D pathway), and m8 and m11 are Tn5 mutants of pJP4 that exhibited decreased ability to convert 4-NPAA to 4-NP in whole cells, yet expressed *tfdA* in cell lysates. The incorporation of carbon from 2,4-D degradation into biomass was measured as DPM from washed culture samples taken at various times (top panel). The cell density of the cultures was measured simultaneously (bottom panel).



Scale: 1" = approx. 1 kb

Figure 4.4. Physical map of plasmid pKm2. This fragment is a portion of the plasmid m8 (pJP4:Tn5)(this work) in the vector pMMB207, which is only partly shown (Morales, 1991). B represents restriaction sites for *Bgl*I, E represents *Eco*RI, H represents *Hin*DIII, and N represents *Not*I; neo^r provides resistance to neomycin and kanamycin. The fragment was isolated by a 'walking' experiment - using the unique *Sal*I site and the kanamycin resistance gene of Tn5 to capture one end of the transposon as well as a portion of the insertion site on pJP4. It is likely that the distal end of this fragment has rearranged leaving the polylinker sites of the vector intact. This plasmid is relatively unstable, being maintained only under antibiotic selection, and is difficult to purify.

using pMMB207 as a vector. This plasmid was difficult to purify in large quantities, having a strong tendency to partition to the hyrophobic phase during phenol/chloroform extractions. Plasmid DNA was purified using the Quiagen mini-plasmid preparation columns. This 5.3 kb subcloned fragment contains the portion of Tn5 that contains the *neoR* gene and extends out past the end of the transposon (Fig. 4.4). This fragment cannot be excised using *Bam*H1, indicating that this fragment may have rearranged at the distal end to the transposon to generate the stable subclone. Since the entire fragment cannot be excised for transfer to a more suitable sequencing vector, direct sequencing of the putative permeation factor using a primer to the IS sequence of the transposon was attempted. Four automated attempts at three facilities and manual single strand PCR sequencing attempts failed at the time this study ended.

Evolutionary and Ecological Implications.

The effect of differential permeability of compounds in nature can have significant ecological and evolutionary impact. If there is a permeation step previous to a cytosolic degradation reaction, the understanding of the kinetics of xenobiotic degradation pathways could be significantly altered. Additionally, if permeation were the slow step in the degradation reaction, this could have significant ecological impact on the competitiveness of the strains in the environment, and, in a practical sense, the predicted utility of the strains in biodegradative processes. Biochemical elucidation of the effect of various targeted or fortuitous permease reactions would also assist in understanding the effect of bioavailability of the substrate in natural or engineered environments. Since it is clear there are substantial differences

in the ability of Gram-negative bacteria to transport 2,4-D into the cell, several significant issues emerge. If differential permeation of compounds occurs in environmental bacteria, it is likely that only a limited number of bacterial strains are suitable for metabolic level 2,4-D degradation. It is known that there are plasmids that carry catabolic pathways that encode transport proteins that increase the level of permeation of specific substrates in their host cells. There have been some indications that this phenomenon may be more widespread that was thought previously. Further, if the level of permeability can be altered by the addition of plasmid borne factors, there may be a limitation on the species that may be affected, perhaps due to poor membrane insertion, inability for periplasmic facilitators to reach the periplasmic space, or the lack of required associated membrane proteins. Non-specific transporters may be more important for the evolution of novel catabolic pathways, and may be a contributing factor in the recalcitrance of certain compounds to microbial degradation.

This work suggests that there are strains that do not transport 2,4-D across their cell membranes in the native state, and therefore the number of strains that may be able to utilize 2,4-D as a significant source of carbon and energy may be limited. It is highly likely that there is some kind of permeation or 2,4-D uptake altering factor present on pJP4, the canonical 2,4-D degradation plasmid, and that factor may be disrupted by site directed mutagenesis (producing significant loss of 2,4-D uptake) without interrupting the expression of *tfdA*. This factor only functions in some of the strains, and is not needed in others, as there are native transporters that allow significant uptake of 2,4-D. 4-NPAA and 2,4-D are transported in the same manner in all the strains that would tested. The mechanism of 2,4-D

uptake in bacterial strains is unknown, but it is likely that 2,4-D is transported by common broad-substrate range ABC import permeases, like *aroP*. The transport of 2,4-D is probably driven by substrate-level phosphorylation, due to the apparent high affinity of the permease and the virtually complete removal of 2,4-D from culture media.

REFERENCES

Ames, G. F.-L. 1986. Bcaterieal periplasmic transport systems: structure, function, mechanism, and evolution. Ann. Rev. Biochem. **55**: 397-425.

Ames, G. F.-L. 1988. Structure and mechanism of bacterial periplasmic transport systems. J. Bioenerg. Biomembr. **20(1)**: 1-18.

Ames, G. F.-L., and A. K. Joshi. 1990. Energy coupling in bacterial periplasmic permeases. J. Bacteriol. **172(8)**: 4133-4137.

Boorsma, A., van der Rest, M. E., Lolkema, J. S., and W. N. Konings. 1996. Secondary transporters for citrate and the Mg^{+2}-citrate complex in *Bacillus subtilis* are homolog proteins. J. Bacteriol. **178(21)**: 6216-6222.

DeLorenzo, V., Herrero, M., Jakubzik, U., and K. Timmis. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative Eubacteria.1990. J. Bacteriol. **172**: 6568-6572.

Fath, M. J., and R. Kolter. 1993. ABC transporters: bacterial exporters. Microbiol. Rev. **57(4)**: 995-1017.

Fukumori, F., and R.P. Hausinger. 1993a. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetic acid monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. **175**:2083-2086.

Fukumori, F., and R. P. Hausinger. 1993b. Purification and characterization of 2,4-dichlorophenoxyacetate/ α-ketoglutarate dioxygenase. J. Biol. Chem. **268**:24311-24317.

Haughland, R. A., Sangodkar, U. M. X., Sferra, P. R., and A. M.

Chakrabarty. 1991. Cloning and characterization of a chromosomal DNA region required for growth on 2,4,5-T by *Pseudomonas cepacia* AC1100. Gene. **100**: 65-73.

Jovell, R. J., Macario, A. J. L., and E. C. deMacario. 1996. ABC transporters in archaea: two genes encoding homologs of the nucleotide-binding components in the methanogen *Methosarcina mazei* S-6. Gene. **174**: 281-284.

Nikaido, H. and M. Vaara. 1985. Molecular basis of bacterial outer membrane permeability. Microbiol. Rev. **49**: 1-32.

Nikaido, H. 1994. Prevention of drug access to bacterial targets: permeability barriers and active efflux. Science. **264**: 382-388.

Nikaido, H. 1996. Multidrug efflux pumps of Gram-negative bacteria. J. Bacteriol. **178(20)**: 5853-5859.

Plésiat, P., and H. Nikaido. 1992. Outer membrances of Gram-negative bacteria are permeable to steroid probes. Mol. Microbiol. **6**: 1323-1333.

Ramos, A., Poolman, B., Santos, H., Lolkema, J. S. , and W. N. Konings. 1994. Uniport of anionic citrate and proton consumption in *Leuconostoc oenos*. J. Bacteriol. **176**: 4899-4905.

Sarsero, J. P., Wookey, P. J., Gollnick, P., Yanofsky, C., and A. J. Pitard. 1991. A new family of integral membrane proteins involved in transport of aromatic amino acids in *Escherichia coli*. J. Bacteriol. **173(10)**: 3231-3234.

Sassanella, T., F. Fukumori, M. M. Bagdasarian, and R. P. Hausinger. Use

of 4-Nitrophenoxyacetic Acid for the Detection and Quantification of 2,4-D/α-Ketoglutarate Dioxygenase Activity in 2,4-D Degrading Microorganisms. Appl. Environ. Microbiol. **63**(3):1189-1191.

Sukhan, A., and R. E. Hancock. 1995. Insertion mutagenesis of the *Pseudomonas aeruginosa* phosphate-specific porin *OprP*. J. Bacteriol. **177(17)**: 4914-4920.

Tisa, L. S. and B. P. Rosen. 1990. Transport systems encoded by bacterial plasmids. J. Bioenerg. Biomembr. **22(4)**: 493-507.

Thayer, J. R., and M. L. Wheelis. 1976. Characterization of a benzoate permease mutant of *Pseudomonas putida*. Arch. Microbiol. **110**: 37-42.

CONCLUSION

This work examines the initial step of the known 2,4-D degradation pathways in 2,4-D degrading soil bacteria. From the canonical pathway, 2,4-D/α-ketoglutarate dioxygenase (TfdA) was the focus of this diverse set of investigations into the microbial ecology and the evolution of catabolic pathways. The significant findings of this work include:

- The α-ketoglutarate-dependent dioxygenase superfamily, to which TfdA belongs, can be divided into several groups that are defined by the conservation of residues thought or known to be essential for iron binding and activity. Though these enzymes are thought to have a very similar mechanism, it is likely that this superfamily is polyphyletic.

- A continuous, quantitative chromogenic assay for TfdA and analogous activities was developed for the screening of environmental and genetically engineered strains using both intact cells and cell lysates. Colony screening assay for petri plates allows rapid testing of large numbers of colonies. This 4-

nitrophenoxyacetic acid (4-NPAA) assay was used to survey a diverse collection of 2,4-D degrading soil bacteria and genetically engineered strains, indicating substantial diversity among Gram-negative bacteria in the ability to transport and/or metabolize 4-NPAA and 2,4-D.

- *Nitrobacter winogradskyi* strain M1 was examined in greater detail, and was found to carry a plasmid that encodes an atypical TfdA-like activity that is both membrane associated and requires different co-factors than TfdA.

- Differential response when using the 4-NPAA assay indicated that there could be differential permeation of 4-NPAA among soil bacteria. Differential permeation was confirmed for 2,4-D using ^{14}C-2,4-D uptake and incorporation assays. Transposon mutagenesis of the canonical 2,4-D degradation plasmid pJP4 indicated that the plasmid encodes an unknown factor that influences the permeation of 2,4-D in some strains.

APPENDICES

APPENDIX A

APPENDIX A

Detailed Methods From Chapter 1.

Strains, enzymatic functions, organisms of origin, and accession numbers are described in Table 1.1. Citations listed here are noted as References in Chapter 1.

Data file assembly. The known α-ketoglutarate-dependent dioxygenase sequences were obtained from world-wide protein databases. All sequences used were amino acid sequences, as more character states were required for a significant analysis due to the complexity of this highly divergent group. The superfamily family was initially defined as those enzymes that bind α-ketoglutarate, ferrous iron, oxygen, and an additional variable substrate, and that theoretically have a similar mechanism (though functionally fall into different categories). This group of enzymes is reviewed by Prescott (1993), and additional sequences from literature searches were added. Using the Genetics Computer Group (GCG) suite of programs (Deveraux, 1987), the sequences were collected and formatted for use. Sequences that did not have the same set of substrates were compared pairwise using **bestfit** or **gap**. Larger groups were tested for sequence similarity using **distances**. Sequences that displayed more than 30% identity of amino acid sequence were automatically placed into the same initial group within the superfamily. The program **pileup** was used to perform multiple sequence alignments of the initial groups, and neighbor-joining analysis was preformed to determine the most divergent members of the group. These

sequences were then used by **blastp** to search for similar sequences. Sequences located by **blastp** were examined for similarity to the search sequence over a minimum of 100 amino acids and for the conservation of residues as evidenced by the members of the initial groups. The sequences of the superfamily were increased as it became apparent that the non-heme iron(II) oxygenases that require a reductant, usually ascorbate, are significantly similar to some of the sequences that require α-ketoglutarate. This 'similarity walking' was done for all of the groups, and resulted in the addition of a number of sequences with unknown function.

Cluster analysis. When the similarity walking was completed, all of the sequences for each group were aligned and conserved residues noted. Considering the breadth of source organisms in several of the clusters, the utilization of such a wide variety of substrates, and the diversity of enzymatic functions that these enzymes catalyze, the retention of specific residues among the group sequences is remarkable and quite characteristic.

NJ Analysis procedure for the α-ketoglutarate-dependent dioxygenases. The multiple sequence alignments for the groups were used to determine the phylogenetic relationships among members of each of the clusters. The MEGA software package (Kamar, 1993)was used for analysis using the Neighbor-Joining (NJ) method. Because the data was strictly protein sequence, p-distance was used as the measure of genetic distance for Neighbor-joining. All analyses were done under the following conditions: NJ analysis, p-distance estimation, 500 replicate bootstrap, and complete deletion option. A table of genetic distances and standard errors was

generated for each of the datasets. The trees were then edited to a presentation format, and saved as text files, typically having the bootstrap values at each node of the tree.

Intracluster analysis. Several of the sequences of each cluster, as well as the sequences that did not fall into a cluster, were then aligned and analyzed using the NJ method. The clusters and individual sequences presented in this work are statistically distinct using this method. It must be noted, however, this analysis is based on the entire amino acid sequence, and may not distinguish the relatedness of specific regions or domains of the sequences. This method also does not detect conservation of three-dimensional structure. Sequences like TfdA and Gbb are remarkably similar in sequence (53% similarity), and may share similar conserved histidine motifs.

APPENDIX B

APPENDIX B

Detailed Methods From Chapter 2.

4-AAP assay for TfdA-like enzymes:

The following assay reagents were prepared immediately before the assay: 5 mM ascorbate, 10 mM α-ketoglutarate, 5 mM Fe^{+2}, 2% 4-aminoantipyrene (4-AAP). Stock reagents include: 50 mM imidazole buffer (pH 6.75), 10 mM 2,4-D, 0.5 M EDTA, Stop Buffer (pH 10), 8% $K_3[Fe(CN)_6]$. The assays were started by mixing 480 µl H_2O, 200 µl imidizole buffer, 100 µl 2,4-D, 10 µl ascorbate, 100 µl α-ketoglutarate, and 10 µl EDTA (EDTA in the negative control only). After vortexing, the following was added: 10 µl Fe^{+2} and 100 µl cell lysate. The amount of cell lysate may be adjusted, equilibrating the volumes with H_2O. Reaction mixtures were incubated for 30 min at 30°C in heating block. After incubation, the following was added 10 µl EDTA (to stop the reaction), 10 µl 4-AAP, 100 µl Stop Buffer, and 10 µl $K_3[Fe(CN)_6]$. Reactions were allowed to sit at room temperature for 20 minutes. The absorbance at 501 nm was measured, centrifuging the reaction mix at top speed in a bench top centrifuge for 1 minute if the reaction mix was cloudy. Cell lysates absorb well at this wavelength, so a reference reaction with EDTA should be done for each strain and/or differing volume of cell lysate..

4-NPAA cell free assay for TfdA:

Cells were grown in minimal inducing medium (M2, MC2, or MC + IPTG), harvested by centrifugation at 10,000 x *g* for 10 min, and washed

once in 20 mM Tris-HCl buffer, pH 7.4. The cells were resuspended in a small volume of buffer (1/20 to 1/50 of the original volume, minimum 1 ml) to equivalent cell densities, chilled in an ice bath, and sonicated while being sure to keep the lysate below 30°C at all times. The lysates were centrifuged for 30 minutes at 100,000 x *g*. The resulting supernatants were used as the cytosolic fractions of the lysates. Cell pellets were washed once, and resuspended in 1 ml of the buffer. The production of 4-NP was monitored continuously at 401 nm for reaction mixtures consisting of 1 mM α-ketoglutarate, 50 μM ascorbate, 50 μM $Fe(NH_4)_2(SO_4)_2$, and 1 mM 4-NPAA in 20 mM Tris-HCl, pH 7.4. The extinction coefficient of 4-NP at pH 7.4 is 17.8 mM^{-1} cm^{-1}. The volume of lysate used in the reaction varied between experimental runs, but the assay results were standardized by adjusting for lysate protein concentration, as determined by the Lowry assay (15).

4-NPAA liquid media/microtiter plate assay for TfdA:

Natural and engineered isolates were grown in M24 or M24+ IPTG media, the cells were pelleted by centrifugation, and the amounts of 4-NP in the supernatants were measured by absorbance at 401 nm. This method was adapted for use in microtiter plates. Cells were grown in the wells of the microtiter plates in appropriate inducing medium. 4-NPAA was added and the cells were incubated at 30°C for 5 minutes to 4 hours. The cells were pelleted by centrifugation of the microtiter plate, the supernatants were transferred to a fresh microtiter plate, NaOH was added to a final concentration of 0.1N, and the amounts of 4-NP in the supernatants were detected using a microtiter plate reader measuring absorbance at 401 nm.

The concentrations of 4-NP were calculated using an experimentally determined absorption coefficient of 18.4 mM^{-1} cm^{-1}.

4-NPAA agar plate colony assay for TfdA:

Petri plates containing colonies of the various isolates were prepared by inoculating M2 agar plates (for those bacteria that will grow on 2,4-D as a sole carbon source) or MC2 agar plates (for those that do not). The cultures were diluted to approximately 200 CFU per plate, or picked onto a grid of similar density. The plates were incubated at 30°C until small to medium-sized colonies were observed. The colonies were assayed for TfdA activity by carefully wetting them with freshly prepared assay solution (10 mM 4-NPAA, 50 μM $Fe(NH_4)_2(SO_4)_2$, and 50 μM ascorbate in 20 mM Tris-HCl, pH 7.4). An atomizer or thin-layer chromatography sprayer was used for the wetting, and this process was repeated once the excess surface liquid had evaporated or been absorbed. Plates were incubated for 5 minutes to 4 hours at 30°C to allow for the production of 4-NP and color development. A yellow 'halo' was observed around positive colonies.

4-NP utilization testing:

All of the strains that were found to convert 4-NPAA to 4-NP were grown on M2 agar plates, inoculated into M24 liquid media, and grown overnight. The cultures were inoculated into M2 medium + 0.250 mM 4-NP, with shaking at 30°C for 48 hours, and 1 ml aliquots of each culture were removed and prepared for HPLC as noted above, using a 0.250 μM 4-NP standard for reference.

APPENDIX C

APPENDIX C

Detailed Methods From Chapter 3.

Assay for Antibiotic Resistances:

Strains used in this assay were previously grown on M2 (Sassanella, 1997) media, and inoculated directly onto 1/10 Luria-Bertani medium containing 16% agar and various concentrations of seven antibiotics (Table 3.1). Inoculated plates were then grown for 36 hours at 30°C and examined for growth.

Auxotrophy Testing:

Strains tested were grown on 1/10 Luria-Bertani medium, then used to inoculate four plates containing minimal salts (Fulthorpe, 1995) and 16% noble agar plates that contained three or more of the following amino acid mixtures: I - alanine, valine, cysteine, threonine, histidine, and methionine; II - arginine, tyrosine, cysteine, serine, and leucine; III - phenylalanine, tryptophan, glutamic acid, hydroxyproline, and isoleucine; IV - aspartic acid, asparginine, glycine, proline, and lysine. If growth occurred on the plate containing all four amino acid mixtures, the strain was tested with several types of the same minimal medium containing a single carbon source, such as 1mM glucose or 1 mM citrate.

16S rDNA Phylogenetic Analysis Methods:

Data file assembly. The *E. coli*, *Rhodopseudomonas palustris,* and M1 partial 16S sequences were obtained C. McGowan. GenBank 16S sequences for *E.coli.* and all *Rhodopseudomonas sp.* were then **stringsearch**ed and **fetched**. The sequences of *Nitrobacter sp.* were **fetch**ed because of the reported closeness in genetic distance. *Agrobacter tumefaciens*, *Pseudomonas dimunata*, and *Erythrobacter longus* were also obtained to represent deeper branches within the same reported clade (Holt, 1994; The Procaryotes). The sequences were then aligned using **pileup**, and edited down to the 300 bp size of the bacterial variable region using **lineup**. This process was repeated several times, including different taxa.

Phylogenetic Analysis of strain M1 and the established Rhodopseudomonas taxa. The data files generated above were then opened in MEGA, and analyzed by the Neighbor-Joining (NJ) method. Twenty-three 16S rDNA sequences of reportedly related bacteria were included in a NJ analysis. This analysis used the Jukes-Cantor method of determining distance, which is more sensitive to varying rates of evolution in taxa, but allows for confidence levels to be determined on the branch lengths generated by NJ. The analysis was then repeated using Maximum Parsimony (MP) analysis. The standard conditions for MP analysis included the use of sequence gaps and additional state (which takes into consideration insertions and deletions normally excluded by the NJ method), the Branch and Bound Algorithm, and a 50% rule for creating a consensus tree if more than one most parsimonious tree is created. MP analysis is the best method for determining the general branching order of a

group of taxa. It does not determine branch lengths, and thus is insensitive to differing rates of evolution among included taxa (which can lead to some inaccuracies when using a breadth of bacterial species).

Confirmation of accuracy of the general branching pattern by using full 16S RNA sequences. Full-length 16S rDNA sequences were used to confirm the general branching pattern found using the bacterial variable region. NJ analysis clusters the *Nitrobacter* species with 87% accuracy. The t-test for confidence was extremely high, with all values over 61 (5 of 8 over 94).

Transposon Mutagenesis Protocols:

All strains were streaked from glycerol stocks onto selective media. Single colonies were grown in Luria-Bertani broth to an optical density of 0.3, and the strains were combined directly on Luria-Bertani agar plate in a 1:4 ratio of donor to recipient (1:1:4 ratio for tri-parental matings). After a 16 h. incubation, the cells were scraped into minimal salts media, vortexed, and plated on appropriate selective media. The pUT (deLorenzo, 1990) and pSUP (Simon, 1983) mutagenesis systems required no additional helper plasmid, but when required, pRK2013 (Figurski, 1979) was used for mobilization. The mating of strain M1 and *A. eutrophus* strains was enhanced by liquid mating variant. The log-phase donor and recipient cells were mixed in a microfuge tube and allowed to incubate for 20 h, then plated as above.

Chromosomal Plug Preparation for CHEF Gel Analysis:

Bacterial cells were grown overnight or longer in appropriate media until good growth was observed. Cells may be pelleted at low speed (7,000x *g*) and frozen at -20°C until needed. Pellets were thawed and resuspended to a chromosomal concentration of (10μg/ml) in 10 mM Tris (pH 7.2)+20 mM NaCl + 100mM EDTA. 1.2 ml of the cell resuspension was warmed to 70C, well mixed with 1.2 ml of prewarmed 1.6% Biorad Ultrapure agarose in 1x TBE, and aliquoted (100 μl) into the gel form, which was chilled to 4°C. Once hardened, the plugs were removed from the gel form using sterile, non-metal tools, and placed in a sterile 50 ml Corning tube. Ten ml of the lysing solution (10 mM Tris (pH7.2), 50 mM NaCl, 100 mM EDTA, 0.2% SDS, 1% *N*-laurylsarcosine) was added to each tube, and they were gently shaken for 90 minutes at 30°C. The plugs were washed twice in 5 ml Wash Solution (20 mM Tris (pH 8.0), 50 mM EDTA), with a 10 minute incubation shaking at room temperature for each wash, and treated with PK Buffer (1 mg/ml proteinase K, 100 mM EDTA, 0.2% SDS, 1% *N*-laurylsarcosine) for 18 h at 42°C. The plugs were washed once, and treated with PSMF solution (1mM PSMF in wash solution) for 1 h with gentle shaking. The plugs were washed twice in Wash Solution, then once in Storage Solution (0.1X Wash Solution). Plugs were stored at 4C in storage solution.

CHEF Gel Analysis:

A 0.8% Biorad Ultrapure agarose in 0.5x TBE buffer was cast. Chromosomal plugs were introduced into the wells of the cell and sealed with warm agarose. The gel was run in a Bio-Rad CHEF gel electrophoresis

unit using 0.5xTBA buffer at 14°C, for 22 hours at 200 mV and a 10 second to 60 second ramp time.

Citations mentioned in this appendix are included in Chapter 3 References.

APPENDIX D

APPENDIX D

Detailed Methods From Chapter 4.

Strains used are described in Table 4.1. Citations included are References as listed in Chapter 4.

Conjugal Matings and Transposon Mutagenesis:

Mating procedures are outlined in Appendix C. The S17/pSUP2021[Tn5] mutagenesis system was used for random transposon mutagenesis of pJP4 (deLorenzo, 1990). *E. coli* CB1360 (Sm^r) carrying pJP4 was mated with S-17/pSUP2021 on LB agar overnight. The cells were resuspended in minimal medium, plated on a selective media containing 100 μg/ml each streptomycin and kanamycin, and incubated 18 h at 37°C. Colonies on these plates were then replicated onto a fresh selective plate and onto an LB plate previously inoculated with log phase *P. putida* PB2442. The selective plate was grown overnight at 37°C and stored at 4°C as a master plate. The *Alcaligenes* mating plate was incubated at 30°C overnight, then replicated twice onto minimal salts medium containing 20 mM benzoate and 100 g/ml kanamycin. Approximately 10,000 colonies were then screened with the 4-NPAA petri plate assay (Chapter 2), and 100 colonies that appeared to have lost the ability to produce 4-NP from 4-NPAA were retained. These colonies, taken from the untested replicate, were assayed again the following day, after transfer to fresh minimal benzoate with kanamycin plate. Eleven colonies were selected as having lost the ability to convert 4-NPAA to 4-NP. These colonies were streaked

for isolation, then tested for conversion of 4-NPAA in liquid cultures over 24 hours. The colonies were also tested for 4-NPAA conversion in lysates after growth in M2 medium supplemented with 0.1% casamino acids. Of the eleven, one strain, m11, had lost the ability to convert 4-NPAA in whole cells and lysates (presumably a *tfdA* insertion), and two strains, m8 and m22, converted 4-NPAA in lysates but not in whole cells (indicating loss of permeation). The *P. putida* strains were then mated into *A. eutrophus* JMP222 by a broth culture mating (Appendix C).

4-NPAA Assay in Whole Cells:

Strains used in this assay were previously grown on M2 (Table 4.1) media. Overnight broth cultures were used to inoculate fresh M2 media supplemented with 100 mM 4-NPAA (pH 7.0). Cultures were then shaken at 30°C for 24 hours. One ml of the culture was transferred to a microfuge tube and the samples were centrifuged at 14,000 rpm in an Eppendorf microfuge for 1 minute. Supernatants were transferred to a cuvette and 10 µl of 10N NaOH added to enhance the sensitivity of the assay. Absorbances were then taken at 401 nm.

2,4-D Initial Uptake Assay:

This assay was based on the benzoate permease assay of Thayer and Wheelis (1976). Two ml of bacterial cells were grown overnight on 1 mM 2,4-D (or 1 mM casamino acids + 1 mM 2,4-D). Uninduced controls are grown on a similar concentration of a non-aromatic amino acid or TCA intermediate. Overnight cultures were added to 40 ml of the same medium in 250 ml Erlenmeyer flasks and grown to mid-log phase. Cultures were

chilled to 4°C, pelleted by centrifugation at 7,000 x *g* for 10 minutes in pre-weighed centrifuge tubes, and washed twice with 20 ml of 20 mM Tris pH 6.8 -7. The cells were resuspended to a concentration of 0.25 g cell/ml in the same buffer at 4°C. Washed cells were incubated in a shaking water bath at 30°C for 10-15 minutes prior to the addition of ^{14}C-2,4-D (~2.6 μCi/ml). For initial uptake, 100 μl of 10 mM 2,4-D (1/10 ^{14}C-2,4-D) was added to 5.0 ml cells, each concentration done consecutively to allow for accurate sampling. One ml samples were transferred onto pre-wetted 0.45 μm Millipore nitrocellulose filters and immediately washed with 5 ml of washing buffer (same as above supplemented with 75 mM sodium azide and 1mM 2,4-D) at 60, 120, 180, 240, and 300 second intervals. The filters were air dried, immersed in a toluene based scintillation fluid, and counted in a scintillation counter set for ^{14}C detection.

2,4-D Incorporation Assay:

Bacterial cells were inoculated from glycerol stocks and streaked on M2 (Table 2.1), if appropriate, or LB agar plates supplemented with appropriate antibiotics and incubated until colonies were evident. The cells were preconditioned by inoculating 10 ml liquid cultures of M2 supplemented with 20 mM citrate and incubation overnight or until the cells had attained a clearly visible cell density. One ml samples of the preconditioned cultures were measured for cell density by taking the spectrophotometric A_{650}. Using these readings, the initial inocula for the next cultures were equalized. Aliquots of the same media were inoculated to a total volume of 20 ml. The inoculated culture was well mixed and divided into two cultures, forming two duplicate starter cultures. To one of the pair,

2.5 mCi (5 µl) of ^{14}C-2,4-D was added. The cultures were then incubated at 30°C, shaking rapidly. At time 0, one ml of the non-labeled culture was used to measure the cell density (A_{650}). One ml of the radioactive culture was filtered rapidly through a Millipore Nybond 0.45 µm filter using a Millipore Vacuum manifold. The filter was washed 3 times with M2 media. The filter was allowed to dry somewhat as the cell density of the non-labeled culture was measured. The filters were removed, placed in scintillation vials, and 5 ml toluene based scintillation fluid was added (for counting and to stop any possible $^{14}CO_2$ evolution due to further metabolism). The sampling procedure was repeated every fifteen minutes for 150 minutes. The samples were then counted in a scintillation counter using setting appropriate for ^{14}C detection, and the results adjusted for background and absorbance of 2,4-D to the membrane.