This is to certify that the

dissertation entitled

CHARACTERIZATION OF GENES INVOLVED
IN THE DEGRADATION OF CARBON TETRACHLORIDE
BY PSEUDOMONAS STUTZERI STRAIN KC

presented by

Lycely del Carmen Sepúlveda-Torres

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Microbiology

Craig S. Criddle
Major professor

Date 6/19/00

CHARACTERIZATION OF GENES INVOLVED IN THE DEGRADATION OF CARBON TETRACHLORIDE BY *PSEUDOMONAS STUTZERI* STRAIN KC

By

Lycely del Carmen Sepúlveda-Torres

AN ABSTRACT OF A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Microbiology

2000

ABSTRACT

Pseudomonas stutzeri strain KC is a denitrifying aquifer isolate that produces and secretes pyridine-2,6-bis(thiocarboxylic acid), a compound that chelates copper to fortuitously transform carbon tetrachloride without producing chloroform. Although *P. stutzeri* strain KC has been successfully used for full-scale bioremediation of carbon tetrachloride contaminated sites, no information was known about the genes responsible for the carbon tetrachloride degradation capacity. The present dissertation describes the methods used for the generation of four *P. stutzeri* strain KC mutants with a reduced ability to degrade carbon tetrachloride. The DNA interrupted in the mutants was sequenced and analyzed using various tools, allowing the assignment of possible gene functions. The information obtained from the aforementioned studies was combined with information about other genes involved in carbon tetrachloride dehalogenation mutated by another research group, in order to propose a possible biosynthesis pathway for the carbon tetrachloride dehalogenation agent pyridine-2,6-bis(thiocarboxylic acid).

To my family:
Papi, Mami, Jorgito, Tony, Coca, Zulma, Yelitza, Yarelis and Elliot.
Thank you for being the wind beneath my wings.

Para mi familia:
Papi, Mami, Jorgito, Tony, Coca, Zulma, Yelitza, Yarelis y Elliot.
Gracias por ser mi fuente de fortaleza e inspiración.

ACKNOWLEDGMENTS

It is my desire to show appreciation and give credit to the people who made possible the fulfillment of my doctoral dream.

I would like to thank my advisor, Dr. Craig Criddle, for supporting me in my studies to the best of his abilities. I feel very fortunate for having a special relationship with my advisor; one that only a few graduate students experience during their graduate careers. You are my supporting mentor, an enthusiastic colleague and a good friend. Thank you for believing in me, even in the moments when I lost faith in myself. I anticipate having a life-long professional relationship with you.

It is my desire to give a special recognition to Dr. James Tiedje, my co-adviser. Thank you very much for making me your protégé since I came to Michigan State University (MSU) for the first time as a Center for Microbial Ecology (CME) summer intern. Thank you for your letter of support for the National Science Foundation pre-doctoral fellowship, for making me the first Puerto Rican woman to be part of the CME and for been my dedicated co-adviser. It has been a great honor working with you. I am one of the few graduate students who have had the opportunity of working with one of the best scientists of their time. I admire you for your scientific achievements and for been such a gracious human being. I am sure that the National Academy of Science will soon give you the recognition you deserve.

Special thanks to my guidance committee members: Dr. Patrick Oriel, Dr. Robert Brubaker and Dr. Loren Snyder for their helpful criticism and advice. I am especially grateful to Dr. Frans deBruijn, Dr. Michael Dybas and Dr. Joan Broderick for providing good guidance when I needed it the most. Thank you all for sharing your knowledge with me. I wish you success in your future endeavors.

I would like to express my gratitude to the graduate students, post-doctoral fellows, faculty and staff of the Department of Microbiology, the CME and the Environmental Engineering Program for lending a helping hand every time I needed help or advice; and for providing a friendly learning environment. Thanks to Mr. Jorge Rodrigues, Dr. Anne Milcamp, Dr. Sabine Rech, Dr. K. Padmanabhan and Dr. Ann Gustafson for helpful discussions about my work and for providing help whenever I asked.

My deepest appreciation to Dr. Ronald Crawford, Dr. Thomas Lewis, Mr. Marc Cortese, Mr. Jon Sebat and the rest of the Crawford group working on *Pseudomonas stutzeri* strain KC. Thank you very much for sharing my passion about this bacterium and for keeping me informed about your findings. I congratulate you on your well-deserved accomplishments in this project.

Special thanks to the National Science Foundation for a pre-doctoral fellowship and to the CME and the Institute for Environmental Toxicology for providing financial support for my studies.

I am indebted to Dr. and Mrs. Stuart Sleight and the rest of the Sleight family for "adopting" me into their family. Thank you very much for providing me a place that I could truly call home away from home. You made my six years away from my homeland and my loved ones less difficult to bear.

Thanks to María Gutiérrez for reminding me that life is more beautiful when seen through the eyes of a child. I would like to recognize the dear friends I met at Michigan State University: Carmen Medina, Vladimir Ferrer, Dr. Olga Hernández-Patino, Héctor Ayala, Verónica Grüntzig and Dr. Paul R. Martin for helping me keep a warm heart during the cold Michigan winters. I would also like to acknowledge my lifelong friends Leslie Strutton, Melissa Colón and Ivelisse Torres for their support and encouragement across the distance.

I would like to express my deepest gratitude to my parents for their unconditional love and support. Thank you for all the years of sacrifice in order to provide my brothers and I the best education you could afford. Thank you for believing in me and for seeing in my achievements the fulfillment of your own truncated dreams. I love you and I am very proud of you. I would be nothing without you. Thanks to my brothers Jorge Rafael and Juan Antonio, my extended family, and countless neighbors and friends for their love, prayers and encouragement.

Sp

sta

sh

my

for

I v

an

my

sur

La

hel

ste

Special thanks to my husband Elliot for his endless love and support. Even though our status changed during the past five years; one thing remained the same: you were there to share my triumphs and defeats with me. Thank you very much for being my soul mate, my confidant, my lover and my best friend. Your love brings out the best of me. I look forward to the rest of our journey together.

I would like to recognize *Pseudomonas stutzeri* strain KC for being such a fascinating and intriguing organism. You have been the protagonist of my scientific melodrama, but my interest in discovering you was stronger than the frustrations you caused me. I am sure that you will be the subject of study for many other Ph.D. students.

Lastly, I would like to thank the Lord for putting the right people on my path, in order to help me accomplish the goals I have set for myself and for showing me His love in every step I take.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION: *PSEUDOMONAS STUTZERI STRAIN KC* AND THE DEGRADATION OF CARBON TETRACHLORIDE

Carbon Tetrachloride: Characteristics and Problems

Carbon tetrachloride (CCl_4) is a non-polar and nonflammable, colorless solvent. It was widely used in the past in the processing of nuclear fuel, in dry cleaning operations, in the manufacture of fire extinguishers, refrigerants, aerosols, and chlorinated organic compounds and as an extractant, a fumigant and a metal degreaser (Verschueren, 1983). An estimated 2.3 million kg/year of CCl_4 were discharged during manufacture and processing, and approximately 27.2 million kg/year were released as solvent emissions (Sittig, 1985). The widespread use and inadequate disposal of carbon tetrachloride caused the contamination of many groundwater supplies. Carbon tetrachloride is very persistent in groundwater due to its resistance to hydrolysis. The estimated half-life of CCl_4 in aqueous solutions is 7,000 years (Vogel *et al.*, 1987a).

Carbon tetrachloride contamination represents a hazard to groundwater reservoirs. A study performed in the 1980's by the United States Environmental Protection Agency (U.S. E.P.A.) indicated that 29 of 113 (25%) public water supply systems tested have a concentration of up to 400 μg CCl_4/L; exceeding the 5 μg CCl_4/L established standard (Sittig, 1985). One of the best characterized sites highly contaminated with carbon tetrachloride is the Hanford site in southwestern Washington (Last & Rohay, 1991; Skeen *et al.*, 1994). The Hanford site is located in an area of approximately 1,500 km^2 that was selected in the 1940's by the U.S. government for the production of nuclear materials to be used by the United States in World War II. Carbon tetrachloride contaminated solutions were discarded in liquid waste disposal facilities during the plutonium recovery

processes. An estimated 1,000 metric tons of carbon tetrachloride were generated over a 20-year period. Carbon tetrachloride permeated through the vadose zone, contaminating at least 7 km^2 with concentrations over 1,000 times the U.S. E.P.A. standard for drinking water. The U.S. Department of Energy is conducting bioremediation experiments in order to remove carbon tetrachloride from the Hanford site and protect the Colombia River located only miles away.

The degradation of CCl_4 is of special interest, not only for its adverse contributions to the environment, but also for the hazards presented to human health. Carbon tetrachloride is highly toxic and is a suspected carcinogen. It can adversely affect different organs including the eyes, kidneys and liver. Excessive exposure can affect the gastrointestinal tract, while acute exposure can cause serious malfunctioning of the liver that may eventually cause death (Sittig, 1985; Vogel & McCarty, 1987b). The manufacture of CCl_4 was prohibited under the terms of the amended Montreal protocol because CCl_4 is also an ozone-depleting agent (Programme, 1994). Even though CCl_4 is now released at a much-reduced rate, the CCl_4 still present in water reservoirs represents an environmental and a health hazard. Therefore, the decontamination of CCl_4 polluted sites is of great significance.

Me

Seve

depe

disse

revie

Tatar

A con

a trich

The for

reductio

electron

trichloro

degradat

cell ma

significa

1980).

Mechanisms of Carbon Tetrachloride Transformation

Several bacteria can transform carbon tetrachloride to different dechlorination states, depending on their metabolic capabilities and the culture conditions. This section of the dissertation provides an overview of CCl_4 degradation mechanisms. More extensive reviews of the CCl_4 degradation pathways are available elsewhere (Hashsham, 1996; Tatara, 1996).

A common step in the transformation of CCl_4 is the addition of a single electron, yielding a trichloromethyl radical and a chlorine ion, as shown in reaction 1.1.

$$CCl_4 + e^- \longrightarrow \bullet CCl_3 + Cl^- \qquad (1.1)$$

The formation of the trichloromethyl radical is believed to be the rate limiting step in the reductions of alkyl halides (Bakac & Espenson, 1985; Wade & Castro, 1973). The electron source can be anything from reduced metals to enzymes and co-factors. The trichloromethyl radical is highly reactive and will further react to produce a variety of degradation products. The trichloromethyl radical can also react directly with insoluble cell materials to form permanent covalent bonds. The latter process accounts for a significant portion of the transformed CCl_4 products in biological systems (Ahr *et al.*, 1980).

The trichloromethyl radical can dimerize to produce hexachloroethane as described in reaction 1.2. The formation of hexachloroethane from carbon tetrachloride was first observed in 1969 when oral administration of CCl_4 to rabbits caused exhalation of hexachloroethane (Fowler, 1969).

$$2[\bullet CCl_3] \longrightarrow C_2Cl_6 \qquad (1.2)$$

Chloroform (CCl_3) is one of the most common products of CCl_4 transformation. Chloroform can be produced from CCl_4 via the trichloromethyl radical by 2 mechanisms: (i) removal of a hydrogen atom from lipids or proteins (Luke *et al.*, 1987), (ii) hydrogenation of the trichloromethyl radical using reducing power from a cometabolic process. These reactions are labeled 1.3 and 1.4, respectively.

$$\bullet CCl_3 + RH \longrightarrow CHCl_3 + \bullet R \qquad (1.3)$$

$$\bullet CCl_3 + H^+ + e^- \longrightarrow CHCl_3 \qquad (1.4)$$

Hydrogenation of the trichloromethyl radical using reducing power from a cometabolic process has been reported in many microbial systems (see Hashsham, 1996; Tatara, 1996 for extensive lists). Several microbial consortia or isolates can form chloroform from CCl_4. Further reductions to dichloromethane and methylchloride are also possible in highly reduced environments such as methanogenic and sulfate respiring microcosms (Vogel *et al.*, 1987a). Nevertheless, as the degree of chlorination decreases from CCl_4 to

methylch

& McCa

differen

(Butler,

coenzyn

Chloro

with a

is alsc

chloro

(Butle

Carbo

Esche

al., 1

dichl

trans

(Krie

carb

carb

follo

inter

methylchloride, the reaction becomes less energetically and kinetically favorable (Vogel & McCarty, 1987b). Chloroform can also be produced from the reaction of CCl_4 with different biologically active compounds like glutathione, cysteine and ascorbic acid (Butler, 1961), and bacterial transition metal – coenzyme pairs including vitamin B_{12}-Co, coenzyme F_{439}-Ni and hematin-Fe (Gantzer & Wackett, 1991).

Chloroform is not a desirable compound because it is highly recalcitrant and persistent with a half life of hydrolysis at room temperature of 1,850 years (Jeffers *et al.*, 1989). It is also highly toxic and a suspected carcinogen (Sittig, 1985). The formation of chloroform from CCl_4 was first reported in 1961 in the breath of animals given CCl_4 (Butler, 1961).

Carbon disulfide (CS_2) was reported as a minor CCl_4 transformation product of *Escherichia coli* cells grown with fumarate as the source of carbon and energy (Criddle *et al.*, 1990a), and a major transformation product in anaerobic enrichments grown in dichloromethane (Hashsham *et al.*, 1995). Carbon disulfide is a major product of CCl_4 transformation in abiotic systems containing a high concentration of bisulfide ion (HS^-) (Kriegman-King & Reinhard, 1992). If the hydroxide ion is also present in the system, carbon disulfide can be hydrolyzed to carbon dioxide yielding the net formation of carbon dioxide from carbon tetrachloride. Adewuyi and Carmichael (1987) proposed the following sequence of reactions for the formation of carbon dioxide from CCl_4 via a CS_2 intermediate (reactions 1.5 to 1.8).

The in

yield t

Reinha

the pro

final pr

If the

abiotic

peroxy

$$CCl_4 + 2HS^- + 2H^+ \longrightarrow CS_2 + 4HCl \quad (1.5)$$

$$CS_2 + OH^- \longrightarrow CS_2OH^- \text{ (slow)} \quad (1.6)$$

$$CS_2OH^- + OH^- \longrightarrow CSO_2H^- + HS^- \quad (1.7)$$

$$CSO_2H^- + OH^- \longrightarrow CO_3H^- + HS^- \quad (1.8)$$

The interactions between the trichloromethyl radical and various sulfur containing ions to yield thiophosgene (CCl_2S) and CO_2 from CCl_4 was studied by Kriegman-King and Reinhard (1992). As seen in reactions 1.9 to 1.11, trichloromethanethiolate (CCl_3S^-) is the proposed common intermediate. A two step reaction will then produce CO_2 as the final product (reactions 1.12 and 1.13).

$$\bullet CCl_3 + HS^- \longrightarrow CCl_3S^- + \bullet H \quad (1.9)$$

$$\bullet CCl_3 + S_2O_3^{2-} + e^- \longrightarrow CCl_3S^- + SO_3^{2-} \quad (1.10)$$

$$\bullet CCl_3 + S_x^{2-} + e^- \longrightarrow CCl_3S^- + S_{x-1}^{2-} \quad (1.11)$$

$$CCl_3S^- \longrightarrow CCl_2S + Cl^- \quad (1.12)$$

$$CCl_2S + 2OH^- \longrightarrow CO_2 + 2HS^- \quad (1.13)$$

If the trichloromethyl radical is produced under aerobic conditions in mammalian or abiotic systems, molecular oxygen reacts with the trichloromethyl radical to give a peroxy radical as demonstrated in reaction 1.14. The peroxy radical is eventually

trans

inter

the m

Pseudo

Pseudo

7136) is

Californ

formate

under i

Crawfo

transformed to CO_2 by a series of reactions where phosgene (CCl_2O) is a major intermediate (Kubik & Anders, 1980; Shah *et al.*, 1979). Reactions 1.15 to 1.21 depict the main steps in this process, starting with a reduced organic compound (RH).

$$\bullet CCl_3 + O_2 \longrightarrow \bullet CCl_3OO \quad (1.14)$$

$$\bullet CCl_3OO + RH \longrightarrow \bullet CCl_3OOH + \bullet R \quad (1.15)$$

$$\bullet CCl_3OOH + H_2O \longrightarrow CCl_2O + H_2O_2 + HCl \quad (1.16)$$

$$2(\bullet CCl_3OOH) \longrightarrow 2CCl_3O + O_2 \quad (1.17)$$

$$\bullet CCl_3O \longrightarrow CCl_2O + \bullet Cl \quad (1.18)$$

$$\bullet CCl_3O + H^+ + e^- \longrightarrow CCl_3OH \quad (1.19)$$

$$CCl_3OH \longrightarrow CCl_2O + HCl \quad (1.20)$$

$$CCl_2O + H_2O \longrightarrow CO_2 + 2HCl \quad (1.21)$$

***Pseudomonas stutzeri* strain KC**

Pseudomonas stutzeri strain KC (ATCC deposit number 55595, DSM deposit number 7136) is a denitrifying bacterium originally isolated from an aquifer in Seal Beach, California. It can rapidly degrade CCl_4 to CO_2 (~ 50%), a cell associated fraction (~5%), formate (~20%) and unidentified nonvolatile compounds (~ 25%), when CCl_4 is provided under iron-limiting conditions (Criddle *et al.*, 1990b; Dybas *et al.*, 1995; Lewis & Crawford, 1993). *Pseudomonas stutzeri* strain KC was the first isolated bacterium

cap

chlc

rates

react

abou

Even

under

contar

only a

Strain

motile

during

strain k

nitrite

were u

biotech

bioaugr

injected

2006).

two yea

capable of degrading CCl_4 under anoxic conditions with minimal accumulation of chloroform. It is also unique in its ability to degrade CCl_4 faster than the degradation rates observed in the reduced methanogenic microcosms that favor dehalogenation reactions (Krone & Thauer, 1992). Appendix A provides more detailed information about the phylogeny and taxonomy of *Pseudomonas stutzeri* strain KC.

Even though the CCl_4 degradation capacity of *P. stutzeri* strain KC is not fully understood, this organism has proven to be beneficial for the remediation of CCl_4 contaminated sites because the CCl_4 dechlorination reaction is rapid, with half-lives of only a few minutes (Tatara *et al.*, 1995) and occurs without accumulation of chloroform. Strain KC attaches to aquifer sediments, but it can also exist in a free-swimming highly motile form that is chemotactic towards nitrate, and it can sustain dechlorination activity during migration (Witt *et al.*, 1999a; Witt *et al.*, 1999b). Emerson (1999) reported that strain KC reproducibly forms complex colony patterns on agar motility plates containing nitrite or nitrate. Five other species of pseudomonads tested under identical conditions were unable to form such patterns. Strain KC has also assumed great significance for biotechnology because of its use in one of the first full-scale field aquifer bioaugmentation applications. Large volumes of strain KC were grown on-site and injected into a CCl_4–contaminated aquifer in Schoolcraft, Michigan (Hyndman *et al.*, 2000). The resulting biocurtain for CCl_4 degradation has now been maintained for over two years, with efficient removal of CCl_4.

The

mad

Univ

CCl_4

the C

closel

of iro

the tra

1993).

associa

secrete

culture

conditi

system

secreted

microflo

transpor

al., 199

The pat

suggest

of grow

approxi

The CCl_4 degradation capacity of strain KC remained a mystery for many years. Efforts made by the Criddle group at Michigan State University and the Crawford group at the University of Idaho, helped elucidate the conditions that allowed the rapid degradation of CCl_4 by *P. stutzeri* strain KC. Results of the first experiments performed to characterize the CCl_4 degradation capacity of strain KC indicated that the degradation of CCl_4 is closely related to the amount of trace elements available in the culture medium. Addition of iron and cobalt inhibit the transformation, but trace amounts of copper are needed for the transformation to occur (Criddle *et al.*, 1990b; Lewis & Crawford, 1993; Tatara *et al.*, 1993). The degradation of CCl_4 by *P. stutzeri* strain KC requires cell-membrane associated reducing power that is independent of the electron transport chain and a small secreted supernatant factor that is produced when the bacterium is grown in iron limiting cultures (Tatara, 1996). This secreted factor is produced under aerobic and anaerobic conditions, but the transformation of CCl_4 only occurs when oxygen is removed from the system (Tatara *et al.*, 1993). Carbon tetrachloride transformation also occurs when the secreted factor is combined with actively respiring bacteria, yeast cells, or aquifer microflora (Tatara *et al.*, 1993; Tatara *et al.*, 1995). The secreted factor can be transported without retardation through aquifer sediments and agar (Sepúlveda-Torres *et al.*, 1999; Tatara *et al.*, 1995).

The pattern of growth of *P. stutzeri* strain KC and the production of the secreted factor suggested that the factor is involved in trace-nutrient delivery. A very rapid initial phase of growth is observed in which no secreted factor production is detected. However, at approximately 12 hours of growth, the rate of total protein production slows and the

secr

due

linki

no C

1990

transf

secon

limitir

(Cridd

known

secrete

1995),

can deg

that sid

(Hohnad

Experim

interme

(1995) r

formate

form a t

the form

secreted factor production accelerates. This pattern suggests that the factor is produced due to the limitation of a required nutrient or metal (Tatara *et al.*, 1993). Observations linking the iron scavenging hypothesis to CCl_4 transformation included the following: (i) no CCl_4 transformation occurs if iron is present in the cell-culture medium (Criddle *et al.*, 1990b), (ii) addition of iron to grown cultures of *P. stutzeri* strain KC inhibited CCl_4 transformation, possibly by competing for a binding site (Tatara *et al.*, 1993), (iii) the second order rate coefficient for CCl_4 transformation increased for cells grown in iron-limiting media (Tatara *et al.*, 1993), (iv) CCl_4 transformation was also inhibited by cobalt (Criddle *et al.*, 1990b; Lewis & Crawford, 1993; Lewis & Crawford, 1999a) which is known to form stable complexes with siderophores (Ekkehardt *et al.*, 1990), (v) the secreted factor had an apparent molecular weight of less than 500 Daltons (Dybas *et al.*, 1995), which correlates with the molecular weights of siderophores, (vi) other cell types can degrade CCl_4 when the secreted factor is provided (Tatara *et al.*, 1995). It is known that siderophores that are produced by one microorganism can be utilized by others (Hohnadel & Meyer, 1988).

Experiments with $^{14}CCl_4$ were performed to determine the identity of the degradation intermediates of the strain KC mediated CCl_4 dehalogenation process. Dybas *et al.* (1995) reported the formation of radioactive formate from $^{14}CCl_4$. The formation of formate requires a two-electron transfer to the CCl_4 molecule. The first electron will form a trichloromethyl radical (reaction 1.1). The addition of a second electron results in the formation of a dichlorocarbene radical (reaction 1.22) that can subsequently be

hyd

Carb

Lewis

electr

static

trichlo

to prod

molecu

formati

N-2-hy

medium

products

hydrolyzed to formate and carbon monoxide, as illustrated in reactions 1.23 and 1.24. Carbon monoxide formation was not tested in the Dybas *et al.* report.

$$\bullet CCl_3 + e^- \longrightarrow :CCl_2 + Cl^- \quad (1.22)$$

$$:CCl_2 + 2H_2O \longrightarrow CH_2O_2 + 2HCl \quad (1.23)$$

$$:CCl_2 + H_2O \longrightarrow CO + 2HCl \quad (1.24)$$

Lewis and Crawford (1995, 1999a) postulated that CCl_4 degradation occurs via a one-electron reduction pathway. They detected phosgene in $^{14}CCl_4$ experiments performed in static serum bottles containing an oxygenic atmosphere. Molecular oxygen reacts to the trichloromethyl radical to produce a peroxy radical (reaction 1.25) that could further react to produce phosgene as indicated in reactions 1.26 to 1.29). Phosgene is a highly reactive molecule that readily hydrolyzes to CO_2 in aqueous solutions (reaction 1.30). The formation of phosgene from CCl_4 was proven by the addition of trapping agents such as N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) or cysteine to the culture medium, followed by the subsequent identification of the corresponding condensation products (reactions 1.31 and 1.32).

$$\bullet CCl_3 + O_2 \longrightarrow \bullet OOCCl_3 \quad (1.25)$$

$$[\bullet OOCCl_3] \longrightarrow 2\ CCl_3O\bullet + O_2 \quad (1.26)$$

$$CCl_3O\bullet \longrightarrow COCl_2 + Cl\bullet \quad (1.27)$$

Thioph

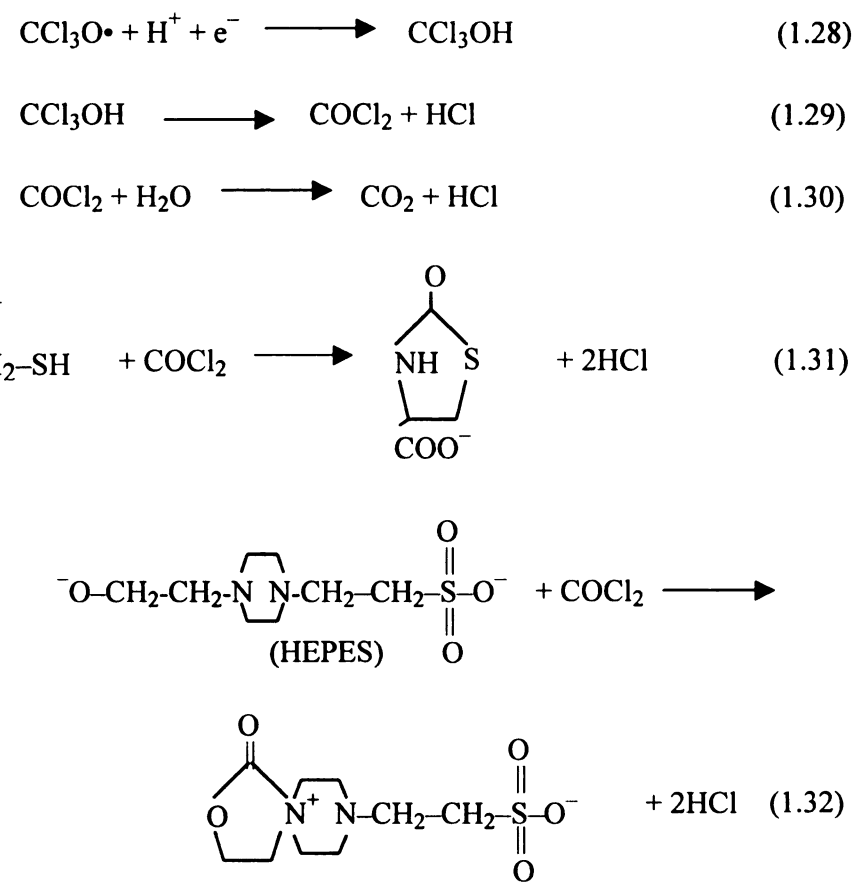
Thioph

(reactio

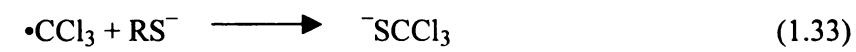
thiophos

(DMED

correspo



Thiophosgene was identified in similar experiments performed under anoxic conditions. Thiophosgene is formed when the trichloromethyl radical reacts with a sulfur species (reactions 1.33 – 1.34). It hydrolyzes to produce CO_2 (reaction 1.35). The formation of thiophosgene from CCl_4 was proven by the addition of N,N'-dimethylethylenediamine (DMED) to the culture medium, followed by the subsequent identification of the corresponding condensation product (reaction 1.36).



Lewis

$^{14}CCl_4$

reporte

excess

them to

of *P. st*

Dybas

accumul

the tota

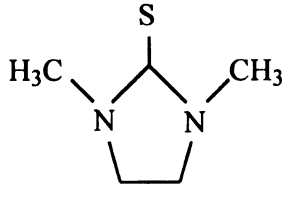
usually

mediate

$$^{-}SCCl_3 \longrightarrow CSCl_2 + Cl^{-} \quad (1.34)$$

$$CSCl_2 + H_2O \longrightarrow CO_2 + HCl \quad (1.35)$$

$CH_3–NH–CH_2–CH_2–NH–CH_3$ (DMED) $+ CSCl_2 \longrightarrow$



$+ 2HCl \quad (1.36)$

Lewis and Crawford (1999a) also looked for the formation of radioactive formate from $^{14}CCl_4$. ^{14}C-formate accounted for approximately 5% of the total $^{14}CCl_4$, as previously reported by Dybas *et al.* (1995). The addition of formate at more than 300-fold molar excess to the CCl_4 addition had no significant effect on the total CO_2 production allowing them to conclude that formate is indeed a final product of the CCl_4 degradation pathway of *P. stutzeri* strain KC. This result confirms that the two-electron pathway proposed by Dybas *et al.* occurs to some extent. The Crawford group observed chloroform accumulation in their cultures. Nevertheless, chloroform corresponds to less than 4% of the total carbon in all their cultures, indicating that the trichloromethyl radical is not usually involved in hydrogen abstraction reactions in the CCl_4 dehalogenation processes mediated by *P. stutzeri* strain KC.

The

Degr

The s

recen

(Figu

secret

et al.,

from t

ability

trichlor

PDTC.

degrade

showed

result is

that iron

partially

indicate

inactiva

The Discovery of Pyridine-2,6-bis(thiocarboxylic acid) as a Carbon Tetrachloride Degrading Compound

The secreted factor responsible for the CCl_4 degradation capacity of *P. stutzeri* KC was recently identified as pyridine-2,6-bis(thiocarboxylic acid) (PDTC) (CAS # 69945-42-2) (Figure 1.1) (Lee *et al.*, 1999). This compound was originally described as a molecule secreted by *Pseudomonas putida* DSM 3601 grown under iron-limited conditions (Ockels *et al.*, 1978). In assays using Ti(III) citrate, both synthetic PDTC and PDTC isolated from the supernatant of strain KC cells grown under iron-limiting conditions, had the ability to degrade CCl_4. Other chlorinated compounds such as chloroform, trichloromonofluoromethane (CCl_3F) or 1,1,1-trichloroacetic acid did not react with PDTC. Similar results were seen in assays when PDTC was added to cells that do not degrade CCl_4 on their own (Lee *et al.*, 1999). PDTC coupled to iron or copper also showed CCl_4 degradation capacity when added to living cells or Ti(III) citrate. This result is consistent with the previous observation made by Dybas *et al.* (1995) showing that iron addition at the reaction level did not inhibit the CCl_4 degradation capacity of partially purified secreted factor. The experiments performed by Lee *et al.* (1999) also indicate that PDTC can be used as a catalyst with a limited turnover, suggesting inactivation upon reaction with CCl_4.

Figu

PDTC

cobalt

Holm,

previo

illustra

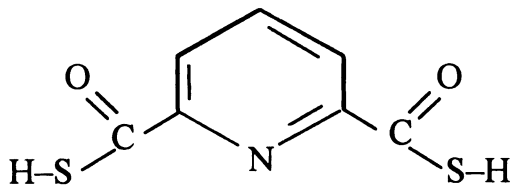


Figure 1.1 Chemical structure of pyridine-2,6-bis(thiocarboxylic acid) (PDTC).

PDTC is known to form a variety of complexes with transition metals including iron, cobalt, nickel and palladium (Espinet *et al.*, 1994; Hildebrand *et al.*, 1984; Krüger & Holm, 1990). The redox potential of Fe-PDTC and Ni-PDTC complexes have been previously described (Hildebrand *et al.*, 1984; Krüger & Holm, 1990). Figure 1.2 illustrates the change in oxidation state of the iron center in Fe-PDTC complexes.

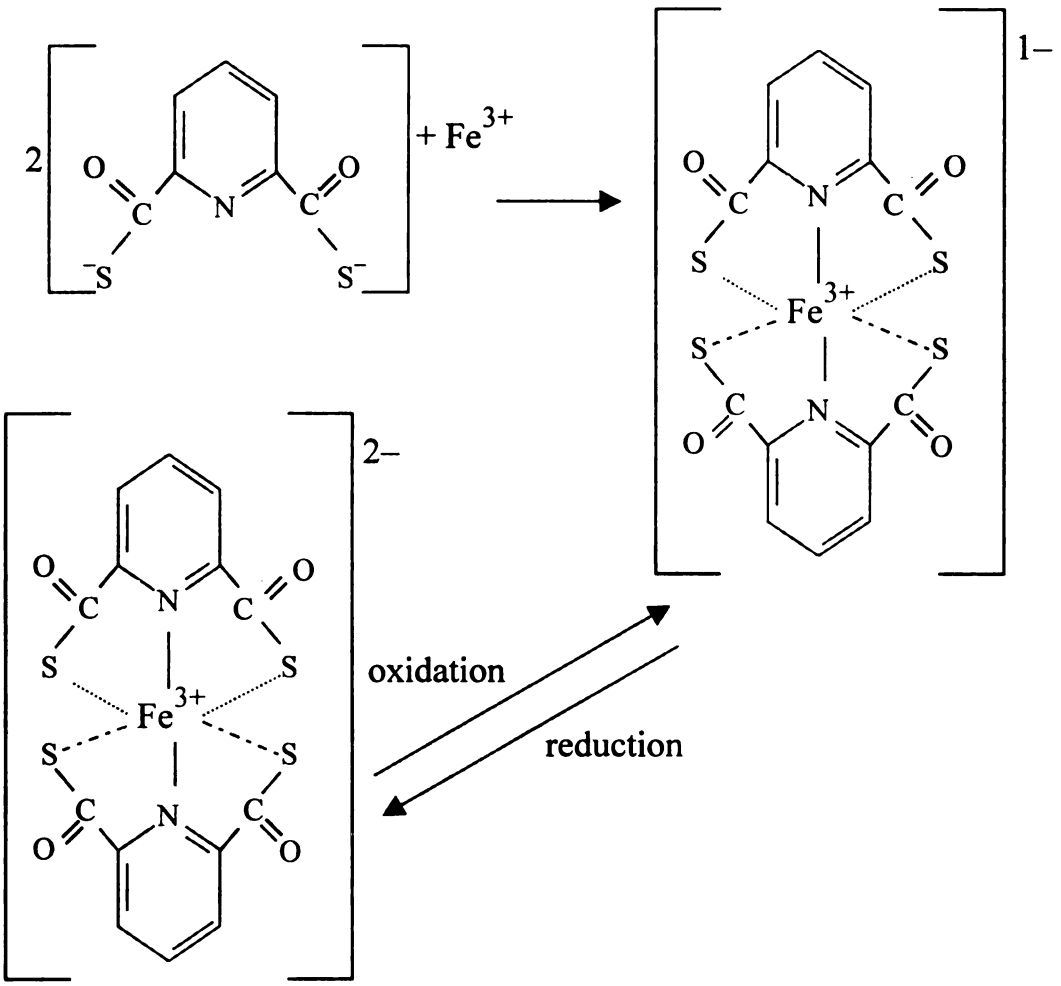


Figure 1.2 Mechanism for iron chelation by pryridine-2,6-bis(thiocarboxylic acid) as proposed by Ockels *et al.* (1978). Adapted from Lee *et al.*(1999).

Upon the discovery of PDTC as the CCl_4 dechlorination agent, Lewis and collaborators (1999b) studied the CCl_4 degradation capacity of synthetic PDTC - metal complexes and presented their work in a preliminary report. Free acid PDTC as well as Fe-PDTC, Cu-PDTC and Ni-PDTC conferred the CCl_4 degradation capacity to *P. stutzeri* ATCC 17588, a bacterium incapable of degrading CCl_4 on its own. On the other hand, Co-PDTC was

ina

inh

19

abl

PD

wa

Res

sho

wh

oth

red

pro

CS_2

The

eluc

The

the

form

radic

inactive. This result is consistent with previous reports showing cobalt-mediated inhibition of the CCl_4 degradation pathway (Criddle *et al.*, 1990b; Lewis & Crawford, 1993). When Na_2S was used as a reductant, only the free acid PDTC and Cu-PDTC were able to remove CCl_4 from the medium. When no reductant was added, 1 molecule of Cu-PDTC reacted with 2 molecules of CCl_4 whereas the removal mediated by the free-acid was less efficient and more erratic.

Results of experiments with $^{14}CCl_4$ in the presence of PDTC with no added reductants showed a carbon distribution similar to the one observed in experiments performed with whole cultures: 70% CO_2, 20% volatile material, 10% non-volatile products. On the other hand, more volatile compounds were observed when Na_2S was added as a reductant. Carbon disulfide (CS_2) and carbonyl sulfide (COS) were identified as volatile products obtained from the PDTC-mediated degradation of $^{13}CCl_4$ and the production of CS_2 was directly correlated to the amount of NaS_2 added to the reaction mixture.

The mechanism of the Cu-PDTC mediated CCl_4 degradation process has not been elucidated and the possible mechanism provided in this section is based on speculation. The data available to date suggest that copper (probably Cu^{1+}) binds to PDTC, allowing the activation of the electron-dense sulfur to donate an electron to CCl_4, causing the formation of a trichloromethyl radical and a pyridine-2,6-bis(thiocarboxylate) thyil radical (Lewis, 2000) (Figure 1.3). The thyil radical may be stabilized with resonance on

the

that

trich

bour

Phos

1.3A

sulfu

thioc

dipic

other

intact

The t

availa

1.3C.

degra

be pro

or pro

trichlo

simulta

form th

chlorot

1.3C w

the copper atom and the function of Cu may be restricted to provide the right conditions that would allow sulfur activation (Broderick, 2000). If the thyil radical and the trichloromethyl radical are in close proximity, they may recombine forming a sulfur-bound trichloromethyl complex that eventually produces phosgene and thiophosgene. Phosgene and thiophosgene are unstable intermediates that hydrolyze to CO_2 (Figure 1.3A). If trichloromethylthiol is an intermediate of the CCl_4 degradation pathway, the sulfur-carbon bonds of PDTC are destroyed, forming a carboxyl group. If both thiocarboxylic acid groups are converted into carboxyl groups, PDTC is transformed into dipicolinic acid, a compound that is unable to catalyze the degradation of CCl_4. On the other hand, if trichloromethanol is produced instead, the thiocarboxylic bond remains intact, allowing the possible regeneration of PDTC.

The trichloromethyl radical is highly reactive and may interact with other compounds available in the medium and may lead to the production of CO_2 as indicated in Figure 1.3C. Of special interest is the lack of chloroform formation during the PDTC-mediated degradation of CCl_4. As indicated in a previous section of this chapter, chloroform may be produced from the trichloromethyl radical by extracting a hydrogen atom from lipids or proteins or by reacting with a proton and an electron (equations 1.3 and 1.4). If the trichloromethyl and the pyridine-2,6-bis(thiocarboxylate) thyil radicals are formed simultaneously, as suggested in Figure 1.3, these radicals may recombine immediately to form the sulfur-bound trichloromethyl group (Figure 1.3A), decreasing the possibility of chloroform formation. These assumptions suggest that the reactions presented in Figure 1.3C would account only for a minimal fraction of the CO_2 produced.

If the

exter

dichlo

imme

first id

al. (19

dichlor

The pa

power

CCl_4 v

prepara

donated

externa

center.

degrad

are inv

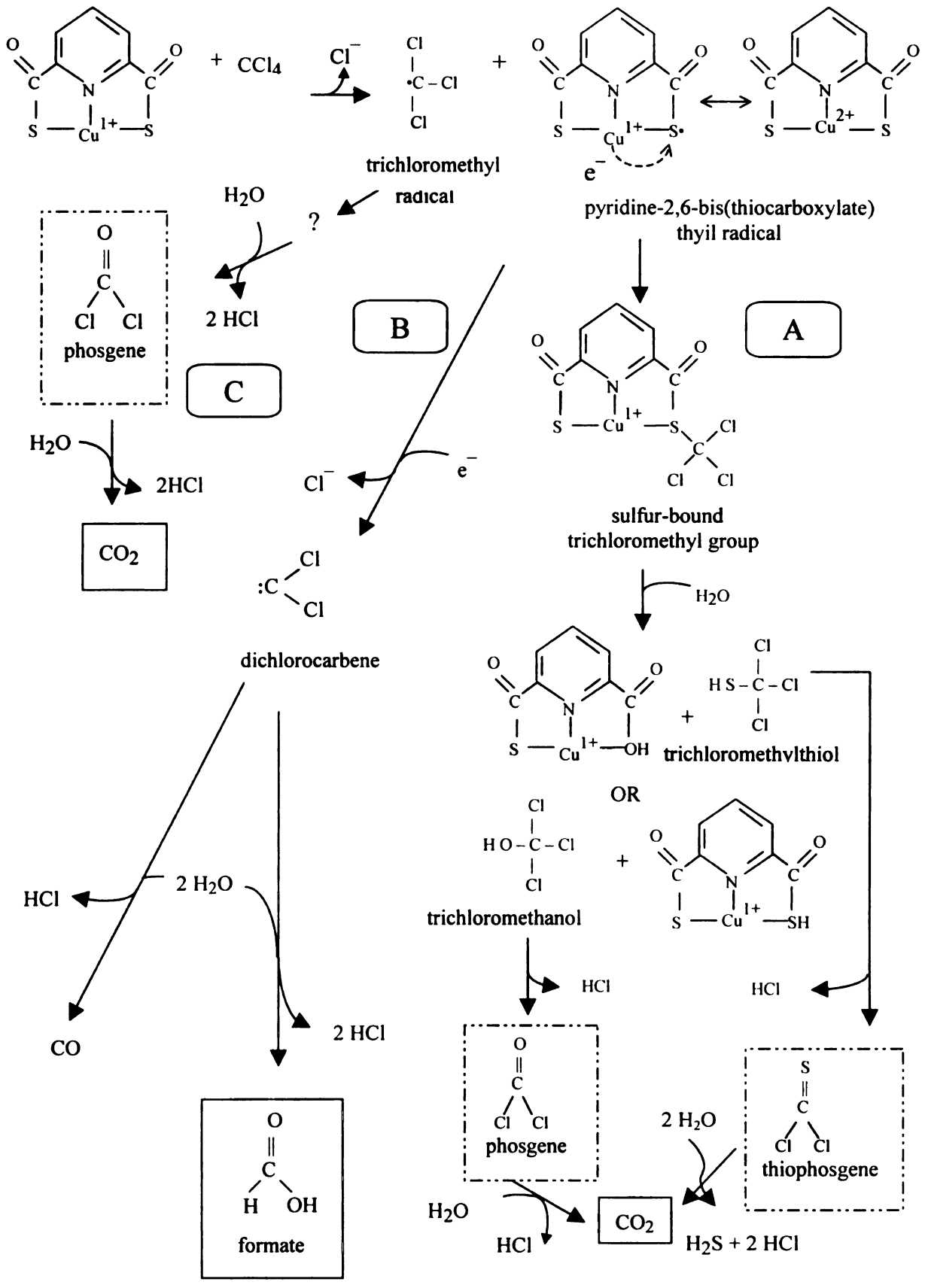
degrad

degrad

If the trichloromethyl radical receives a second electron (perhaps from copper or from an external electron donor), it would trigger the elimination of chloride and the formation of dichlorocarbene (Figure 1.3B). If dichlorocarbene is formed at some extent, it immediately reacts with water to produce carbon monoxide and formate. Formate was first identified as a final product of the strain KC-mediated CCl_4 degradation by Dybas *et al.* (1995) and was later confirmed by Lewis and Crawford (1999a), suggesting that dichlorocarbene formation may occur at some extent.

The pathway presented in Figure 1.3 does not account for the role of external reducing power in the degradation of CCl_4. Tatara (1996) demonstrated that the degradation of CCl_4 was enhanced when the reaction occurred in the presence of cell membrane preparations amended with NADH + H^+. He also demonstrated that the electrons donated to CCl_4 did not come from the electron transport chain. A possible role of external reducing power may be to maintain the proper oxidation state of the metal center. More experiments are needed in order to elucidate how Cu-PDTC catalyzes the degradation of CCl_4, what are the intermediates of the process and what external factors are involved in the degradation pathway. The discovery of Cu-PDTC as the CCl_4 degrading agent is only another step in the elucidation of the fascinating CCl_4 degradation capacity of *Pseudomonas stutzeri* stain KC.

Figure 1.3 Proposed pathway for the degradation of CCl_4 by Cu-PDTC. A) Formation of CO_2 via the condensation of a trichloromethyl radical and a thyil radical. B) Formation of formate and carbon monoxide from a dichlorocarbene intermediate. C) Formation of CO_2 by the interaction of the trichloromethyl radical with unknown compounds. Final products identified in the dechlorination pathway are indicated by solid boxes. Known pathway intermediates are indicated by dashed boxes. Both sulfurs present in PDTC can theoretically participate in the reactions presented in branch A. Only one cycle is presented for the benefit of simplicity. This figure was modified from (Tatara, 1996) and (Lewis, 2000).


+ CCl4
Cl⁻
trichloromethyl radical
pyridine-2,6-bis(thiocarboxylate) thyil radical
e⁻
H2O
?
2 HCl
phosgene
B
A
C
H2O
2HCl
CO2
Cl⁻
e⁻
sulfur-bound trichloromethyl group
dichlorocarbene
H2O
trichloromethylthiol
OR
trichloromethanol
HCl
HCl
HCl
2 H2O
2 HCl
CO
phosgene
2 H2O
thiophosgene
H2O
HCl
CO2
H2S + 2 HCl
formate

Outli

The p

the C

how n

provid

mutant

P. stut

involv

provid

for PD

may er

for effi

how *P.*

Outline of This Dissertation

The primary objective of this research project was to characterize the genes that confer the CCl_4 degradation capacity to *Pseudomonas stutzeri* strain KC. Chapter 2 describes how mutants unable to degrade CCl_4 were created by transposon mutagenesis. It also provides some information about phenotypic characterization of the aforementioned mutants. Chapter 3 provides an exhaustive genetic analysis of the genes mutated in the *P. stutzeri* strain KC mutants described in Chapter 2. Information about other genes involved in CCl_4 degradation, mutated independently by another research group, is also provided. The information available about all the genes was used to propose a pathway for PDTC biosynthesis. Lastly, Chapter 4 provides suggestions for future research that may enhance our knowledge of the physiological role of PDTC and how it could be used for effective remediation of CCl_4 contaminated sites. Experiments for the elucidation of how *Pseudomonas stutzeri* strain KC controls PDTC synthesis are also provided.

Ade

Ahr,

Baka

Brode

Butler

Criddl

Criddl

Dybas,

REFERENCES

Adewuyi, Y. G. & Carmichael, G. R. (1987). Kinetics of hydrolysis and oxidation of carbon disulfide by hydrogen peroxyde in alkaline medium and application to carbonyl sulfide. *Environ Sci Technol* **21,** 170-177.

Ahr, H. J., King, L. J., Nastainczyk, W. & Ullrich, V. (1980). The mechanism of chloroform and carbon monoxide formation from tetrachloromethane by microsomal cytochrome P-450. *Biochem Pharmachol* **29,** 2855-2861.

Bakac, A. & Espenson, J. H. (1985). Kinetics and mechanism of the alkylnickel formation in one-electron reductions of alkyl halides and hydroperoxides by a macrocyclic nickel (I) complex. *J Am Chem Soc* **108,** 713-719.

Broderick, J. (2000). Personal communication, Michigan State University.

Butler, T. C. (1961). Reduction of carbon tetrachloride *in vivo* and reduction of carbon tetrachloride and chloroform *in vitro* by tissue constituents. *J Pharmaco Exp Theory* **134,** 311.

Criddle, C. S., DeWitt, J. t. & McCarthy, P. L. (1990a). Reductive dehalogenation of carbon tetrachloride by *Escherichia coli* K-12. *Appl Environ Microbiol* **56,** 3247-3254.

Criddle, C. S., DeWitt, J. T., Grbic-Galic, D. & McCarthy, P. L. (1990b). Transfromation of carbon tetrachloride by *Pseudomonas* sp. strain KC under denitrification conditions. *Appl Environ Microbiol* **56,** 3240-3246.

Dybas, M. J., Tatara, G. M. & Criddle, C. S. (1995). Localization and chracterization of the carbon tetrachloride transformation activity of *Pseudomonas* sp. strain KC. *Appl Envrion Microbiol* **61,** 758-762.

Ekkehardt, H. F., McMurry, T. J., Hugi, A. & Raymond, K. N. (1990). Coordination chemistry of microbial iron transport: structural and spectroscopic characterization of diastereometric Cr(III) and Co(III) complexes of desferriferrithiocin. *J Am Chem Soc* **112,** 1854-1860.

Emerson, D. (1999). Complex pattern formation by *Pseudomonas* strain KC in response to nitrate and nitrite. *Microbiology* **145,** 633-641.

Espinet, P., Lorenzo, C. & Miguel, J. A. (1994). Palladium complexes with the tridentate dianionic ligand pyridine-2,6-bis(thiocarboxylate), PDTC. Crystal structure of (n-Bu_4N)[Pd(pdtc)Br]. *J Am Chem Soc* **33,** 2052-2055.

Fowler, S. J. L. (1969). carbon tetrachloride transformation in the rabbit. *Br J Pharm* **37,** 733-737.

Gantzer, C. J. & Wackett, L. P. (1991). Reductive dechlorination catalyzed by bacterial transition-metal coenzymes. *Environ Sci Technol* **25,** 715-722.

Hashsham, S. A., Scholze, R. & Freedman, D. L. (1995). Cobalamin-enhanced anaerobic biotransformation of carbon tetrachloride. *Environ Sci Technol* **29,** 2856-2863.

Hashsham, S. A. (1996). Cobalamin-enhanced anaerobic biotransformation of carbon tetrachloride. Doctoral Dissertation, University of Illinois, Urbana-Champaign.

Hildebrand, U., Lex, J., Taraz, K., Winkler, S., Ockels, W. & Budzikiewicz, H. (1984). Untersuchungen zum redox-system bis(pyridin-2,6-dicarbothioato)-ferrat(II)/-ferrat(III). *Z Naturforsch (B)* **39B,** 1607-1613.

Hohnadel, D. & Meyer, J. M. (1988). Specificity of pyoverdin-mediated iron uptake among fluorescent *Pseudomonas* strains. *J Bacteriol* **170,** 4865-4873.

Hyn

Jeffe

Krieg

Krone

Krüge

Kubik

Last,

Lee, C

Hyndman, D. W., Dybas, M. J., Forney, L., Heine, R., Mayotte, T., Phanikumar, M. S., Tatara, G., Tiedje, J., Voice, T., Wallace, R., Wiggert, D., Zhao, X. & Criddle, C. S. (2000). Hydraulic characterization and design of a full scale biocurtain. *Ground Water* **In Press.**

Jeffers, P. M., Ward, L. M., Woytowitch, L. M. & Wolfe, N. L. (1989). Homogeneous hydrolysis rate constants for selected chlorinated methanes, ethanes, ethenes, and propanes. *Environ Sci Tech* **23,** 965-969.

Kriegman-King, M. R. & Reinhard, M. (1992). Transformation of carbon tetrachloride in the presence of sulfide, biotite, and vermiculite. *Environ Sci Technol* **26,** 2198-2206.

Krone, U. E. & Thauer, R. K. (1992). Dehalogenation of trichlorofluoromethane (CFC-11) by *Methanosarcina barkeri. FEMS Microbiol Lett* **90,** 210-214.

Krüger, H. J. & Holm, R. H. (1990). Stabilization of trivalent nickel in tetragonal NiS_4N_2 and NiN_6 environments: synthesis, structures, redox potentials, and observations related to [NiFe]-hydrogenases. *J Am Chem Soc* **112,** 2955-2963.

Kubik, V. L. & Anders, M. W. (1980). Methabolism of carbon tetrachloride to phosgene. *Life Sciences* **26,** 2151-2155.

Last, G. & Rohay, V. (1991). Carbon tetrachloride contamination, 200 West Area, Hanford Site. *PNL-SA-19564, Pacific Northwest Laboratory, Richland, Washington* .

Lee, C. H., Lewis, T. A., Paszczynski, A. & Crawford, R. L. (1999). Identification of an extracellular catalyst of carbon tetrachloride dehalogenation from

Lew

Lew

Lewi

Lewis

Lewis,

Luke,

Pseudomonas stutzeri strain KC as pyridine-2,6-bis(thiocarboxylate). *Biochem Biophys Res Commun* **261,** 562-566.

Lewis, T. A. & Crawford, R. L. (1993). Physiological factors affecting carbon tetrachloride dehalogenation by the denitrifying bacterium *Pseudomonas* sp. strain KC. *Appl Environ Microbiol* **59,** 1635-1641.

Lewis, T. A. & Crawford, R. L. (1995). Transformation of carbon tetrachloride via sulfur and oxygen substitution by *Pseudomonas* sp. strain KC. *J Bacteriol* **177,** 2204-2208.

Lewis, T. A. & Crawford, R. L. (1999a). Chemical studies of carbon tetrachloride transformation by *Pseudomonas stutzeri* strain KC. In *Novel Approaches for Bioremediation of Organic Pollution* (ed. R. Fass, Y. Flashner and S. Reuveny), pp. 1-11. Plenum Publishers, New York.

Lewis, T. A., Paszczynski, A., Lee, C. H. & Crawford, R. L. (1999b). The dechlorination agent for *Pseudomonas stutzeri* strain KC toward carbon tetrachloride is pyridine-2,6-bis(thiocarboxylic acid). In *Pseudomonas* '99, Maui, Hawaii.

Lewis, T. (2000). Personal communication, University of Vermont.

Luke, B. T., Loew, G. H. & McClean, A. D. (1987). Theorethical investigations of theanaerobic reduction of halogenated alkanes by cytochrome p450. I. Structures, inversion barriers, and heats of formation of halomethyl radical. *J Am Chem Soc* **109,** 1307-1317.

Ockels, W., Römer, A. & Budzikiewicz, H. (1978). An Fe(II) complex pryridine-2,6-di-(monothiocarboxylic acid) - a novel bacterial metabolic product. *Tetrahedron Lett* **36,** 3341-3342.

United Nations Environmental Programme (1994). Adjustment to the Montreal Protocol on substances that deplete the ozone layer. In *Principles of International Environmental Law.: 2A. Documents in International Environmental Law* (ed. P. Sands, R. Tarasofsky and M. Weiss), pp. 208-223. Manchester University Press, Manchester.

Sepúlveda-Torres, L. del C., Rajendran, N., Dybas, M. J. & Criddle, C. S. (1999). Generation and initial characterization of *Pseudomonas stutzeri* KC mutants with impaired ability to degrade carbon tetrachloride. *Arch Microbiol* **171,** 424-429.

Shah, H., Hartman, S. P. & Weinhous, S. (1979). Formation of carbonyl chloride in carbon tetrachloride metabolism by rat liver *in vitro*. *Cancer Research* **39,** 3942-3947.

Sittig, M. (1985). Handbook of Toxic and Hazardous Chemicals and Carcinogens. Noyes Publications, New York.

Skeen, R., Amos, K. & Petersen, J. (1994). Influece of nitrate concentration on carbon tetrachloride transformation by a Ddenitrifying microbial consortium. *Water Res* **28,** 2433-2438.

Tatara, G. M., Dybas, M. J. & Criddle, C. S. (1993). Effects of medium and trace elements on kinetics of carbon tetrachloride transforamtion by *Pseudomonas* sp. strain KC. *Appl Environ Microbiol* **59,** 2126-2131.

Tatara, G. M., Dybas, M. J. & Criddle, C. S. (1995). Biofactor mediated transformation of carbon tetrachloride by diverse cell types. In *Bioremediation of Chlorinated Solvents*, vol. 3(4) (ed. R. E. Hinchee, A. Leeson and L. Semprini), pp. 69-76. Battelle Press, Columbus, Ohio.

Tatara, G. M. (1996). Physiology of carbon tetrachloride transformation by *Pseudomonas stutzeri* KC. Doctoral Dissertation, Michigan State University.

Verschueren, K. (1983). Handbook of Environmental Data on Organic Chemicals, pp. 341-344. Van Nostrand Reinhold, New York.

Vogel, T. M., Criddle, C. S. & McCarty, P. L. (1987a). Transformations of halogenated aliphatic compunds. *Environ Sci Technol* **21,** 722-726.

Vogel, T. M. & McCarty, P. L. (1987b). Abiotc and biotic transformations of 1,1,1-trichloroethane aliphatic compounds. *Environ Sci Technol* **21,** 1208-1213.

Wade, R. S. & Castro, C. E. (1973). Oxidation of iron (II) porpyrins by alkyl halides. *J Am Chem Soc* **95,** 226-230.

Witt, M. E., Dybas, M. J., C, W. D. & Criddle, C. S. (1999b). Use of bioaugmentation for continuous removal of carbon tetrachloride in model aquifer columns. *J Environ Eng Sci* **16,** 475-485.

Witt, M. E., Dybas, M. J., Worden, R. M. & Criddle, C. S. (1999a). Motiltity-enhanced bioremediation of carbon tetrachloride-contaminated aquifer sediments. *Environ Sci Technol* **33,** 2958-2964.

CHAPTER 2

GENERATION AND INITIAL CHARACTERIZATION OF *PSEUDOMONAS STUTZERI* KC MUTANTS WITH IMPAIRED ABILITY TO DEGRADE CARBON TETRACHLORIDE

Published in *Archives of Microbiology* (1999) 171: 424-429

Arch M

ORI

Lycel
Nara
Craig

Gen
mut

Receiv

Abstr
stutzer
that tra
volatil
membr
types a
genera
strain
of agar
screene
the pla
combin
but fou
nonvol
degrade
rescens
and we
C-labe
rescens
CT* co
activate
colonie
observe
four repe
lated by

L. Del C
Departm
East Lan

L. Del C
Center fo
East Lan

N. Raier
Departm
Michigar

C. S. Cr
Departme
Stanford
e-mail:
Tel:

Arch Microbiol (1999) 171:424–429

ORIGINAL PAPER

Lycely Del C. Sepúlveda-Torres
Narayanan Rajendran · Michael J. Dybas
Craig S. Criddle

Generation and initial characterization of *Pseudomonas stutzeri* KC mutants with impaired ability to degrade carbon tetrachloride

Received: 23 November 1998 / Accepted: 15 March 1999

Abstract Under iron-limiting conditions, *Pseudomonas stutzeri* KC secretes a small but as yet unidentified factor that transforms carbon tetrachloride (CT) to CO_2 and nonvolatile products when activated by reduction at cell membranes. *Pseudomonas fluorescens* and other cell types activate the factor. Triparental mating was used to generate kanamycin-resistant *lux*::Tn*5* recombinants of strain KC. Recombinants were streaked onto the surface of agar medium plugs in microtiter plates and were then screened for carbon tetrachloride degradation by exposing the plates to gaseous ^{14}C-carbon tetrachloride. CT$^+$ recombinants generated nonvolatile ^{14}C-labeled products, but four CT$^-$ recombinants did not generate significant nonvolatile ^{14}C-labeled products and had lost the ability to degrade carbon tetrachloride. When colonies of *P. fluorescens* were grown next to colonies of CT$^+$ recombinants and were exposed to gaseous ^{14}C-carbon tetrachloride, ^{14}C-labeled products accumulated around the *P. fluorescens* colonies, indicating that the factor secreted by CT$^+$ colonies had diffused through the agar and become activated. When *P. fluorescens* was grown next to CT$^-$ colonies, little carbon tetrachloride transformation was observed, indicating a lack of active factor. Expression of *lux* reporter genes in three of the CT$^-$ mutants was regulated by added iron and was induced under the same iron-limiting conditions that induce carbon tetrachloride transformation in the wild-type.

Key words Transposon mutagenesis · Carbon tetrachloride · Biotransformation · Biodegradation · Luciferase · *Pseudomonas stutzeri* KC · Reporter genes · Mutants

Abbreviations *CT* Carbon tetrachloride · *Km* Kanamycin · *LB* Luria broth · *RLU* Relative light units · *Rf* Rifampicin · *TSB* Tryptic soy broth

L. Del C. Sepúlveda-Torres
Department of Microbiology, Michigan State University, East Lansing, MI 48824, USA

L. Del C. Sepúlveda-Torres · M. J. Dybas · C. S. Criddle
Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA

N. Rajendran · C. S. Criddle
Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI 48824, USA

C. S. Criddle (✉)
Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA
e-mail: criddle@ce.stanford.edu,
Tel.: +1-650-723-9032, Fax: +1-650-725-9474

Introduction

Carbon tetrachloride (CT) is a suspected carcinogen that also causes acute liver toxicity in animals (Sittig 1985). Its production has been banned under the terms of the amended Montreal protocol because it is an ozone-depleting agent (United Nations Environmental Programme 1994). Most denitrifying organisms that degrade carbon tetrachloride do so slowly with the accumulation of chloroform, a compound that can be even more persistent than carbon tetrachloride (Egli et al. 1988; Semprini et al. 1992).

Pseudomonas stutzeri strain KC is an aquifer isolate that transforms carbon tetrachloride to carbon dioxide (Criddle et al. 1990; Lewis and Crawford 1993; Tatara et al. 1993), formate (Dybas et al. 1995), and other nonvolatile products without the formation of chloroform (Criddle et al. 1990; Lewis and Crawford 1993; Tatara et al. 1993). Rapid carbon tetrachloride transformation requires a small (500 Da) factor secreted by exponential-stage strain KC cells grown under Fe^{3+}-limiting conditions, together with actively growing cells capable of regenerating the secreted factor. The transformation occurs only under anoxic conditions, although the factor is produced under both oxic and denitrifying conditions (Dybas et al. 1995). Organisms that usually do not degrade carbon tetrachloride are able to do so when the factor is provided (Tatara et al. 1995). Of special interest are the genes re-

quired for production, secretion, and activation of the factor. To identify these genes, methods are needed for rapid and efficient screening of large numbers of mutants.

Most of the microorganisms that degrade volatile, halogenated hydrocarbons produce CO_2 and nonvolatile products (Jain and Criddle 1995). The nonvolatile products accumulate or are utilized for growth (Nielsen 1990; Chaudhry and Chapalamadugu 1991; Fetzner and Lingens 1994; Jain and Criddle 1995). When ^{14}C-labeled hydrocarbons are degraded, nonvolatile ^{14}C-labeled products are often generated (Wackett and Gibson 1983; Lewis and Crawford 1993; Sepúlveda et al. 1997). To obtain mutants with impaired ability to degrade carbon tetrachloride, we screened a large number of *P. stutzeri* KC *lux*::Tn*5* mutants for failure to produce nonvolatile ^{14}C-labeled products from ^{14}C-carbon tetrachloride. Mutants were then characterized for *lux* expression by assaying luciferase under different growth conditions.

Materials and methods

Strains and plasmids

P. stutzeri strain KC (DSM 7136, ATCC 55595), isolated from an aquifer in Seal Beach (Calif., USA) (Criddle et al. 1990), was maintained on nutrient agar. *Pseudomonas fluorescens* (ATCC 13525) was obtained from the Department of Microbiology, Michigan State University. Rifampicin-resistant (Rf^R) *P. stutzeri* KC strains were selected in our laboratory on nutrient agar plates containing rifampicin (100 μg/ml) from a set of spontaneous Rf^R strain KC mutants. An Rf^R strain KC isolate was used in the triparental matings described below in order to counterselect for the *Escherichia coli* donor and helper strains. One Rf^R strain, KC137, was selected for subsequent studies because its carbon tetrachloride degradation capability matched that of the wild-type. All *Pseudomonas* strains were grown at 20–25 °C with constant shaking (150 rpm). *E. coli* DH5α containing pRL1063a, a kanamycin-resistant (Km^R) transposon delivery plasmid for a *lux*AB::Tn*5* construct (Wolk et al. 1991), and *E. coli* containing the Km^R helper plasmid pRK2013 (Ditta et al. 1980) were grown in Luria Broth medium (Sambrook et al. 1989) with 35 μg Km/ml at 37 °C and with shaking at 200 rpm.

Chemicals and media

Nonradioactive carbon tetrachloride (99% pure) was purchased from Aldrich Chemical (Milwaukee, Wis., USA). ^{14}C-carbon tetrachloride (> 99% pure; 250 μCi, specific activity of 4.3 mCi/mmol) purchased from NEN Dupont Research Products (Boston, Mass., USA) was dissolved in iso-octane to a concentration of 0.136 μmol/μl (1.4 μCi/μl) and was stored at –20 °C.

Precipitate-free, iron-free Simulated Groundwater Medium (SGM) was prepared as described by Dybas et al. (1995) with some modifications. The modified medium used half of the sodium hydroxide and potassium phosphate specified in the original procedure, and Fe-free TN2 trace element solution (Criddle et al. 1990) instead of SGM trace element solution. The medium was prepared in acid-washed glassware, adjusted to an initial pH of 8.2, autoclaved at 121 °C for 25 min, transferred to a laboratory bench for quiescent settling of precipitate, and decanted after 24 h. The precipitate-free decanted medium was reautoclaved for 25 min at 121 °C and cooled before use. Precipitate-free SGM contained 23 mM acetate, 19 mM nitrate, and 0.1 M phosphate as determined using a Dionex model 2000i-SP ion-chromatography system (Sunnyvale, Calif., USA). No iron was detected using a Perkin Elmer (Norwalk, Conn., USA) model 3110 graphite furnace atomic absorption spectroscopy system.

Anoxic medium D adjusted to an initial pH of 8.0 was prepared in acid-washed glassware as previously described (Criddle et al. 1990). Noble agar, nutrient agar, nutrient broth, and tryptic soy broth (TSB) were obtained from Difco (Detroit, Mich., USA).

Transposon mutagenesis

Triparental matings (Simon et al. 1983) were used to mobilize plasmid pRL1063a into Rf^R *P. stutzeri* KC137. *E. coli* DH5α containing the donor plasmid pRL1063a and *E. coli* DH5α containing the helper plasmid pRK2013 were grown overnight in LB medium supplemented with 35 μg Km/ml at 37 C and with shaking at 200 rpm. These cultures were transferred to fresh LB without antibiotics using a 4% (by vol.) inoculum and were then grown at the same temperature and shaking speed for 4 h (final cell concentration, ~10^9 cfu/ml). Cultures of the recipient strain KC137 were grown under aerobic conditions for 16 h in TSB or LB containing 35 μg Rf/ml, or under denitrifying conditions for 2 days in medium D containing 35 μg Rf/ml (final cell concentration, ~10^9 cfu/ml). All cultures were washed twice and then resuspended to the original volume using sterile 0.9% NaCl solution. Mixes were prepared by combining 100 μl of *E. coli* DH5α (pRL1063), 100 μl of *E. coli* DH5α (pRK2013), and 500 μl of strain KC137 in sterile microcentrifuge tubes. Individual strains and 100 μl/100 μl pairwise mixes were included as controls to ensure that Km^R/Rf^R double mutants were not present in the mating plates. Fifty microliters of each mixture was spotted on 13-mm filters (containing pores of 0.45 μm in diameter; Millipore, Bedford, Mass., USA) on nutrient agar plates. The plates were incubated at 37 °C for 10 h and then at room temperature for 10 h. The filters were transferred to microcentrifuge tubes containing 1 ml of a sterile 0.9% NaCl solution, and the bacteria grown on the filter were resuspended in the buffer. Resuspended cells were distributed in 200-μl aliquots onto nutrient agar plates containing 50 μg Km/ml and 50 μg Rf/ml and were then incubated at room temperature for 3 days.

Microtiter plate assay for ^{14}C-carbon tetrachloride degradation

Sterile microtiter plates (96 wells) were filled with 200 μl of medium D (pH 8) containing 15 g Noble agar l^{-1} and 25 μg Km/ml. Recombinants obtained by triparental mating were streaked on the agar surfaces of the microtiter plate wells using sterile toothpicks. Inoculated microtiter plates were transferred to a 3.8-l steel paint can (Freund Can Company, Chicago, Ill., USA) containing clean glass marbles to minimize the gas volume within the can (working vol. ≈ 1.9 l). Air in the can was replaced by a 95% nitrogen/5% hydrogen mixture by passing the open can through the interlock of an anaerobic glove box (Coy Laboratories, Detroit, Mich., USA). A sealed 12 × 32-mm glass vial containing 1.4 μmol (14 μCi) of ^{14}C-carbon tetrachloride was attached to the inner wall of the container. The septum of the vial was punctured with a needle, releasing ^{14}C-carbon tetrachloride into the can, and the can was immediately sealed and removed from the glove box for incubation. Plates inside the steel can were exposed to vapor phase ^{14}C-carbon tetrachloride for 5 days at 20 °C and then were vented overnight in a chemical hood. Agar plugs from the microtiter plates were transferred to 10 ml of Safety-solve scintillation cocktail (RPI, Prospect, Ill., USA) and were assayed for nonvolatile radioactivity for 3 min with a 1500 Tri-carb liquid scintillation counter (Packard Instrument, Downers Grove, Ill., USA).

The recombinants that did not produce significant ^{14}C-labeled nonvolatile products in the microtiter plate assay were analyzed by gas chromatography (GC) to confirm loss of carbon tetrachloride degradation capabilities. Recombinants were grown overnight in 5 ml TSB containing 35 μg Km/ml. These cultures were used as a 1% inoculum (by vol.) for 10 ml of anoxic medium D (pH 8) vials containing 35 μg Km/ml and 350 ng of sterile carbon tetrachloride. Cultures were incubated upside down for a week under anoxic conditions at 20 °C and with shaking at 150 rpm. The carbon tetrachloride remaining in the vials was assayed by GC as described by Tatara et al. (1993).

Table 1 Results of the ^{14}CT microtiter plate assay and the gas chromatography assay

Strain	dpm per well in microtiter plate assay	% CT degraded after 1 week in gas chromatography assay[c]
Noninoculated controls	82 ± 15[a]	9 ± 7
Wild-type *Pseudomonas stutzeri* KC	2,218 ± 437[b]	100 ± 0
Recombinants capable of degrading CT in both assays		
KC300	1,416	100 ± 0
Recombinants with impaired ability to degrade CT in both assays		
KC657	465	28 ± 26
KC1896	511	18 ± 8
KC2753	291	28 ± 8
KC3164	459	21 ± 13

[a] Average ± one standard deviation for seven replicates
[b] Average ± one standard deviation for nine replicates
[c] Average ± one standard deviation for triplicates

Petri plate assay for secreted factor production

By itself, the secreted factor produced by strain KC does not reliably transform carbon tetrachloride; however, reliable transformation is observed when it is combined with viable whole cells (Dybas et al. 1995). Tatara (1996) has demonstrated that carbon tetrachloride transformation also occurs when the secreted factor is added to crude cell membrane preparations supplemented with NADH. Decreased transformation rates were observed when the factor was added to crude cell membranes lacking NADH, with little or no transformation when the secreted factor was added to cytoplasmic fractions. Both cell membranes and NADH were required for maximum activation. Of interest is the fact that *P. fluorescens* and many other cell types can mediate activation of the secreted factor (Tatara et al. 1995). In order to determine if the CT$^-$ mutants were impaired for secreted factor production, secretion, or activation, CT$^-$ mutants and cells of *P. fluorescens* were grown on Petri plates containing solid medium D (pH 8). Colonies of CT$^-$ mutants and *P. fluorescens* were grown on Petri plates in an alternating "checkerboard" layout with each colony separated by 1.5 cm from adjacent colonies of the other organism. Petri plates were incubated following the same protocol used for the microtiter plates. After the incubation period, 1-cm^2 agar squares including and surrounding each colony were transferred to 10 ml scintillation cocktail and were counted by the liquid scintillation counter.

Expression of *lux* genes in strain KC mutants

CT$^-$ and selected CT$^+$ recombinants were grown in 5 ml TSB containing 35 μg Km/ml for 24 h at 25 °C and with shaking at 150 rpm. These cultures were used as 0.5% inocula (v/v) for 25-ml cultures of TSB or Fe-free, precipitate-free SGM containing 35 μg Km/ml supplemented with 0, 5, 10, or 20 μM Fe^{3+} as $FeNH_4(SO_4)_2 \cdot 12H_20$. Cultures were grown for 29 h at 25 °C and with shaking at 150 rpm, and then were used as 0.5% inocula (v/v) for 100-ml cultures containing the same antibiotic and Fe^{3+} concentrations. Growth of these cultures was monitored over a 48-h period by periodically removing 1-ml aliquots and measuring optical density at 600 nm.

The light emission assay used to detect luciferase activity was performed on a Berthold Lumat LB 9501 luminometer (EG&G Wallac, Gaithersburg, Md., USA) by combining 5-μl culture aliquots with 50 μl of a solution containing 20 mg bovine serum albumin/ml and 1 μl N-decyl aldehyde/ml. Samples were vortexed for 30 s and assayed for light emission (relative light units) for 1 min.

Results

Generation of mutants with impaired ability to degrade carbon tetrachloride

Three thousand five hundred recombinants were obtained from 44 independent triparental mating events. Most recombinants accumulated roughly the same level of ^{14}C-labeled nonvolatile products in the agar plugs as wild-type cells (1,500–3,000 dpm). Only 38 recombinants accumulated fewer than 1,000 dpm: 4 were auxotrophs that could not grow in the defined medium used for the GC assay; 30 grew poorly in solid medium D, but grew and degraded carbon tetrachloride in liquid medium; and 4 were classified as putative CT$^-$ mutants based on their growth in liquid and solid media and on their inability to degrade significant CT in either assay. A PCR probe specific to strain KC was used to confirm that the putative CT$^-$ mutants were strain KC cells (Dybas et al. 1998). Table 1 summarizes results for noninoculated controls, the wild-type, one of the recombinants that retained the ability to degrade carbon tetrachloride (strain KC300), and the four mutants that failed to transform carbon tetrachloride in either assay (CT$^-$ mutants). Although the CT$^-$ mutants were only tested once in the microtiter plate assay, the results of the GC assay provided statistically significant proof that these mutants were impaired in carbon tetrachloride degradation.

Table 2 Petri plate assay of secreted factor production by wild-type and recombinant KC colonies grown in the presence of *Pseudomonas fluorescens* and ^{14}C- carbon tetrachloride. *P. fluorescens* was grown alone or in an alternating checkerboard configuration with wild-type KC, KC300, KC657, KC1896, KC2753, or KC3164. All colonies were separated by 1.5 cm. Radioactivity is reported as disintegrations per minute per square centimeter of agar for eight replicates

Incubation condition	dpm/cm^2 (avg ± sd)
No-cell control	42 ± 5
P. fluorescens alone	198 ± 35
P. fluorescens+wild-type strain KC	1,583 ± 335
P. fluorescens+autoclaved wild-type strain KC	215 ± 32
P. fluorescens+KC300 (CT$^+$)	1,905 ± 179
P. fluorescens+KC657 (CT$^-$)	95 ± 13
P. fluorescens+KC1896 (CT$^-$)	84 ± 13
P. fluorescens+KC2753 (CT$^-$)	114 ± 9
P. fluorescens+KC3164 (CT$^-$)	102 ± 14

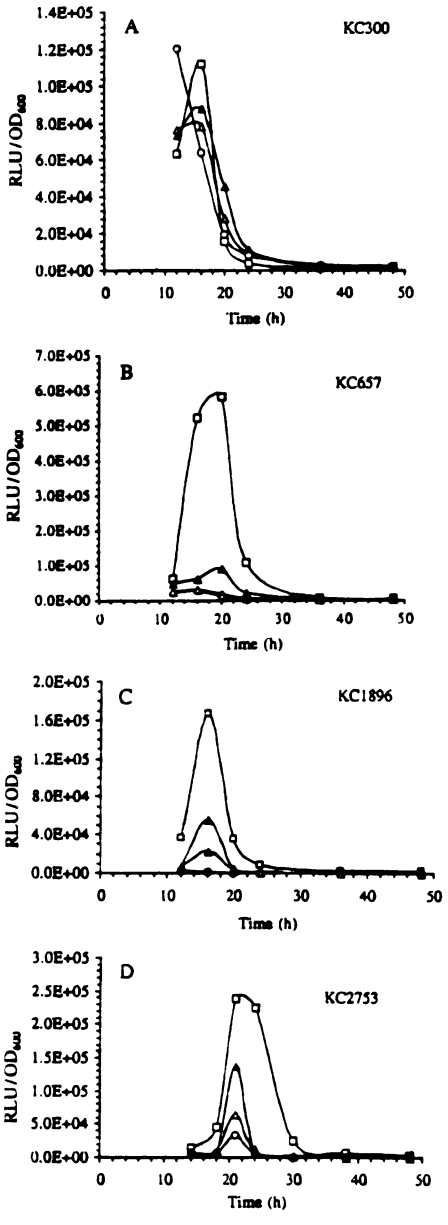


Fig. 1 Luciferase activity, expressed as relative light units divided by OD_{600}, during growth of **A** strain KC 300, **B** strain KC 657, **C** strain KC 1896, and **D** strain KC 2753 in Simulated Groundwater Medium containing 1 □, 5 ▲, 10 △, and 20 ○ μM Fe^{3+}. Due to the low values of OD_{600} prior to 12 h of growth and in 0 μM Fe^{3+}, this ratio was only computed for samples after 12 h of growth and for Fe^{3+} concentrations > 0 μM. Plotted values are the average of two independently grown cultures. Data is not shown for strain KC 3164, which exhibited very low *lux* expression and no clear pattern with increasing iron concentration (similar to strain KC 300)

Evidence that the CT⁻ mutants were impaired in secreted factor production

Table 2 summarizes the results of experiments using an agar-based assay to detect secreted factor production. When *P. fluorescens* was grown in the presence of wild-type strain KC or strain KC *lux*::Tn*5* recombinants capable of degrading carbon tetrachloride (such as strain KC300), the radioactivity that accumulated in the agar plugs around the *P. fluorescens* colonies was more than sevenfold the amount that accumulated when *P. fluorescens* was grown alone (Table 2). When *P. fluorescens* was grown in the presence of any of the putative CT⁻ mutants, the nonvolatile radioactivity was only twice that of noninoculated controls. This indicates that the four putative CT⁻ mutants either do not produce significant levels of CT-degrading factor or secrete a largely inactive form of it.

Expression of *lux* genes in CT⁻ mutants

There was no significant difference between the time course for growth of *lux*::Tn*5* recombinants that were impaired in carbon tetrachloride degradation (e.g., strain KC657) and the time course for growth of strain KC300 and many other recombinants that retained carbon tetrachloride degradation activity. The maximum optical density was 0.8 for SGM cultures and 1.6 for TSB cultures after 20 h (data not shown). However, for all of the recombinants except strain KC3164, addition of Fe^{+3} resulted in dramatically reduced luciferase activity (Fig. 1). This effect was not observed with CT⁺ recombinants such as strain KC300 (Fig. 1 A). Complex media supported low levels of *lux* expression in the recombinants (Fig. 2), but defined mineral medium supported high levels of expression in three of the four CT⁻ mutants (Fig. 1 B–D). For cells grown in SGM containing only 1 μM Fe^{3+}, maximal *lux* expression was sevenfold the level achieved by cells grown in TSB. In contrast, the CT⁺ recombinant strain

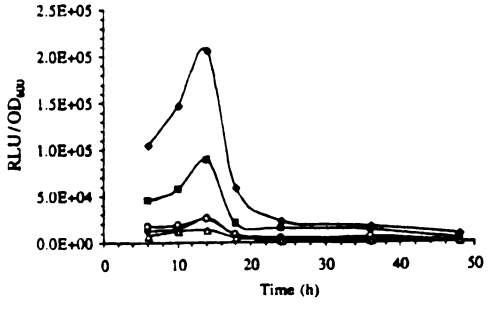


Fig. 2 Luciferase activity, expressed as relative light units divided by OD_{600}, during growth of strain KC 300 ◆, strain KC 657 ■, strain KC 1896 ●, strain KC 2753 ○, and strain KC 3164 △ in tryptic soy broth. Plotted values are the average of two independently grown cultures

KC300 had strong *lux* expression in SGM regardless of Fe^{3+} concentration (Fig. 1 A), and expression in TSB was only 1.7-fold the level achieved during growth in SGM.

Discussion

The ^{14}C-carbon tetrachloride assay was a reliable and efficient means of rapidly screening large numbers of recombinants for carbon tetrachloride degradation. GC analysis confirmed that four mutants were impaired in carbon tetrachloride degradation. GC confirmation was important because some recombinants grew poorly on agar, giving artificially low counts. It is also important because a mutant might degrade carbon tetrachloride without generating nonvolatile products, and such a mutant would be incorrectly classified as CT^-.

An interesting result was the finding that the factor secreted by strain KC could diffuse through agar and still transform carbon tetrachloride when "activated" by viable whole cells such as colonies of *P. fluorescens* (Table 2). Elevated levels of nonvolatile ^{14}C-labeled products were obtained for *P. fluorescens* colonies adjacent to colonies of wild-type strains KC or KC300. Low levels of nonvolatile ^{14}C-labeled products were obtained for *P. fluorescens* colonies adjacent to colonies of CT^- mutants (strains KC657, KC1896, KC2753, and KC 3164). This result indicates that the mutants were defective in production and/or secretion of the secreted factor.

The *lux*::Tn*5* construct on pRL1063a generates a promoterless *lux* transcriptional fusion. Therefore, *lux* is expressed only when the native promoter of the interrupted gene is activated. For three of the four CT^- mutants (strains KC657, KC1896, and KC2753), *lux* expression studies confirmed that the genes required for carbon tetrachloride degradation are expressed under Fe^{3+}-limiting conditions in the late exponential phase of growth (Fig. 1). Mutant KC3164 did not show any significant luciferase expression because the transposon inserted in the wrong orientation with respect to the native promoter, as indicated by subsequent DNA sequence analysis (data not shown). The Fe^{3+}-dependent response of the CT^- transformants was not an artifact introduced by the transposon. CT^+ mutant KC300 did not show significant differences in *lux* expression when grown at different iron concentrations (Figs. 1 A, 2), presumably because the transposon inserted itself in a gene that is not regulated by Fe^{3+} availability. For the putative CT^- mutants, light emission was greater for cultures grown in defined mineral medium SGM (Fig. 1) than for cultures grown in complex medium TSB (Fig. 2). In contrast, light emission for the CT^+ mutant KC300 was greater in complex media (Figs. 1 A, 2). Further details related to the identity of the genes interrupted in the CT^- strains will be provided in a separate publication.

The conditions that induce luciferase expression in the CT^- strains correspond to conditions that induce secreted factor production in the wild-type strain (Tatara 1996). Expression of the *lux* genes was inversely correlated to the amount of Fe^{3+} added for CT^- mutants (Fig. 1). This pattern was consistent with previous reports indicating that Fe^{3+} addition stimulated cell growth but reduced carbon tetrachloride degradation (Criddle et al. 1990; Tatara et al. 1995).

This is the first report of mutants with impaired ability to degrade carbon tetrachloride. Initial characterization indicates that these mutants either do not produce the carbon-tetrachloride-degrading factor or they produce a nonfunctional form of it. The observed interaction of the secreted factor with other cell types and its apparent ability to obtain reducing equivalents thereby suggest a possible role in cell/cell communication. On the other hand, its known regulation by ferric iron and cobalt suggests a possible role in trace metal acquisition, perhaps not unlike that recently reported for the cytochrome secreted by *Geobacter sulfurreducens* (Seeliger et al. 1998). Further studies are under way to characterize the genes inactivated by the *lux*::Tn*5* transposon in the strain KC mutants. Sequencing and identification of the interrupted genes is expected to provide insight into the identity and physiological role of the carbon-tetrachloride-transforming secreted factor and its mechanism of transformation.

Acknowledgements L. Sepulveda-Torres thanks the National Science Foundation for a predoctoral fellowship. We thank F. DeBruijn and A. Milcamps for providing the *lux*::Tn*5* delivery plasmids and for their helpful advice. Special thanks to M. Tomashow and A. Gustafson for providing assistance with the luciferase expression assays. We are grateful to L. Dybas and J. Nguyen for helping with the ion chromatography work and the atomic absorption determination, respectively. This work was supported by the National Science Foundation Center for Microbial Ecology (grant no. BIR-9120006) and by the NIEHS Superfund Basic Research Program of the Institute for Environmental Toxicology (grant no. ES04911).

References

Chaudhry GR, Chapalamadugu S (1991) Biodegradation of halogenated organic compounds. Microbiol Rev 55:59–79

Criddle CS, DeWitt JT, Grbić-Galić D, McCarty PL (1990) Transformation of carbon tetrachloride by *Pseudomonas sp.* strain KC under denitrification conditions. Appl Environ Microbiol 56:3240–3246

Ditta G, Stanfield S, Corbin D, Helinski DR (1980) Broad host range cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc Natl Acad Sci USA 77:7347–7351

Dybas MJ, Tatara GM, Criddle CS (1995) Localization and characterization of the carbon tetrachloride transformation activity of *Pseudomonas* sp strain KC. Appl Environ Microbiol 61:758–762

Dybas MJ, et al (1998) Pilot-scale evaluation of bioaugmentation for *in-situ* remediation of a carbon tetrachloride-contaminated aquifer. Environ Sci Technol 32:3598–3611

Egli C, Tschan T, Scholtz R, Cook AM, Leisinger T (1988) Transformation of tetrachloromethane to dichloromethane and carbon dioxide by *Acetobacterium woodii*. Appl Environ Microbiol 54:2819–2823

Fetzner S, Lingens F (1994) Bacterial dehalogenases: biochemistry, genetics, and biotechnological applications. Microbiol Rev 58:641–685

Jain M, Criddle CS (1995) Metabolism and cometabolism of halogenated C-1 and C-2 hydrocarbons. In: Singh VP (ed) Biotransformations: microbial transformations of health risk compounds. Elsevier, Amsterdam, The Netherlands, pp 65–102

Lewis TA, Crawford RL (1993) Physiological factors affecting carbon tetrachloride dehalogenation by the denitrifying bacterium *Pseudomonas* sp. strain KC. Appl Environ Microbiol 59: 1635–1641

Lewis TA, Crawford RL (1995) Transformation of carbon tetrachloride via sulfur and oxygen substitution by *Pseudomonas* sp. strain KC. J Bacteriol 177: 2204–2208

Nielsen AH (1990) The biodegradation of halogenated organic compounds. J Appl Bacteriol 69: 445–470

Sambrook J, Fritsch EL, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Seeliger S, Cord-Ruwisch R, Schink B (1998) A periplasmic and extracellular c-type cytochrome of *Geobacter sulfurreducens* acts as a ferric ion reductase and as an electron carrier to other acceptors or to partner bacteria. J Bacteriol 180: 3686–3691

Semprini L, Hopkins GD, McCarty PL, Roberts PV (1992) In-situ transformation of carbon tetrachloride and other halogenated compounds resulting from biostimulation under anoxic conditions. Environ Sci Technol 26: 2454–2461

Sepúlveda-Torres L del C, Dybas MJ, Criddle CS (1997) Solid phase bioremediation of carbon tetrachloride by *Pseudomonas stutzeri* strain KC. In: Alleman BC, Leeson A (eds) The Fourth International *in-situ* and On-site Bioremediation Symposium. Battle Press, Columbus, pp 33–37

Simon R, Priefer U, Pühler A (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram-negative bacteria. Biotechnology 1: 784–790

Sittig M (1985) Handbook of toxic and hazardous chemicals and carcinogens, 2nd edn. Noyes, New York

Tatara GM (1996) PhD Thesis, Michigan State University, MI, USA

Tatara GM, Dybas MJ, Criddle CS (1993) Effects of medium and trace metals on kinetics of carbon tetrachloride transformation by *Pseudomonas* sp. strain KC. Appl Environ Microbiol 59: 2126–2131

Tatara GM, Dybas MJ, Criddle CS (1995) Biofactor-mediated transformation of carbon tetrachloride by diverse cell types. In: Hinchee RE, Leeson A, Semprini L (eds) Bioremediation of chlorinated solvents. Battle Press, Columbus, pp 69–76

United Nations Environment Programme (1994) Adjustment to the Montreal protocol on substances that deplete the ozone layer. In: Sands P, Tarasofsky R, Weiss M (eds) Principles of international environmental law. 2A. Documents in international environmental law. Manchester University Press, Manchester, pp 208–223

Wackett LP, Gibson DT (1983) Expression of naphthalene oxidation in *Escherichia coli* results in the biosynthesis of indigo. Science 222: 167–169

Wolk CP, Cai Y, Panoff JM (1991) Use of a transposon with luciferase as a reporter to identify environmental responsive genes in a cyanobacterium. Proc Natl Acad Sci USA 88: 5355–5359

CHAPTER 3

SEQUENCE AND ANALYSIS OF THE GENES INTERRUPTED IN FOUR *PSEUDOMONAS STUTZERI* STRAIN KC MUTANTS WITH IMPAIRED ABILITY TO DEGRADE CARBON TETRACHLORIDE

ABSTRACT

As indicated in Chapter 2, transposon mutagenesis was used to create isolates of *Pseudomonas stutzeri* strain KC unable to degrade carbon tetrachloride. The present chapter presents the strategies used to sequence the genes interrupted in four mutants with impaired CCl_4 degradation capabilities. The DNA sequence was analyzed using various tools, allowing the creation of a possible biosynthesis pathway for pyridine-2,6-bis(thiocarboxylate), a compound that chelates copper to fortuitously degrade carbon tetrachloride.

INTRODUCTION

The primary objective of this research project was to identify genes involved in the degradation of carbon tetrachloride (CCl_4) by *Pseudomonas stutzeri* strain KC. In order to achieve that goal, 3,500 transpositional mutants of stain KC were created via tri-parental matings using a transposon Tn*5* containing promoterless luciferase reporter genes from *Vibrio fischeri* (Wolk *et al.*, 1991). Four mutants named KC657, KC1896, KC2753 and KC3164 showed an impaired CCl_4 degradation ability. Three of the four mutants strongly expressed luciferase under iron-limiting conditions and gene expression was attenuated when cells were grown in cultures containing iron. These experiments are described in detail in Chapter 2 of this dissertation (Sepúlveda-Torres *et al.*, 1999).

The agent responsible for CCl_4 degradation was recently identified as pyridine-2,6-bis(thiocarboxylate) (PDTC) chelated to copper (Lee *et al.*, 1999). PDTC is a strong metal chelator that was first described in 1978 as a metabolite produced by a *Pseudomonas putida* strain grown under iron-limiting conditions (Ockels *et al.*, 1978). A recent report by Lewis *et al.* (2000) described the characterization of a *P. stutzeri* strain KC mutant that lost the ability to produce PDTC and degrade CCl_4 upon the spontaneous deletion of a 170 kb fragment. This mutant was named CTN1. The CCl_4 degradation capacity of strain CTN1 was restored when a 25,746 bp piece of DNA (called T31, GenBank accession number AF196567) containing 16 predicted ORFs, was introduced back into the cell using the wide host range cosmid pRK311.

The present chapter explains the methodology utilized to determine the sequence and possible functions of the genes mutated in the four strains described in Chapter 2. I also analyzed the sequence of the genes independently identified by Lewis *et al.* (2000) in order to generate my own conclusion of how the genes sequenced by Lewis are involved in the CCl_4 degradation process. I analyzed the mutants generated by both research groups in order to explain how *P. stutzeri* strain KC synthesizes pyridine-2,6-bis(thiocarboxylate) (PDTC) when grown under iron-limiting conditions.

MATERIALS AND METHODS

Organisms and culture conditions. Bacteria and plasmids used in this study are listed in Table 3.1. *Pseudomonas* strains were propagated aerobically at 25 °C and 150 revolutions per minute (rpm) in tryptic soy broth (TSB) (Difco Laboratories, Detroit, MI). *Escherichia coli* strains were propagated under aerobic conditions at 37 °C and 200 rpm in Luria broth (Sambrook *et al.*, 1989). Antibiotics used were from Sigma (St. Louis, MO) at the following concentrations: ampicillin (Ap) 50 μg ml^{-1}, kanamycin (Km) 70 μg ml^{-1}, rifampicin (Rf) 100 μg ml^{-1} streptomycin (St) 100 μg ml^{-1}, and tetracycline (Tc) 15 μg ml^{-1}. The selection medium utilized for conjugation was a modified DRM medium (Lee *et al.*, 1999) containing the following (per liter): K_2HPO_4, 6g; sodium citrate dihydrate, 6 g; sodium nitrate, 0.5 g; ammonium chloride, 1 g; adjusted to pH 7.9 prior to the addition of agar, 15 g. The medium was autoclaved and cooled to 60 °C prior to the addition of 1 mL of 1 M $MgSO_4$, 666 μl $Ca(NO_3)_2$ and 1 ml of trace elements solution TN2 (Criddle *et al.*, 1990) from sterile solutions. Medium D (Criddle *et al.*, 1990) was used for the determination of CCl_4 degradation capacities.

Table 3.1 List of strains and plasmids used in this study

Strain / Plasmid	Comments	Source / Reference
Pseudomonas stutzeri strain KC	CCl_4 degrader, aquifer isolate	(Criddle *et al.*, 1990)
Pseudomonas stutzeri strains KC657, KC1896, KC2753 and KC3164	Mutants created by transpositional insertion of pRL1063a, Km^R, Rf^{Ra}, impaired in CCl_4 degradation	(Sepúlveda-Torres *et al.*, 1999)
P. stutzeri CCUG[b] 11256	Type strain for *P. stutzeri* species, clinical isolate, genomovar 1	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Stanier *et al.*, 1966)
P. stutzeri ATCC[b] 17591	Clinical isolate, genomovar 2	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Stanier *et al.*, 1966)
P. stutzeri DSM[b] 50227	Clinical isolate, genomovar 3	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Stanier *et al.*, 1966)
P. stutzeri 19SMN4	Marine isolate, naphthalene degrader genomovar 4	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Rosselló *et al.*, 1991)
P. stutzeri DNSP21	Wastewater isolate, genomovar 5	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Rosselló *et al.*, 1991)
P. balearica DSM 6083	Wastewater isolate, naphthalene degrader, genomovar 6	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Bennasar *et al.*, 1996; Rosselló *et al.*, 1991)
P. stutzeri DSM 50238	Soil isolate, genomovar 7	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Stanier *et al.*, 1966)
P. stutzeri JM300	Soil isolate, genomovar	Dr. J. Lalucat, Universitet de les Illes Balears, Spain (Carlson & Ingraham, 1983)
P. putida DSM 3601	Tomato plant isolate, produces 2,6-bis(pyridine thiocarboxylic acid)	Dr. R. Crawford, University of Idaho (Ockels *et al.*, 1978)
P. fluorescens ATCC 13525	Type strain, used to corroborate CCl_4 degradation by non – CCl_4 degraders in the presence of strain KC supernatant	Michigan State University Department of Microbiology
Escherichia coli DH5α	Used for plasmid propagation	Dr. F. deBruijn, Michigan State University (Hanahan, 1983)
E. coli JM109	Used for plasmid propagation	Promega (Madison, WI) (Yanish-Perron *et al.*, 1985)
E. coli S-17-1	St^{Ra} RP4 mobilization genes integrated in the chromosome (mob^+)	Dr. F. deBruijn, Michigan State University (Simon *et al.*, 1983)
pBluescript SK(−)	DNA cloning vector	Stratagene (La Joya, CA)

Table 3.1 (continuation)

Strain / Plasmid	Comments	Source / Reference
pRK311	Tc^{R} [a] wide host range cosmid	Dr. R. Crawford, University of Idaho (Ditta *et al.*, 1985)
pT31	25.7 kb DNA fragment of *P. stutzeri* strain KC cloned into the *Bam*HI site of pRK311	Dr. R. Crawford, University of Idaho (Lewis *et al.*, 2000)
pBlue8.3	Ap^{R}, 8.3 kb DNA fragment of *P. stutzeri* strain KC cloned into the *Eco*RI site of pBluescript SK(−)	This study
pRKblue	Ap^{R} [a], Tc^{R}, pBluescript SK(−) inserted in the *Bam*HI site of pRK311	This study
pRKblue8.3	Ap^{R}, Tc^{R}, pBluescript SK(−) containing a 8.3 kb DNA fragment of *P. stutzeri* strain KC cloned into the *Bam*HI site of pRK311	This study

[a] Km, kanamycin; Rf, rifampicin; St, streptomycin; Tc, tetracycline; Ap, ampicillin

[b] ATCC, American Type Culture Collection, Rockville, MD; CCUG, Culture Collection University of Göteborg, Sweden; DSM, Duetsche Sammulung von Microorganismen und Zellkulturen, Braunschweig, Germany

[R] resistant

Generation and isolation of plasmids containing the genes interrupted by *lux*AB::Tn*5*. Genomic DNA from the four *P. stutzeri* strains KC657, KC1896, KC2753 and KC3164 were isolated with the QIAamp® tissue kit (Qiagen, Valencia, CA), following the manufacturer's instructions. Five micrograms of DNA were restricted with 20 U *Eco*RI (Gibco BRL, Rockville, MD) for 2 hr in a final volume of 50 µl. The restricted DNA samples were extracted with phenol:chloroform, precipitated with ethanol, washed with 70% ethanol and dissolved in 88 ul of water as previously described (Sambrook *et al.*, 1989). The suspended DNA samples were combined with 10 µl 10X T4 ligase buffer and 2 units of T4 ligase (1 U / µl) (Roche Molecular Biochemicals,

Indianapolis, IN) and incubated at 16 °C for 20 hr. Ten microliters of T4-ligated DNA were transformed to $CaCl_2$-competent *E. coli* DH5α using the heat-shock method (Sambrook *et al.*, 1989). The cells were transferred to LB-agar medium amended with Km and incubated overnight at 37 °C. Colonies were streaked once, to assure purity, before the plasmids were isolated using the Quantum® prep plasmid isolation kit (Bio-Rad Laboratories, Hercules, CA). Figure 3.1 shows a graphical representation of how the plasmids containing the mutated DNA were generated.

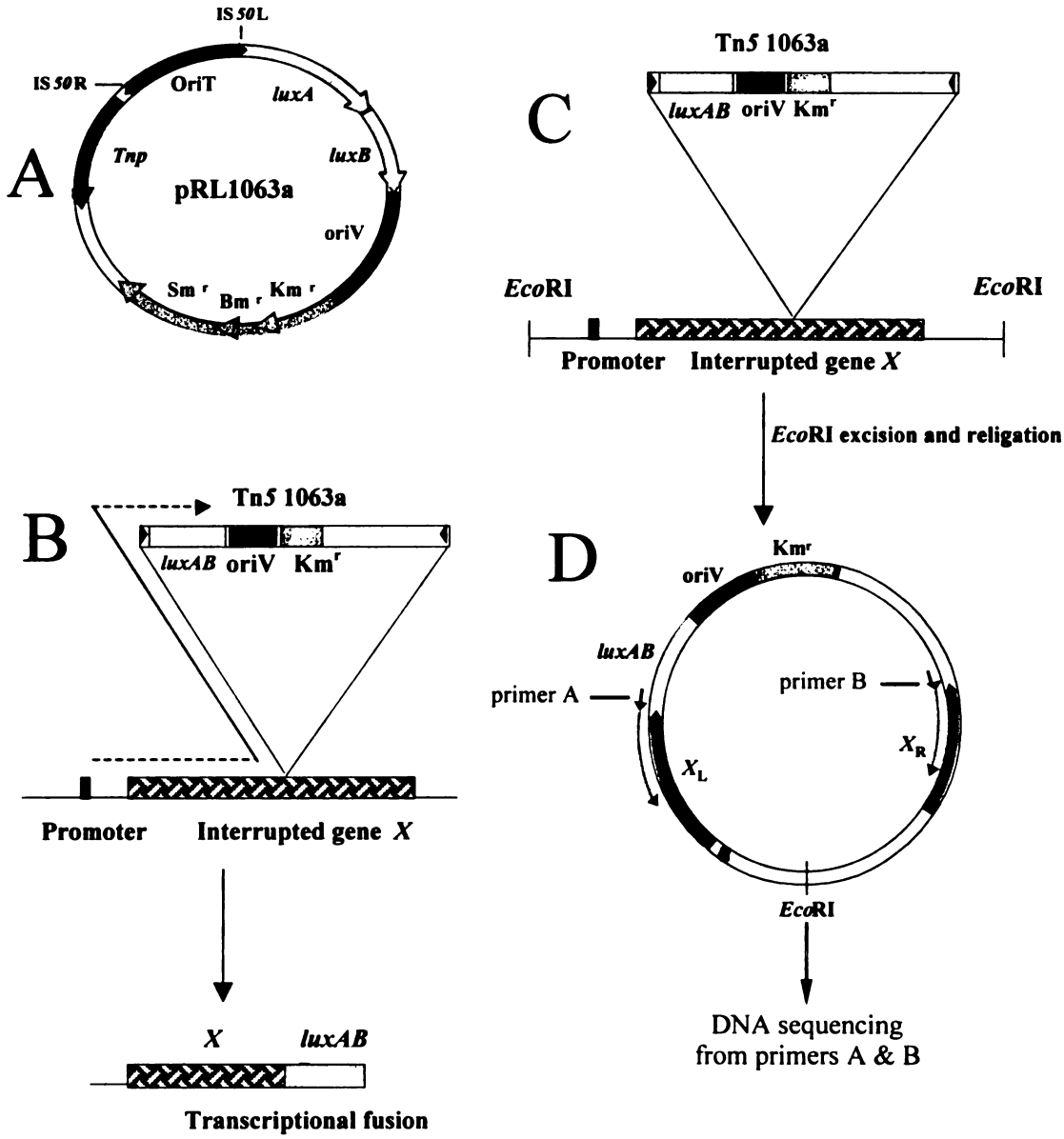


Figure 3.1 Cloning and sequencing of genes mutated *P. stutzeri* strain KC. A) Map of pRL1063a, the plasmid containing the *lux*AB::Tn*5* transposon (Wolk *et al.*, 1991). B) Interruption of gene *X* by *lux*AB::Tn*5*. The *P. stutzeri* strain KC native promoter now controls the luciferase genes. C) Cloning of the interrupted gene by restricting the DNA with *Eco*RI, followed by a ligation reaction. D) Sequencing of the *lux*AB::Tn*5* flanking regions using primers A and B. Figure kindly provided by Frans deBruijn.

Sequencing of the plasmids containing the genes interrupted by *lux*AB::Tn*5*. Double-stranded plasmid DNA was sequenced with a modified dideoxy method of Sanger *et al.* (1977). Automated sequencing readings were performed using the ABI model 377 automated DNA sequencer (Applied Biosystems, Foster City, CA) following the protocols recommended by the manufacturer for *Taq* DNA polymerase cycle sequencing reactions with flourescently labeled BigDyes™ dideoxynucleotide terminators. To determine the sequence of the *P. stutzeri* strain KC DNA on both sides of the transposon insertion, two oligonucleotides derived from the Tn*5*-pRL1063a sequence were used as sequencing primers. One primer, corresponding to positions 110-86 of the Tn*5*-pRL1063a DNA sequence (5'-TACTAGATTCAATGCTATCAATGAG-3'), was designed to determine the upstream sequence from the Tn*5* in the anti-sense direction. The other primer, corresponding to positions 7758-7781 of the Tn*5*-pRL1063a DNA sequence (5'-AGGAGGTCACATGGAATATCAGAT-3') was designed to determine the downstream DNA sequence in the sense direction. These primers were modified from previously described sequencing primers for Tn*5*-pRL1063a (Black *et al.*, 1993; Fernández-Piñas *et al.*, 1994). The sequences of the internal fragments of the DNA inserts were determined, in both orientations, by primer walking. The DNA sequence was deposited to GenBank under accession number AF149851 and is provided in Appendix B.

Cloning of *Eco*RI fragment containing the wild type genes. Genomic DNA from wild type strain KC was isolated with the GenomicPrep cells and tissue DNA isolation kit (Amersham Pharmacia Biotech, Piscataway, NJ). Five micrograms of DNA were

restricted with 10 units (U) of *Eco*RI (Life Technologies, Rockville, MD). The restricted fragments were separated by agarose gel electrophoresis and visualized by staining with ethidium bromide. The fragments ranging in size from 8 to 9 kb were purified from the gel using the QIAEX II gel extraction system (QIAGEN, Valencia, CA) and ligated to pBluescript KS(–) (Stratagene, La Jolla, CA) previously restricted with *Eco*RI and dephosphorilated with calf intestine phosphatase (CIP) (Life Technlogies). The resulting plasmids were introduced into *Escherichia coli* JM109 competent cells (Promega, Madison, WI).

All plasmids were screened for the presence of the fragment of interest by the polymerase chain reaction (PCR) using primers that amplify fragments containing the transposon insertion sites. The PCR were performed in a volume of 50 µl containing 50 ng DNA, 20 mM Tris pH 8.4, 50 mM KCl, 3 mM MgCl, 0.2 mM of each dNTP, 10 pmole of each primer and 2 U *Taq* DNA polymerase (Life Technologies). The following PCR program was used: initial denaturation at 95 °C for 3 min, 30 cycles consisting of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 2 min. A single final extension step was performed at 72 °C for 10 min in order to assure chain termination. Two sets of primers were used: primers CC109f (5' - GTT ACA GCC GCC ACC TAC TGA T – 3') and CC110r (5'- GCT AGG CAG AGA AGA GTC CAC G – 3') were used to amplify a 1.1 kb fragments spanning from position 2,493 to position 3,604; primers CC111f (5' – GGC TGC TCA GTA TCG GCA GTA T – 3') and CC112r (5' – GGG GCG TTG ACA GAG AAG TAA G – 3') were used to amplify a 1.4 kb fragment spanning from position 4,892 to position

6,276. PCR products were separated by electrophoresis in 1.5% agarose and stained with ethidium bromide.

pBluescript SK(−) containing the fragment of interest (pBlue8.3) was restricted with *Bam*HI, treated with CIP and ligated to *Bam*HI-restricted wide host range cosmid pRK311 (Ditta *et al.*, 1985) using T_4 DNA ligase (Life Technologies). The resulting Tc^R Ap^R plasmid (pRKblue8.3) was introduced to *E. coli* S-17-1 by transformation using the $CaCl_2$ – heat shock method (Sambrook *et al.*, 1989), and was transferred to *P. stutzeri* CCUG 11256 by conjugation. Plasmids pRK311, pT31and pRKblue were also introduced to *P. stutzeri* CCUG 11256.

DNA sequence analyzes. Alignments of the DNA fragments obtained from sequencing were done using Sequencher (Gene Codes, Ann Arbor, MI). Open reading frames (ORFs) were determined by the program CodonUse 3.1 (Conrad Halling, University of Chicago) using the codon use tables of *Pseudomonas putida, Pseudomonas fluorescens and Pseudomonas aeruginosa* as references, with a codon window of 33 bases and a logarithmic range of 3. Comparison of DNA and protein sequences with the sequences available in the databases was done using the basic local alignment search tool (BLAST) programs available in the internet search engines of the National Center for Biotechnology Information (Altschul *et al.*, 1997) using the address listed in Table 3.2. Possible promoter regions were localized with the promoter prediction by neural network software administered by the Lawrence Berkeley Laboratory available through the Internet (Reese & Eeckman, 1995), as indicated in Table 3.2. Transcriptional terminators

were predicted by a program available in the Wisconsin Package (Genetic Computer Group, Madison, WI) (Brendel & Trifonov, 1984; Butler, 1998; Devereaux *et al.*, 1984). Molecular weights and isoelectic points of proteins were determined by the Lasergene Package (DNAstar, Madison, WI). Transmembrane helices were predicted by the Internet servers of the seven programs included in Table 3.2. Signal peptides were determined by the SPScan program of the Wisconsin Package and two other Internet resources as indicated in Table 3.2. The localization of motifs was performed by comparing the proteins against the five libraries provided in Table 3.2.

Table 3.2 Internet-based programs used for DNA and protein analyses

Resource type	Resource name (Internet address)	Reference
DNA and protein comparisons	BLAST (www.ncbi.nlm.nih.gov/BLAST)	(Altschul *et al.*, 1997)
Promoter determination	Promoter (www.fruitfly.org/seq_tools/promoter.html)	(Reese & Eeckman, 1995)
Transmembrane helices	TMHMM (www.cbs.dtu.dk/services/TMHMM-1.0/) HMMTOP (www.enzim.hu/hmmtop) SOSUI (azusa.proteome.bio.tuat.ac.jp/sosui/submit.html) TMPred (www.ch.embnet.org/software/TMPRED_form.html) Psort (psort.nibb.ac.jp/form.html) TopPred2 (www.Biokemi.su.se/~server/toppred2/toppredServer.cgi) DAS (www.Biokemi.su.se/~server/DAS)	(Sonnhammer *et al.*, 1998) (Tusnady & Simon, 1998) (Hirokawa *et al.*, 1998) (Hofmann & Stofell, 1993) (Nakai & Kaneshisa, 1991) (von Heijne, 1992) (Cserzo *et al.*, 1994)
Signal peptide	SignalP (www.cbs.dtu.dk/services/SignalP) Psort (psort.nibb.ac.jp/form.html)	(Nielsen *et al.*, 1997) (Nakai & Kaneshisa, 1991)
Protein motifs	Motif finder (motif.genome.ad.jp/) – searches all the following: PROSITES BLOCKS ProDom PRINTS Pfam	 (Hoffmann *et al.*, 1999) (Henikoff *et al.*, 1999) (Corpet *et al.*, 1999) (Attwood *et al.*, 1999) (Bateman *et al.*, 1999)

Corroboration of CCl_4 degradation by P. stutzeri KC657, KC1896, KC2753 and KC3164 when the supernatant from wild type strain KC is provided. In order to asses the CCl_4 degradation capacity of the four mutant strains when the CCl_4 degrading factor is provided externally, mutant cells were combined with partially purified PDTC using the bioassay developed by Dybas *et al.* (1995).

Corroboration of CCl_4 degradation by P. stutzeri CCUG 11256 containing plasmids with DNA from P. stutzeri strain KC. Triplicate 5-ml cultures of P. stutzeri CCUG 11256 harboring pRK311, pT31, pRKblue or pRKblue8.3 were incubated overnight at room temperature and 250 rpm in TSB amended with Ap and Tc as needed. The cultures were used to inoculate 20 ml serum vials containing 10 ml of medium D and 500 ng CCl_4. The vials were incubated in an inverted position for 24 hr before the concentration of CCl_4 remaining in the vials was determined by a gas chromatograph equipped with an electron capture detector (Tatara *et al.*, 1993).

Screening of *Pseudomonas* strains for the presence of the genes interrupted in mutants impaired in CCl_4 degradation. Eight *Pseudomonas stutzeri* strains, representing the eight *P. stutzeri* genomic groups identified to date (Rosselló *et al.*, 1991; Rosselló-Mora *et al.*, 1996), were screened for the presence of the genes interrupted in the mutants that lost the ability to degrade CCl_4. The strains were analyzed by PCR and by Southern hybridization. The PCR primer pairs used for the study were CC109f – CC110r and CC111f – CC112r and the reactions were performed as described in a

previous section of this chapter. The probe for Southern hybridization was a DIG-labeled 3.4 kb *Hind*III fragment corresponding to positions 1653 to 5043 of the 8.3 kb *Eco*RI fragment containing the wild type genes interrupted in the mutants. The probe was hybridized to 10 μg of genomic DNA as recommended by the manufacturer (Roche Molecular Biochemicals). The strains used for this study were *P. stutzeri* CCUG 11256, *P. stutzeri* ATCC 17591, *P. stutzeri* DSM 50227, *P. stutzeri* 19SMN4, *P. stutzeri* DNSP21, *P. stutzeri* DSM 50238, *P. stutzeri* JM300, *P. balearica* DSM 6083 and *P. putida* DSM 3601. The carbon tetrachloride degradation capacity of the strains was tested in medium D containing 500 ng CCl_4 as described previously in this chapter.

RESULTS

Sequencing of genes interrupted by *lux*AB::Tn*5*. Sequence analysis determined that all four mutants had insertions within a 3.1 kb sequence contained within a 8,274 bp *Eco*RI fragment. Furthermore, two of the four mutants received the transposon insertion in the same position, separated only by 9 bp. Figure 3.2 shows a map of the transposon insertion points in the four mutants.

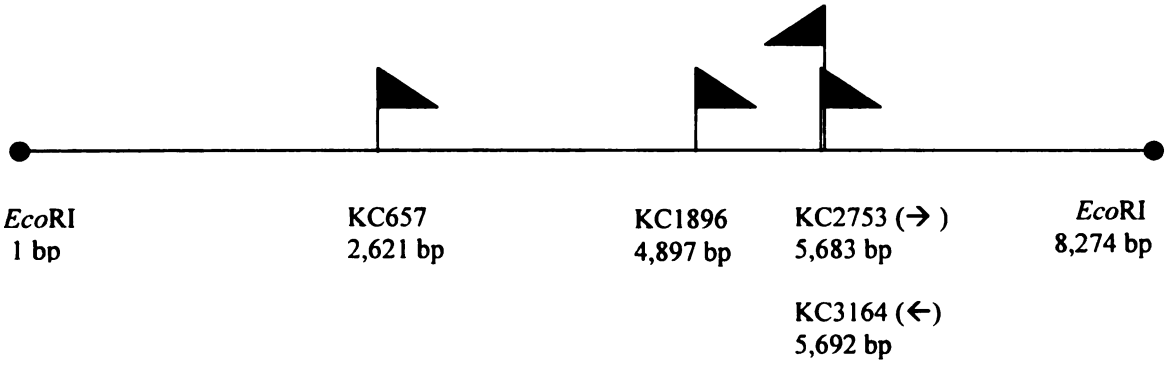


Figure 3.2 Mapping of the transposition insertion points for transpositional mutants KC657, KC1896, KC2753 and KC3164. The black flags represent the *lux*AB::Tn*5* transposon. The direction of the flag indicates the direction of the luciferase genes. The exact positions of the transposon insertion sites are provided.

Determination of open reading frames. The codon use tables of *P. putida, P. fluorescens* and *P. aeruginosa* generated the same open reading frame profile. As seen in Figure 3.3, eight ORF's were found in the forward direction. Mutant KC657 mapped in ORF 2435–3610 while mutants KC1896, KC2753 and KC3164 mapped in ORF 4460-6291.

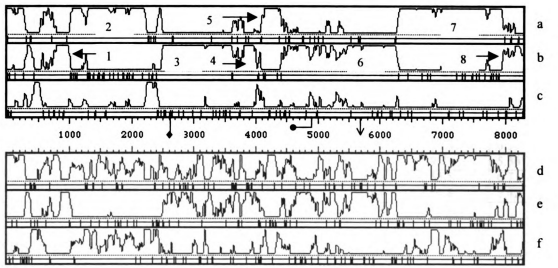


Figure 3.3 Open reading frames (ORFs) found in the 8,274 bp *Eco*RI fragment mutated in KC657, KC1896, KC2753 and KC3164. Panels a, b and c represent the 3 forward reading frames while panels d, e and f represent the 3 reverse reading frames. A small vertical line depicts the methione start codon and a small bold vertical line represents the stop codons. 1, ORF-446; 2, ORF-1009; 3, ORF-2435; 4, ORF-3626; 5, ORF-4099; 6, ORF-4460; 7, ORF-689; 8, ORF-7985. The insertion points in KC657 (♦), KC1896 (•) and KC2753/KC3164 (↓) are also provided.

Determination of physical characteristics of the open reading frames. The start and stop positions of the ORFs, the putative ribosomal binding site (RBS) (Shine & Delgarno, 1974), as well as the number of amino acids in the encoded proteins, their molecular weights and isoelectric points are indicated in Table 3.3.

Table 3.3 Physical characteristics of the open reading frames encoded in the *P. stutzeri* strain KC 8.3 kb *Eco*RI fragment interrupted in mutants impaired in CCl_4 degradation.

Frame	Start Base	Ending Base	RBS[a] (position)	aa[b]	kDa[c]	pI[d]
2	446	1012	AGGA (437 – 439)	188	20.1	8.65
1	1009	2262	AGAGGA (994 – 999)	417	45.4	10.52
2	2435	3610	AGGA (2423 – 2426)	391	42.8	4.93
2	3626	4036	AGGA (3612 – 3615)	136	15.6	6.50
1	4099	4371	AGGA (4088 – 4091)	90	9.7	5.56
1	4460	6291	GG (4450 – 4451)	610	65.5	6.71
1	6289	7983	GGAGG (6279 – 6283)	564	60.8	5.42
2	7985	> 8274	GGAG (7970 – 7974)	> 96	> 10.1	

[a] RBS, ribosomal binding site
[b] aa, number of amino acids in the protein encoded by the gene
[c] kDa, protein molecular weight in kiloDaltons
[d] pI, isoelectric point of protein

Sequence analysis of the 8.3 kb DNA fragment. The information gathered by comparing the ORFs with DNA and protein databases, motif databases, transmembrane helices prediction and leader peptide determination programs was used to assign possible gene functions. This information is summarized in Table 3.4 and provided in detail in Appendix C.

Table
strain

F
446 - 1
1009 – - Putati mem with funct
2455 – - Interru KC65 - Memb protei involv transf
3626 – 4 - Functio
4099 – 4 - May be sulfur t
4460 – 6 - Interrup KC189 and KC - Possible protein involve addition of H_2O

Table 3.4 Putative functions of open reading frames (ORFs) found in the *P. stutzeri* strain KC 8.3 kb fragment interrupted in mutants impaired in CCl_4 degradation.

ORFs Functions	Similarities	Motifs
446 - 1012	- No similarities found	- No motifs found
1009 – 2262 - Putative integral membrane protein with unknown function	- No similarities found	- Imperfect matches found with ATP-phosphoribosyl transferase, Fe-containing alcohol dehydrogenase, homoserine dehydrogenase - 12 transmembrane helices predicted by 7 programs
2435 – 3610 - Interrupted in KC657 - Membrane bound protein – may be involved in sulfur transfer	- 10^{-122} [a] MoeZ[b] from *Mycobacterium tuberculosis* (Z95120[c], 71% p[d] and 57% i[e], 389 aa[f]) - 10^{-104} MoeB from *M. tuberculosis* (molybdopterin synthase sulfurylase, Z95150, 67% p, 51% i, 379 aa) - 10^{-42} MoeB from *Escherichia coli* (D90720, 56% p, 39% i, 245 aa) - 10^{-35} ThiF from *E. coli* (activates protein involved in thiamin biosynthesis, P30138, 56% p, 37% i, 235 aa)	- Imperfect matches found with chemotaxis sensory transducer, dihydroxyacid dehydratase, amino acid dehydrogenase, pyridine nucleotide disulfide oxidoreductase, ATP-phosphoribosyl transferase, purine phosphorilase, purivate kinase, - Transmembrane domains: 0 (SOSUI), 1 (TMHMM), 2 (HMMTOP, TMPred, PSort, DAS, TopPred2) - 11 bp signal peptide found by SignalP
3626 – 4036 - Function unknown	- 10^{-30} hypothetical 16.5 kDa protein Rv1334 from *M. tuberculosis* (Q10645, 67% p, 47% i, 134 aa) - 10^{-14} Mec^{+} from *Streptomyces kasugaensis* (restores cysteine methionine nutritional deficiency, M29166, 62% p, 42% i, 96 aa)	- Imperfect match found with bacterial ribonuclease P, an enzyme that cleaves extra nucleotides from tRNA precursors
4099 – 4371 - May be involved in sulfur transfer	-10^{-20} hypothetical protein Rv1335 from *M. tuberculosis* (Z73902, 74% p, 51% i, 90 aa) - 10^{-4} MoaD from *Archeaglobus fulgidus* (adds sulfur to molybdopterin precursor, AE000990, 50% p, 32% i, 74 aa)	-Perfect match for TonB-dependent protein N-terminus recognition sequence (79-DSLTVXPA-86).
4460 – 6291 - Interrupted in KC1896, KC2753 and KC3164 - Possible membrane protein. May be involved in addition or removal of H_2O or H_2.	- 10^{-23} hypothetical protein Rv3272 from *M. tuberculosis* (Z92771, 49% p, 33% i, 278 aa) - 10^{-22} BaiF from *E. coli* (bile acid dehydroxylase, D90867, 52% p, 30% i, 242 aa) - 10^{-6} CaiB from *E. coli* (L-carnitine dehydratase, P31572, 39% p, 30% i, 197 aa)	- Imperfect matches with alanine dehydrogenase and glutamyltranspeptidase. - Transmembrane domains: 0 (TMHMM, HMMTOP, SOSUI), 1 (PSort), 5 (TopPred2), 6 (TMPred), 7 (DAS) - 124 amino acid signal peptide predicted by SignalP

Table 3.4 (continuation)

ORFs Functions	Similarities	Motifs
6289 – 7983 - Possible AMP-ligase involved in substrate activation	- 10^{-55} DhbE from *Bacillus subtilis* (2,3-dihydroxybenzoate-AMP ligase, P40871, 48% p, 30% i, 530 aa) - 10^{-52} PchD from *Pseudomonas aeruginosa* (AMP-ligase in pyochelin biosynthesis, X82644, 44% p, 30% i, 529 aa) - 10^{-47} SnbA from *Streptomyces pristinaespiralis* (3-hydroxydipicolinic acid-AMP ligase, X98515, 45% p, 30% i, 545 aa) - 10^{-39} EntE from *E. coli* (2,3-dihydroxybenzoate-AMP ligase, P10378, 45% p, 26% i, 526 aa)	- Perfect AMP binding domain (194-LLVSSGTESEPK-205). - Transmembrane domains: 0 (TMHMM, 1 (SOSUI), 2 (PSort), 3, HMMTOP, 4 (TMPred, TopPred2), 5 (DAS).

[a] Match probability
[b] protein name
[c] GenBank accession numbers
[d] p, positive matches (amino acids with the same functional side chains)
[e] i, identify matches (identical amino acids)
[f] aa, amino acids (number of amino acids overlapping in match alignment)

Determination of transcriptional terminators and promoters. A stable stem – loop structure was predicted from positions 2,295 to 2,331, between the termination codon of ORF-1009 and the initiation codon of ORF-2435. As illustrated in Figure 3.4, this sequence forms a stable G-C rich hairpin followed by a string of 5 uridine residues in the resultant mRNA. This structure resembles the Rho-independent terminators of *E. coli* (Adhya & Gottesman, 1978; Platt, 1986). Similar structures have also been found in *P. aeruginosa* (Gray *et al.*, 1984) and *Burkholderia cepacea* (Zylstra *et al.*, 1989). Two transcriptional promoters were predicted between the termination codon of ORF-1009 and the initiation codon of ORF-2435, one corresponding to positions 2343 – 2388 (87%

certa

predi

2.

Figure
Primar

Corrob

KC316

Table 3

strain K

observa

indepen

certainty) and the other corresponding to positions 2362 – 2407 (94% certainty). The predicted transcription starts are cyidine-2,368 and guanosine-2,397, respectively.

2,295 –AATGGCCTGAAATGCGCGGTCCTTTTGCATAGTTTTT – 2,331

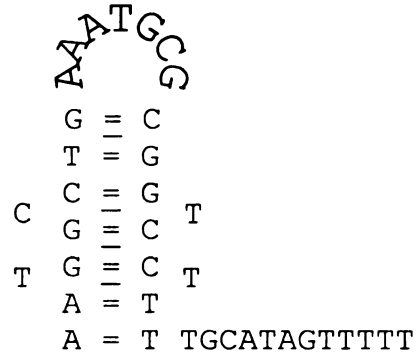
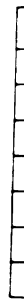


Figure 3.4 Stem – loop secondary structure observed between positions 2,295 and 2,331. Primary sequence value = 3.99, secondary sequence value = 55.

Corroboration of CCl_4 degradation by P. stutzeri KC657, KC1896, KC2753 and KC3164 when the supernatant from wild type strain KC is provided. As seen in Table 3.5, the mutants were able to degrade CCl_4 when the supernatant of a wild type strain KC culture capable of degrading CCl_4 was provided. This result confirms previous observations indicating that the ability to utilize the CCl_4 degradation agent is totally independent of its production (Dybas *et al.*, 1995; Tatara *et al.*, 1995)

Tabl
the su



Corro

with I

degrad

Screen

mutan

primer

strain c

produce

Table 3.5 Degradation of CCl_4 by strains KC657, KC1896, KC2753 and KC3164 when the supernatant from a wild type strain KC culture capable of degrading CCl_4 is provided.

Strain	% CCl_4 degraded (T = 2.5 h)
Non-inoculated culture medium	0 ± 1
< 10 kDa m.w. supernatant alone	0 ± 14
P. fluorescens + supernatant	61 ± 3
KC657 + supernatant	67 ± 1
KC1896 + supernatant	62 ± 7
KC2753 + supernatant	61 ± 3
KC3164 + supernatant	63 ± 2

Corroboration of CCl_4 degradation by P. stutzeri CCUG 11256 containing plasmids with DNA from P. stutzeri strain KC. Only the strain harboring pT31 was able to degrade CCl_4.

Screening of *Pseudomonas* strains for the presence of the genes interrupted in mutants impaired in CCl_4 degradation. The strains tested did not amplify with the two primer sets used and no signal was observed in the Southern hybridization. The only strain capable of degrading carbon tetrachloride was *P. putida* DSM 3601, a strain that produces PDTC when grown in iron-limiting conditions.

The

Sepúl

ORF-

transp

with i

lucifer

inserte

Lewis

degrad

When

Sepúlv

that the

spannin

The re

the 25

detecta

CTN1

ORFs

produc

DISCUSSION

The four transpositional mutants of *Pseudomonas stutzeri* strain KC created by Sepúlveda *et al.* (1999) that showed an impaired ability to degrade CCl_4 had mutations in ORF-2435 (KC657) or ORF-4460 (KC1896, KC2753 and KC3164). Furthermore, the transposon insertion in mutants KC2753 and KC3164 mapped to the same location but with inverted orientations. These results explain why KC3164 was unable to translate the luciferase reporter genes at any significant levels because the luciferase genes were inserted in the wrong orientation with respect to the promoter.

Lewis *et al.* (2000) sequenced a 25.7 kb fragment that was able to restore the CCl_4 degradation ability to a spontaneous strain KC mutant that lost a 170 kb DNA fragment. When the 8.3 kb *Eco*RI fragment containing the four mutation points characterized by Sepúlveda was compared with the 25.7 kb sequence reported by Lewis, it was evident that the 8.3 kb fragment corresponds to positions 4,041 to 12,314 of the Lewis sequence, spanning from the end of Lewis' ORF-C to the beginning of ORF-K.

The report by Lewis *et al.* (2000) included the results of a semi-saturation mutagenesis of the 25.7 kb fragment with transposon mini-Tn*5*::*lacZ*1 or mini-Tn*5*::*pho*A. No detectable decrease in PDTC production and CCl_4 degradation was observed when strain CTN1 harbored a T31 fragment mutated in ORFs A, B, C, D, E, M, or the space between ORFs N and O. On the other hand, when ORFs F, J or P were mutated, PDTC production was impaired and no significant degradation of CCl_4 was observed.

Mut

tran

only

conta

the r

Mutations in ORF-K showed a variable phenotype, depending on the position of the transposon insertion. When smaller DNA fragments were introduced to strain CTN1, only partial restoration of the CCl_4 degradation capacity was seen when a fragment containing ORFs A to the beginning of O was used. Figure 3.5 illustrates and compares the results independently obtained by Sepúlveda and by Lewis.

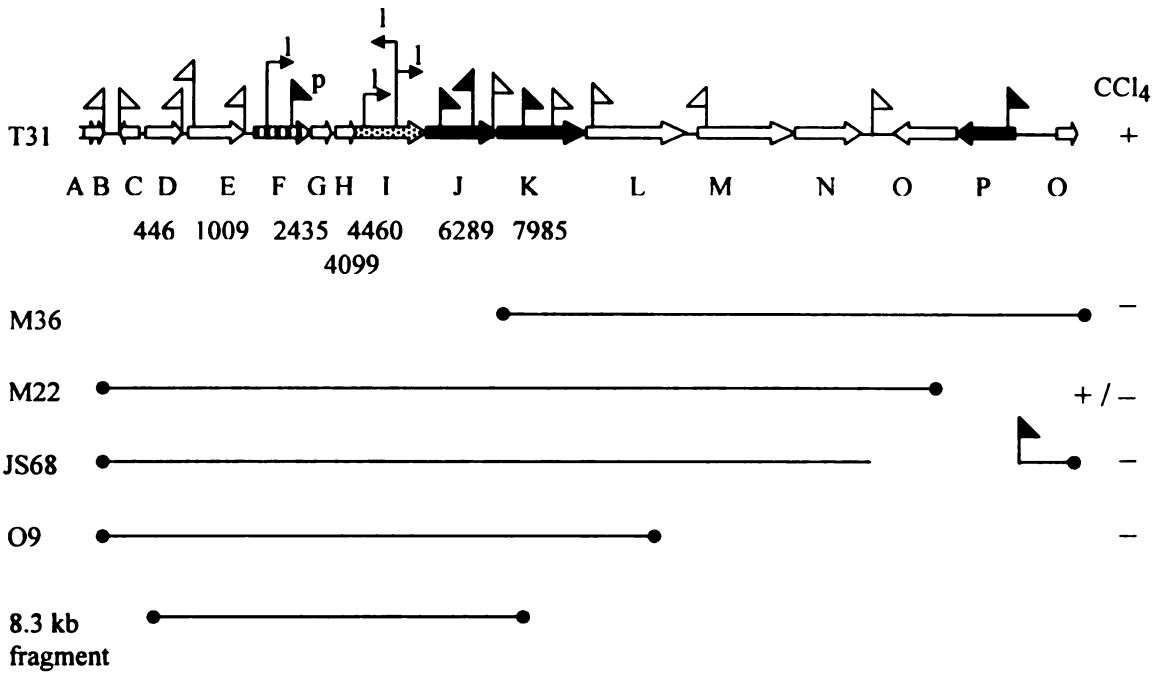


Figure 3.5 Organization of 16 open reading frames (ORFs) in a 25.7 kb fragment capable of restoring the CCl_4 degradation capacity in strain CTN1. Block arrows with letter designations below indicate ORFs identified by Lewis. The numbers underneath the ORF letters indicate the equivalent nomenclature used by Sepúlveda. The directions of the arrows indicate the transcription/translation orientations of the genes. White arrows designate ORFs not mutated by Lewis *et al.* (2000) or ORFs that did not affect CCl_4 degradation when mutated. Black arrows designate ORFs that caused a CCl_4-negative phenotype. The stripped arrow designates an ORF mutated by Sepúlveda and by Lewis while the dotted arrow designates an ORF mutated only by Sepúlveda. Tn*5*::*lac*1 insertions are represented by vertical lines with flags. The orientations of the flags indicate the orientation of the *lacZ* gene and the color indicates the mutation's effect on the CCl_4 phenotype as described above. The flag designated by a letter (p) indicates a Tn*5*::*pho*A insertion. Tn*5*::*lux*AB insertions made by Sepúlveda are indicated by lines with linear arrows identified by a letter (l). The orientation of the linear arrows represents the direction of the luciferase genes. Partial fragments introduced to strain CTN1 by Lewis and their ability to complement the CCl_4 degradation mutation are also provided. This figure was modified from (Lewis *et al.*, 2000).

The

tran

som

that

PDT

termi

predi

codon

ORF-

uptake

promo

AAT C

when i

scaveng

factor o

The results reported by Lewis *et al.* (2000) confirmed that the four mutants obtained by transpositional insertion using Tn*5*::lux are impaired in the degradation of CCl_4 because some of the structural genes needed for the production of PDTC were truncated. The fact that all mutations upstream of Lewis' ORF-F (Sepúlveda's ORF-2435) had no effect on PDTC production confirmed the findings obtained by Sepúlveda in the promoter and terminator analyses of the 8.3 kb *Eco*RI fragment. As seen in Figure 3.6, there is a predicted Rho-independent terminator followed by a promoter between the termination codon of ORF-1009 (Lewis' ORF-E) and the initiation codon of ORF-2435 (Lewis' ORF-F). Further analysis of that region also showed the presence of an imperfect ferric uptake regulator protein binding site (Fur box), an A-T rich sequence found in the promoter region of many iron-controlled genes in *E. coli* (*E. coli* consensus = 5'–GAT AAT GAT AAT CAT TAT C–3') (Braun & Hantke, 1991). Fur is believed to bind DNA when iron is abundant in order to repress the synthesis of proteins involved in iron scavenging. Sequences similar to the -10 and -35 sequences recognized by transcription factor σ^{70} were also identified.

225

229

233

237

241

Figu
Sequ
initia
bold

ORF

et al

10^{-1}

from

Sign

Moe

sulfu

(Moa

synth

adeny

adeny

from

2251 GAGCGGGTTT GAAGGCTGAA GTGACCGGCC ATGCCCCTTC
ORF-1009 stop codon

2291 GGACAATGGC CTGAAATGCG CGGTCCTTTT GCATAGTTTT
predicted Rho-independent terminator

2331 TCATGCTCAC GTCATATGAA AGAACAGCCA ACGGCAATTG
-35 sequence

2371 CTATAGTCAT CACCACGAAC GATAATGATT ATCGTTACCA
-10 sequence Fur Box

2411 TTGAAATCAA ACAGGATAAG CGATATGCCA CTATCAGCGC
RBS ORF-2435 start codon

Figure 3.6 Organization of the DNA region between ORF-1009 and ORF-2435. Sequences associated to termination of transcription and translation of ORF-1009 and the initiation of transcription and translation of ORF-2435 are indicated. Cysteine residues in bold correspond to possible transcription initiation sites for ORF-2435.

ORF-2435 (Lewis' ORF-F) was mutated in strain KC657 and was also mutated by Lewis *et al.* (2000) using a Tn*5*::*pho*A transposon. It is very similar to MoeZ (probability of 10^{-122}), a putative *Mycobacterium tuberculosis* protein. It is also very similar to MoeB from *M. tuberculosis* (probability of 10^{-104}) and *Escherichia coli* (probability of 10^{-42}). Significant similarities were also observed with ThiF from *E. coli* (probability of 10^{-35}). MoeZ is a hypothetical protein, named after MoeB, the molybdopterin synthase sulfurylase protein from *E. coli*. MoeB transfers sulfur to the molybdopterin synthase (MoaD/MoaE heterodimer) in order to sulfurylate precursor Z of the molybdopterin synthesis pathway (Rajagopalan, 1996). ThiF is also similar to MoeB and it catalyses the adenylation by ATP of the carboxyl-terminal (C-terminus) glycince of ThiS. The adenylation of ThiS is likely to be involved in the activation of ThiS for sulfur transfer from cysteine or from a cysteine-derived sulfur donor in the thiamine biosynthetic

path

cyst

tuber

cyste

transf

encod

transf

The p

progra

domai

progra

Signal

termin

Lewis

2435

phosp

into o

the pr

No p

simil

prote

pathway (Taylor *et al.*, 1998). MoeB, ThiF and other MoeB-like proteins contain two cysteines separated by two amino acids (C-X-X-C motifs) that are missing in both the *M. tuberculosis* MoeZ protein, and the protein encoded by ORF-2435. These paired cysteines have been proposed to bind zinc and form an active center involved in sulfur transfer (Rajagopalan, 1997). The lack of the C-X-X-C motifs in MoeZ and the protein encoded in ORF-2435 may indicate that these proteins are not directly involved in sulfur transfer.

The protein encoded by ORF-2435 is a putative cell membrane bound-protein since five programs (HMMTOP, TMPred, PSort, DAS and TopPred2) predicted 2 transmembrane domains in the vicinity of amino acids 40 to 70 and 200 to 230, and an additional program (TMHMM) predicted only the first transmembrane domain. The program SignalP predicted a possible ten-residue signal sequence at the amino terminus (N-terminus) that may be removed upon protein translocation through the cell membrane. Lewis *et al.* (2000) confirmed the membrane topology of the protein encoded by ORF-2435 by the expression of *pho*A on an iron-limited medium. PhoA is an alkaline phosphatase from *E. coli* used to identify genes whose protein products are transported into or through the inner membrane because it is only active when a significant portion of the protein is located in the periplasmic space (Gutiérrez *et al.*, 1987).

No perfect motifs were found in the protein encoded in ORF-2435 but it has signatures similar to motifs found in different enzymes. Such signatures may provide insights on protein function. Signatures similar to amino acid dehydrogenases, amino acid

amir

regic

conv

disso

inter-

1990

bond

et al.,

amino

protei

concer

adapta

other r

phosph

The n

chemo

or ami

Neithe

ORFs

and R

tander

genon

aminotransferases and purine phosphorilases were found in the first membrane-spanning region predicted by all seven programs. Amino acid dehydrogenases are involved in the conversion of an amine group into a carbonyl group, with the release of ammonia and the dissociation of water (Link *et al.*, 1997). Amino acid aminotransferases catalyze the inter-conversion of amino acids by the exchange of the amino group (Scofield *et al.*, 1990). Purine phosphorilases catalyze the phosphorolytic breakdown of the N-glycosidic bond with the formation of the corresponding free base and pentose-1-phosphate (Seeger *et al.*, 1995). Two signatures were found near the C-terminus end of the protein, between amino acids 330 and 370. One of the signatures corresponds to integral membrane proteins that serve as chemotaxis transducers. Such proteins respond to changes in the concentrations of attractant and repellents in the environment and facilitate sensory adaptation through the variation of the level of methylation (Boyd *et al.*, 1983). The other motif corresponds to ATP phosphoribosyltransferase involved in the activation of a phosphorilated ribose involved in the biosynthesis of histidine (Javanovic *et al.*, 1994). The motifs suggest that the protein encoded by ORF-2435 may be involved in chemotaxis signal transduction and/or substrate modification by the transfer of phosphate or amino groups.

Neither Sepúlveda nor Lewis isolated mutants in ORF-3626 and ORF-4099 (Lewis' ORFs G and H). The highest similarities were seen with hypothetical proteins Rv1334 and Rv1335, respectively. These proteins are found in the *M. tuberculosis* genome in tandem, which indicates that similar protein pairs may be found elsewhere in bacterial genomes, suggesting the possibility of similar protein functions and the acquisition of

such genes by lateral transfer. ORF-3626 shows similarity to an unidentified protein from *Streptomyces kasugaensis* that complements a methionine-cysteine double auxotroph through an uncharacterized process (Hirasawa *et al.*, 1985). ORF-3626 also has a signature corresponding to an enzyme involved in tRNA maturation by endonucleolitic cleavage (Hansen *et al.*, 1985).

ORF-4099 shows low similarity (10^{-4}) to MoaD-like proteins from *Archeaglobus fulgidus* and even lower similarities to MoaD-like proteins from *Helicobacter pilori* and *M. tuberculosis*. No similarities were observed with the *E. coli* MoaD. These MoaD-like proteins were named for their similarity to the *E. coli* MoaD and there is no empirical evidence of their function. MoaD is part of the molybdopterin synthase heterodimer involved in the synthesis of molydopterin from precursor Z and is the substrate for MoeB. MoeB forms a thiocarboxylate at the MoaD C-terminus glycine-glycine (G-G) sequence that is believed to serve as the sulfur donor for molybdopterin (Rajagopalan, 1997). The G-G sequence is also shared with ubiquitin, which in eukaryotes is attached to a series of transfer proteins via thioester linkage at its terminal glycine. ORF-4099 ends with two glycines in tandem.

Motif analyses of ORF-4099 revealed that it encodes for a protein with a perfect signature for TonB-dependent receptor signature ($X_{10\ to\ 115}$ – DSLTVXPA, X_{78} in the case of ORF-4099), even though it appears to be a cytoplasmic protein. TonB is an inner membrane-anchored protein that extends through the periplasmic space in order to couple cytoplasmic membrane energy to active transport of iron-bearing siderophores across the

out

beli

of tl

has

Ton

rath

coup

phys

expl

durin

signa

ORF-

Sepúl

highe

(prob

of 10

introd

carnit

were

TopP

acid s

with

outer membrane. The interactions between TonB and the outer membrane receptors are believed to cause a transformational change in the loaded receptor, leading to the release of the siderophore into the periplasmic space (Braun, 1995; Postle, 1990). The TonB box has been emphasized as an important mediator of some physical interactions between TonB and TonB-dependent receptors. It is likely that the conformation of the TonB box rather than the specific amino acid sequence dictates productive interactions of energy coupling with TonB (Kadner, 1990; Larsen *et al.*, 1997). Since MoeB and MoaD interact physically during the biosynthesis of the molybdopterin cofactor, it would be worth exploring the possibility of physical interactions between ORF-2435 and ORF-4099 during the production of PDTC. Such interactions may be mediated by the TonB signature and/or the adjacent G-G carboxyl-terminus motif present in ORF-4099.

ORF-4460 (Lewis' ORF I) was mutated in strains KC1896, KC2753 and KC3164 by Sepúlveda but Lewis *et al.* (2000) were unable to identify any mutants in this region. The highest similarity was found with a hypothetical protein Rv3272 from *M. tuberculosis* (probability of 10^{-23}). ORF-4460 is also similar to bile acid dehydroxylase (probability of 10^{-22}) and L-carnitine dehydratase (probability of 10^{-6}), two enzymes from *E. coli* that introduce double bonds in their substrates. Higher similarities were observed with L-carnitine dehydratase - like proteins from different organisms. No transmembrane helices were predicted by TMHMM, HMMTOP or SOSUI while TMPred, Psort, DAS and TopPred2 predicted 6, 1,7 and 5 transmembrane domains, respectively. A 124 amino acid signal peptide was predicted by SignalP. Imperfect motif matches were observed with alanine dehydrogenase, phosphoribosylamine-glycine ligase and

glutamyltranspeptidase. A conclusive function cannot be derived from the motif information. On the other hand, a function related to the addition or removal of water or hydrogen can be suggested based on the functions of the two *E. coli* proteins similar to this ORF. The topology of this protein cannot be completely elucidated from the prediction programs due to the differences in the number of transmembrane domains predicted.

ORF-6289 (Lewis' ORF-J) is similar to several AMP-ligases involved in the activation of siderophore precursors by adenylation. The activated precursors are then combined with other intermediates to produce the mature siderophore. The highest similarity was observed with 2,3-dehydroxybenzoate-AMP ligase (probability of 10^{-55}) from *Bacillus subtilis* (Rowland *et al.*, 1996) and *E. coli* (probability of 10^{-39}) (Rusnak *et al.*, 1989), involved in the biosynthesis pathway of the siderophore enterobactin. A significant similarity was also observed with an AMP-ligase involved in the biosynthesis of the siderophore pyochelin in *P. aeruginosa* (probability of 10^{-52}) (Serino *et al.*, 1997). A perfect AMP binding domain was observed in residues 194 to 205. No transmembrane helices were predicted by TMHMM, but several membrane spanning segments were predicted by SOSUI (1 domain), PSort (2 domains), HMMTOP (3 domains), TMPred (4 domains), TopPred2 (4 domains) and DAS (5 domains). The BLAST similarities and the motif searches indicate that the protein encoded by ORF-6289 may be an integral membrane protein involved in reagent activation by adenylation.

ORI

K (6

recep

43% o

1994

simil

FyuA

proce

Moec

that it

induci

sidero

the in

memb

C-term

indirec

that in

interac

In add

(2000)

TonB-

A 25

ORF-7985, truncated in Sepúlveda's 8.3 kb *Eco*RI fragment, corresponds to Lewis' ORF-K (687 amino acids, molecular weight 75.3 kD). It is similar to the ferric yersiniabactin receptor (FyuA) from *E. coli* (GenBank accession number Z38065, 26% identities and 43% positives over a 645 amino acid stretch with a probability of 10^{-34}) (Rakin *et al.*, 1994) and *Yersenia pestis* (GenBank accession number Z35104, 25% identities and 42% similarities over a 645 amino acid stretch with a probability of 10^{-33}) (Rakin *et al.*, 1995). FyuA is an outer membrane protein involved in the uptake of ferric yersiniabactin in a process that requires TonB and energy from the proton motive force (Klebba *et al.*, 1993; Moeck & Coulton, 1998; Rakin *et al.*, 1994). One popular concept of TonB function is that it spans the periplasm and physically interacts with the siderophore-loaded receptor, inducing conformational changes in the receptor that lead to the release of the siderophore to periplasmic binding proteins. TonB has also been hypothesized to induce the intake of loaded siderophores by transiently juxtaposing the inner and outer membranes when the N-terminus of TonB is anchored to the inner membrane while the C-terminus spans the outer membrane bilayer. A third hypothesis suggests that TonB indirectly influences siderophore uptake by energizing a mobile periplasmic messenger that interacts with outer membrane receptors (Klebba *et al.*, 1993). The possible interactions between ORF-K and TonB may be the topic for future investigations.

In addition to the predicted TonB box at amino acids 84 to 92 reported by Lewis *et al.* (2000), an additional sequence similar to signatures found towards the C-terminus of TonB-dependent siderophore receptors was found in amino acids 543 to 585 of ORF-K. A 25 to 50 amino acid signal peptide was predicted for this protein, depending on the

program used. Lewis observed a variable effect when this ORF was mutated. A mutation near the beginning of the ORF impaired PDTC production and CCl_4 degradation while a mutation on the last third of the gene was neutral. This result indicates that the first 2 thirds of the gene are enough to produce a functional protein. The BLAST similarities and the motif searches suggest that the protein encoded by ORF-K is a putative TonB-dependent outer membrane siderophore receptor.

Lewis' ORF-L encodes for a 743 amino acid long protein (molecular weight 83.3 kD) of unknown function, whose mutation does not affect the CCl_4 degradation ability. This protein does not have any perfect motifs but has several sequences similar to motifs with known functions including a chaperonin cpn60 - like sequence and a sequence similar to bacterial type II secretion system protein E . Chaperonins are cytoplasmic ATPases that prevent misfolding and promote refolding and assembly of unfolded polypeptides generated under conditions of stress in a variety of organisms from bacteria to mammals (Hemmingsen *et al.*, 1988). The proteins belonging to bacterial type II secretion system protein E are cytoplasmic ATPases that facilitate protein secretion, DNA uptake and assembly of type-IV fimbriae in *E. coli* and other bacteria (Whitchurch & Mattick, 1994). The protein encoded by ORF-L lacks an ATP binding site. If this protein is indeed a facilitator of protein folding or translocation, the energy needed for the process should be provided from an external source.

ORF-M encodes for a 764 amino acid long protein (molecular weight 83 kD) that does not affect PDTC production and CCl_4 degradation upon mutation. It is similar to putative

aminotransferases from *Streptomyces coelicolor* (GenBank accession number CAB39702, 50% positives and 35% identities over a 466 amino acid overlap with a probability of 10^{-70}) (Redenbach *et al.*, 1996) and *E. coli* (GenBank accession number P42588, 54% positives and 39% identities over a 407 amino acid overlap with a probability of 10^{-65}) (Blattner *et al.*, 1997). Lower similarities were observed with proven aminotransferases. ORF-M has a perfect aminotransferases class III pyridoxal phosphate attachment site in amino acids 524 to 561. Pyridoxal phosphate is an essential cofactor for reactions which act on the C-2 (alpha carbon) atom of amino acids and involve cleavage of any of the amino acid bonds. It is the coenzyme in a large number of amino-acid conversions such as transaminase, decarboxylase and dehydratase reactions. Pyridoxal phosphate is covalently attached to the enzyme by the amino group of a lysine side chain (Michal, 1999). Both BLAST and motif searches suggest that the protein encoded by ORF-M is a putative aminotransferase.

ORF-N encodes for a 392 amino acid long protein (molecular weight 40.6 kD) that was not mutated by Lewis. It is similar to a putative transmembrane efflux protein from *Streptomyces coelicolor* (GenBank accession number CAB66188, 40% positives and 26% identities over a 466 amino acid overlap with a probability of 10^{-9}) (Redenbach *et al.*, 1996) and *Bacillus subtilis* (GenBank accession number CAB16085, 43% positives and 25% identities over a 308 amino acid overlap with a probability of 10^{-9}) (Ogasawara *et al.*, 1994). This ORF has 8 to 12 predicted transmembrane helices, depending on the program used, and there are three predicted sites for the cleavage of a leader sequence, which range from 30 to 120 amino acids. Even though this ORF does not have perfect

m

So

19

(Ya

for

sug

ORF

54.7

dehy

60%

10^{-70}

positi

(Tosa

obser

first s

hydro

yield

tricar

of ce

the ß-

motifs, sequences similar to motifs found in integral membrane transporters were found. Some examples include hemolysin and alkaline protease secretion proteins (Hess *et al.*, 1986), sugar transport proteins (Blattner *et al.*, 1997), sodium - galactoside symporters (Yazyu *et al.*, 1984), glycerol-3-phosphate transporter (Eiglmeier *et al.*, 1987), and formate – nitrite transporters (Sawers & A, 1989). The aforementioned information suggests that the protein encoded in ORF-N is a permease involved in solute transport.

ORF-O encodes for a 513 amino acid long protein with a predicted molecular weight of 54.7 kD. This protein was not mutated by Lewis and it is similar to an acyl-CoA dehydrogenase from *Deinococcus radiolarians* (GenBank accession number AAF10499, 60% positives and 39% identities over a 376 amino acid overlap with a probability of 10^{-70}) (White *et al.*, 1999) and *B. subtilis* (Genbank accession number CAA74221, 56% positives and 39% identities over a 375 amino acid overlap with a probability of 10^{-60}) (Tosato *et al.*, 1997). Two perfect motifs found in acyl-CoA dehydrogenases were observed in residues 256 to 268 and 464 to 483. Acyl-CoA dehydrogenase catalyses the first step in the degradation of fatty acids by the process of ß-oxidation. It removes two hydrogens, leading to the formation of a double bond that will eventually be cleaved to yield acetyl-CoA and an acyl-CoA shortened by two carbons. Acetyl-CoA is used in the tricarboxylic acid cycle for the production of reducing power that leads to the generation of cellular energy (Michal, 1999). The protein encoded in ORF-O is apparently linked to the ß-oxidation pathway.

O

kD

is

(G

aci

et a

in t

(Gus

ubiq

simil

meth

The l

dicarb

produ

a sho

metab

acid

weigh

dipic

Z225

first i

acid a

ORF-P encodes a 351 amino acid long protein with a predicted molecular weight of 38.2 kD. Mutations in this ORF impaired production of PDTC and the degradation of CCl_4. It is similar to hydroxyneurosporene methyltransferase from *Rhodobacter spheroides* (GenBank accession number P54906, 43% positives and 31% identities over a 180 amino acid overlap with a probability of 10^{-9}) involved in the biosynthesis of carotenoids (Lang *et al.*, 1995). Several sequences similar to motifs found in methyltransferases were seen in the protein encoded by ORF-P. Some of the motifs include RNA methyltransferase (Gustafsson *et al.*, 1991), C-5 cytosine-specific methylase (Som *et al.*, 1987), and ubiquinone – menaquinone methyltransferase (Daniels *et al.*, 1992). The sequence similarity and motif information suggest that the protein encoded in ORF-P is a putative methyltransferase.

The last intermediate in the production of PDTC should be dipicolinic acid (pyridine-2,6-dicarboxilic acid, CAS registration number 499-83-2). Dipicolinic acid is a compound produced by *B. subtilis* and a few other bacteria genera, and its production is confined to a short period during bacterial sporulation. It is not essential for the structure or metabolism of growing cells, but its absence results in heat-sensitive spores. Dipicolinic acid combines with calcium in the endospore core, representing about 10% of the dry weight of the endospore (Paulus, 1993). Dipicolinic acid is produced by the heterodimer dipicolinate synthase (DpaA/DpaB, also called SpoV/SpoF, GenBank accession number Z22554), in a single step, from 2,3-dihydrodipicolinate. 2,3-dihydrodipicolinate is the first intermediate in the biosynthesis branch leading to the production of diaminopimelic acid and L-lysine from L-aspartate (Chen *et al.*, 1993; Daniel & Errington, 1993) (Figure

3.

fr

sy

sh

3.7). None of the predicted proteins found in the 8.3 kb *Eco*RI fragment or the 25.7 kb fragment identified by Lewis show similarity to dipicolinate synthase. If *P. stutzeri* KC synthesizes dipicolinic acid from 2,3-dihydrodipicolinate, a different set of enzymes should be involved in the process.

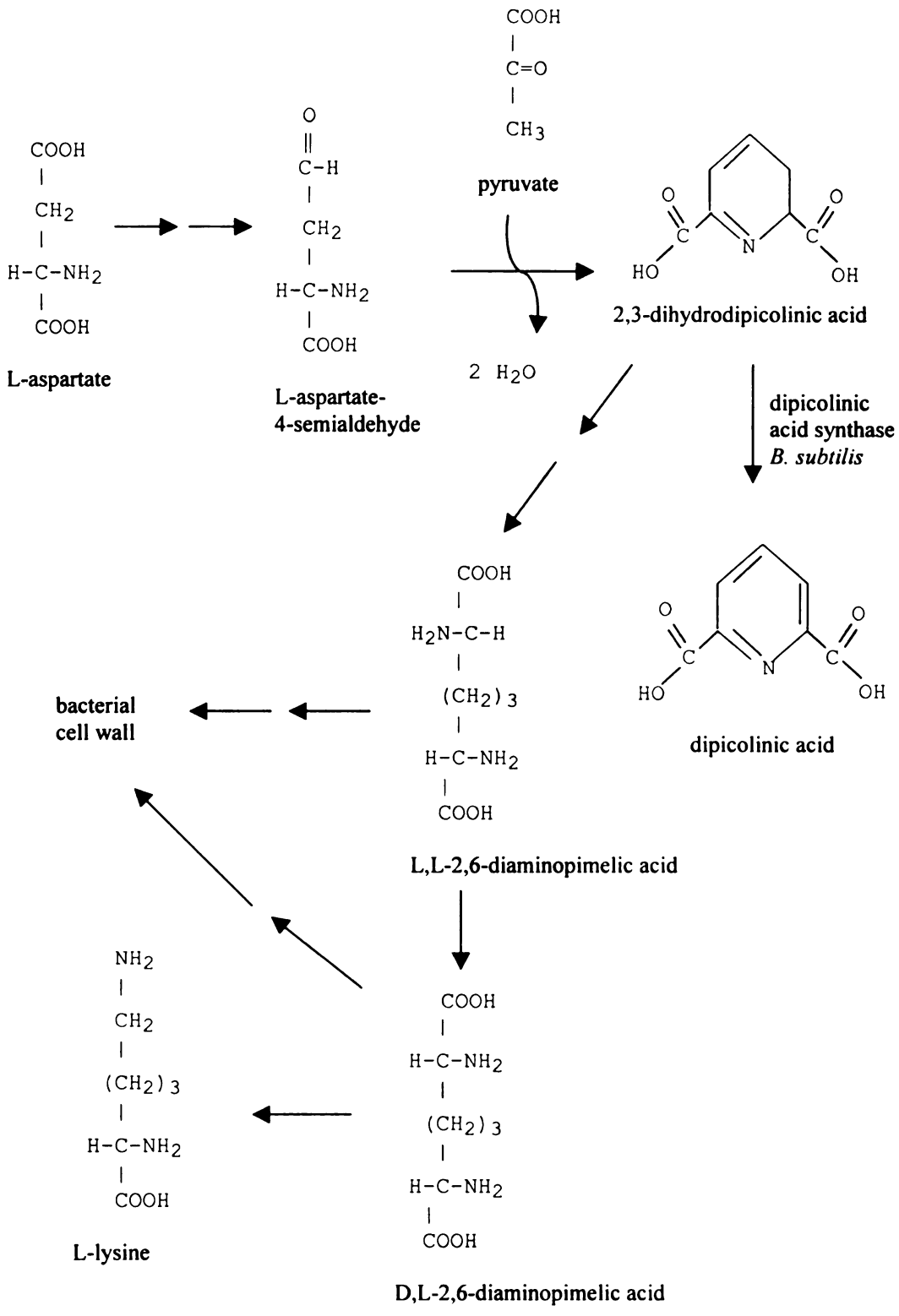


Figure 3.7 Pathway leading to bacterial cell wall precursors and L-lysine from L-aspartate. The branch leading to dipicolinic acid in *B. subtilis* is also provided. Modified from (Michal, 1999) and (Paulus, 1993).

Even though there is no empirical data to explain how the essential ORFs found by Sepúlveda and by Lewis are involved in the PDTC biosynthesis pathway, a speculative pathway can be proposed based on the similarity profiles. Lewis *et al.* (2000) proposed that ORF-2435, ORF-3626 and ORF-4099 (Lewis' ORFs F-H) serve to effect sulfur transfer to an oxygen-substituted (acyl or hydroxyl) carbon, and that ORF-6289 (Lewis' ORF-J) activates an acyl-group by adenylation. They proposed that the thiocarboxylate groups of PDTC are formed by condensation of a sulfur carrier (ORF-4099) and the adenylated precursor.

I propose a more explicit pathway that takes into consideration motif information. An overview is provided in Figure 3.8. A detailed explanation of a speculative biosynthesis pathway is proposed in Figure 3.9 and in the subsequent paragraphs.

Figure
explanat
3.5.

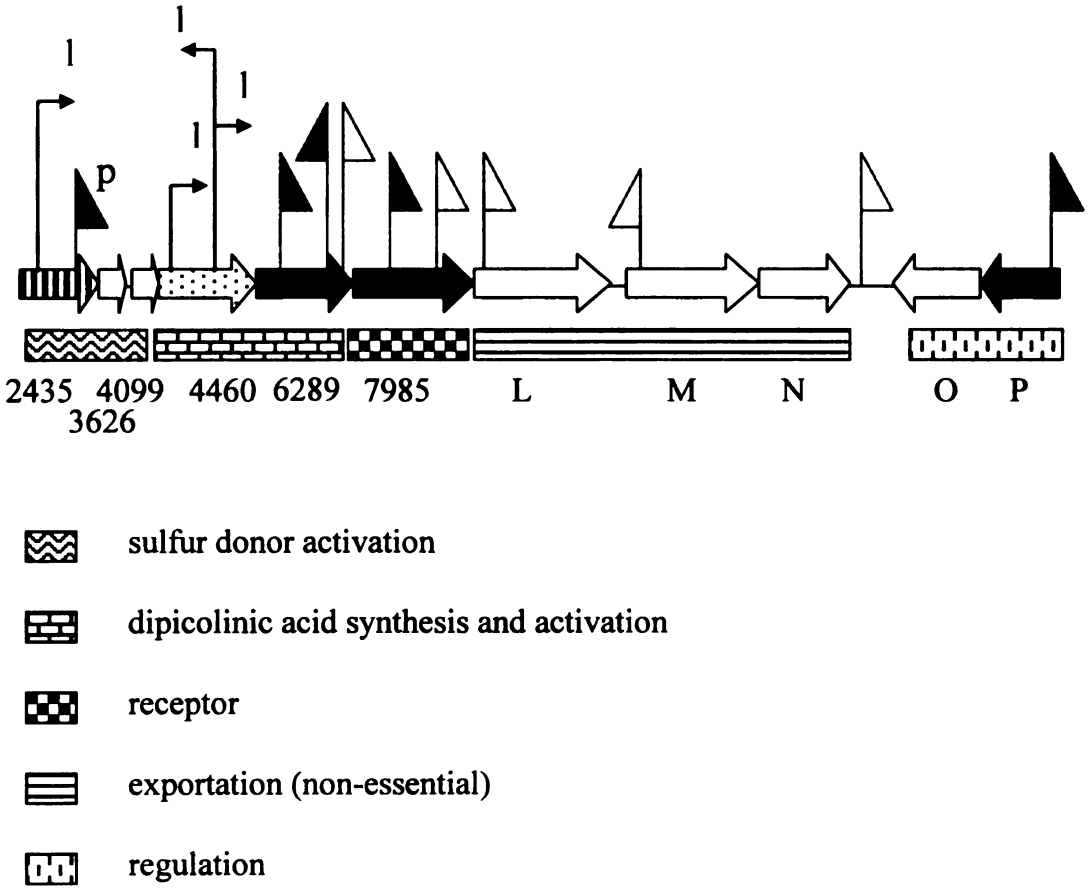
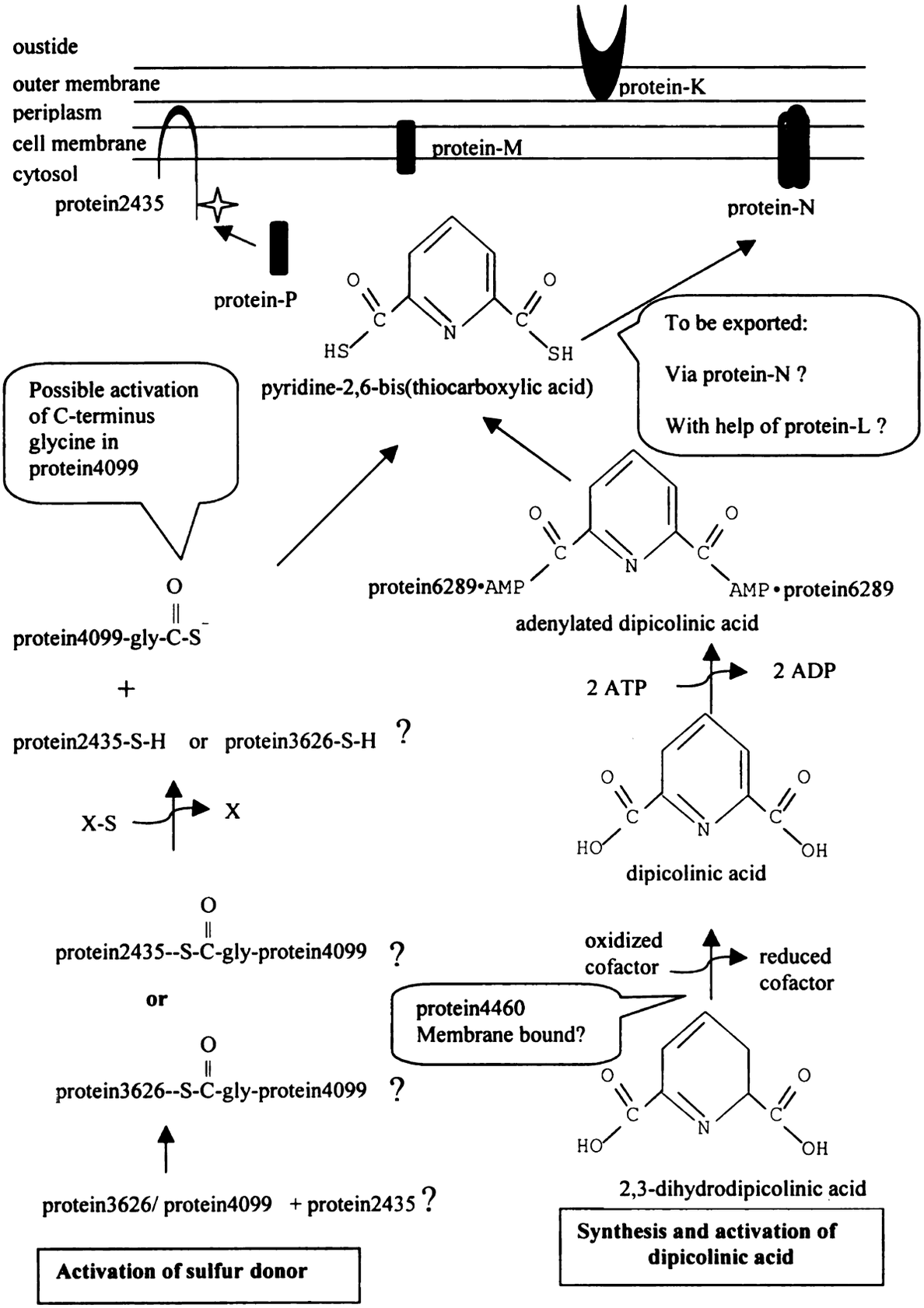


Figure 3.8 Overview of the proposed biosynthesis pathway for PDTC. A detailed explanation of the information provided in this figure is provided in the caption of Figure 3.5.

Figure 3.9 Speculative pathway for the synthesis of PDTC in *Pseudomonas stutzeri* strain KC. Upon the detection of environmental stimuli by ORF-2435, ORF-P methylates the C-terminus of ORF-2435 (indicated by a star), triggering the PDTC biosynthetic pathway. ORF's 3626 and 4099 activate the sulfur donor while ORF's 4460 and 6289 synthesize and activate dipicolinic acid from 2,3-dihydrodipicolinate. PDTC is exported out of the cell, possibly by ORF-N with help of ORF-L. For simplicity, ORF numbers are used to represent the polypeptides encoded by the corresponding ORF.


oustide
outer membrane
periplasm
cell membrane
cytosol
protein2435
protein-P
protein-K
protein-M
protein-N
pyridine-2,6-bis(thiocarboxylic acid)
To be exported:
Via protein-N ?
With help of protein-L ?
Possible activation of C-terminus glycine in protein4099
protein4099-gly-C-S
+
protein2435-S-H or protein3626-S-H ?
X-S
X
protein6289•AMP
AMP•protein6289
adenylated dipicolinic acid
2 ATP
2 ADP
dipicolinic acid
protein2435--S-C-gly-protein4099 ?
or
protein3626--S-C-gly-protein4099 ?
protein3626/ protein4099 + protein2435 ?
oxidized cofactor
reduced cofactor
protein4460
Membrane bound?
2,3-dihydrodipicolinic acid
Activation of sulfur donor
Synthesis and activation of dipicolinic acid

Th

in F

Sepu

pred

term

the I

termi

The f

produ

The s

(Lewi

sulfur

need c

The p

hydrog

dihydr

the su

dehydr

The pro

is highl

The mutagenesis experiment performed by Lewis indicates that the first protein involved in PDTC synthesis is encoded by ORF-2435 (Lewis' ORF-F) which was mutated by Sepúlveda and by Lewis. This gene encodes for an integral membrane protein with two predicted transmembrane helices. It has a sequence similar to motifs found near the C-terminus of chemotaxis sensors that facilitate sensory adaptation through the variation of the level of methylation of glutamate and glutamine residues located towards the C-terminus of the protein. The level of methylation is regulated by a methyltransferase. The function of this protein may be to sense the levels of the stimuli that trigger PDTC production from 2,3-dihydrodipicolinate.

The similarity profiles obtained for the proteins encoded by ORF-3626 and ORF-4099 (Lewis' ORFs G and H) suggest that these proteins may work together in the process of sulfur transfer. The interactions between protein-3626, protein-4099 and protein-2435 need clarification and may be the subject of future research.

The protein encoded by ORF-4460 (Lewis' ORF-I) is similar to proteins that remove hydrogen or water from their substrates, yielding a double bond. The conversion of 2,3-dihydrodipicolinate to dipicolinic acid requires the removal of two hydrogen atoms and the subsequent formation of a double bond. ORF-4460 may encode for the dehydrogenase involved in that process.

The protein encoded by ORF-6289 (Lewis' ORF-J) has a perfect AMP-binding motif and is highly similar to enzymes that activate siderophore precursors by adenylation. This

pro

aci

thio

tran

proc

The

trans

se. C

boun

simila

chape

is sim

essent

perme

aminot

transpo

facilita

ORF-O

involve

has a pe

ß-oxida

protein may be responsible for the adenylation of the two carboxyl groups of dipicolinic acid, activating this compound for the transformation of the carboxyl groups into thiocarboxyl groups. Once the adenylated intermediate is formed, sulfur can be transferred from the protein-3626/protein-4099 complex to the adenylated intermediate, producing PDTC.

The proteins encoded downstream of ORF-6289 appear to be involved in PDTC transport, reception and synthesis regulation, rather than in the biosynthetic pathway per se. ORF-K encodes for a TonB-dependent siderophore receptor. This outer membrane bound protein may be responsible for bringing metal-PDTC into the cell. Protein-L is not similar to known proteins but contains sequences similar to motifs found in molecular chaperons and proteins that facilitate translocation through the cell membrane. Protein-M is similar to aminotransferases and has a perfect motif for pyridoxal phosphate, an essential cofactor in many aminotransferase reactions. ORF-N encodes for a putative permease that facilitates solute transport through the cell membrane. The role of an aminotransferase is not clear at this point, but the permease may be responsible of transporting PDTC out of (and maybe into) the cell with the help of cytoplasmic facilitators like Protein-L.

ORF-O and ORF-P are transcribed in the opposite orientation than the rest of the genes involved in PDTC production. Protein-O is similar to acyl-CoA dehydrogenases and it has a perfect acyl-CoA binding site. Such an enzyme will be involved in the process of ß-oxidation of fatty acids. The function of this ORF in PDTC production/processing is

not

poss

Prot

knov

synt

initi

biosy

respo

plasr

(Figu

are c

produ

One

Sepúl

half c

systen

a mod

retaini

Tn5::

trappe

compe

assay c

not clear. On the other hand, the physiological roles of PDTC are yet to be elucidated. A possible role for protein-O might me assigned when more information is available. Protein-P is similar to methyltransferases and has sequences similar to motifs found in known methyltransferases. Protein-P may be involved in the regulation of PDTC synthesis. This putative methylase may modify Protein-2435 (Lewis' ORF-F) in order to initiate or attenuate PDTC synthesis. Protein-2435 is the first enzyme in the PDTC biosynthesis pathway and has sequences similar to chemotaxis transducers that trigger responses to environmental stimuli, depending on their level of methylation. When plasmids M22 and JS68 containing a partial fragment of T31 lacking ORF's O and P (Figure 7.6) are introduced into strain CTN1, only 10% or no CCl_4 degradation capacity are observed, respectively, indicating that proteins O and P are important for PDTC production. Their possible roles as regulators come to mind once again.

One of the main differences in the mutants found by Sepúlveda and by Lewis is that Sepúlveda was unable to find mutants corresponding to the proteins found in the upper half of the pathway. This result can be explained when the differences between their systems and experimental conditions are taken into consideration. Sepúlveda started with a modified *P. stutzeri* KC strain capable of growing in the presence of rifampicin while retaining the wild type CCl_4 degradation capacity. The bacterium was mutated with Tn*5*::*lux* and exposed to $^{14}CCl_4$ for 5 days before ^{14}C-labeled non-volatile products trapped in the culture medium were analyzed. If other receptors and permeases can compensate for the lack of a functional protein-K or protein-N, for example, Sepúlveda's assay could not be used to identify such a mutant, even if the affinity of the other proteins

for PD

over a l

CCl_4 d

compen

compen

mutatin

and utili

Sepúlve

Tn*5*::*lux*

reasons.

for *P. st*

interesti

that lost

the Tn*5*

construc

receptor

some of

producti

As indic

able to

genes in

for PDTC is lower, because the assay only shows the non-volatile product accumulated over a long period of time. Lewis reported that mutations in ORFs L and M do not affect CCl_4 degradation, which suggests that other proteins present in *P. stutzeri* strain KC can compensate for the lack of functional proteins encoded by ORFs L and M. The compensatory action of other proteins present in the strain KC genome could be tested by mutating ORFs K, L, M or N in wild type strain KC and monitoring PDTC production and utilization by these cells.

Sepúlveda tried to introduce pRK311, pT31, pRKblue and pRKblue8.3 into the four Tn*5*::*lux* transpositional mutants but the mutants did not accept any DNA, for unknown reasons. When these plasmids were introduced into *P. stutzeri* CCUG 11256 (type strain for *P. stutzeri*), only the cells harboring pT31 were able to degrade CCl_4. It would be interesting to see if strain CCUG 11256 and strain CTN1 (the spontaneous KC mutant that lost a 170 kb fragment) are able to degrade CCl_4 when the 8.3 kb *Eco*RI mutated in the Tn*5*::*lux* mutants and Lewis' ORF's O and P are introduced into the cell. Such a construct will leave out the upper half of the PDTC biosynthesis pathway encoding for a receptor and a permease. If these experiments are successful, they would confirm that some of the genes present in the upper half of the pathway are not essential for PDTC production and CCl_4 degradation.

As indicated in the results section, *P. putida* DSM 3601, a PDTC producing bacteria, was able to degrade CCl_4 but its DNA did not hybridize with a 3.4 kb probe containing the genes interrupted in KC657, KC1896, KC2753 and KC3164. No amplification of strain

DSM

KC

resul

diffe

synth

The

inves

how

resear

P. stu

DSM 3601 DNA was observed when primers that amplify the genes interrupted in strain KC were used, even when the annealing temperature was decreased by 5 °C. These results suggest that *P. stutzeri* strain KC and *P. putida* DSM 3601 produce DPTC using different mechanisms. The elucidation of the similarities and differences of the PDTC synthesis pathways of these organisms may be the subject of future research.

The genetic information presented in this chapter has opened new doors for the investigation of the physiological role of PDTC, the mechanisms of PDTC regulation and how this molecule can be efficiently used for the decontamination of CCl_4. More research is needed in order to fully understand the fascinating capabilities of PDTC and *P. stutzeri* strain KC.

Adhy

Altsc

Attw

Bate

Beni

Blac

Blat

REFERENCES

Adhya, S. & Gottesman, M. (1978). Control of transcription. *Annu Rev Biochem* **47,** 967-996.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST; a new generation of protein database search programs. *Nucleic Acids Res* **25,** 3389-3402.

Attwood, T. K., Flower, D. R., Lewis, A. P., Mabey, J. E., Morgan, S. R., Scordis, P., Selley, J. & Wright, W. (1999). PRINTS prepares for the new millenium. *Nucleic Acid Res* **27,** 220-225.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D. & Sonnhammer, E. L. (1999). Pfam3.1: 1313 multiple alignments match the majority of proteins. *Nucleid Acids Res* **27,** 260-262.

Bennasar, A., Rosselló-Mora, R. A., Lalucat, J. & Moore, E. R. (1996). 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stutzeri* and proposal of *Pseudomonas balearica* sp. nov. *Int J Syst Bacteriol* **46,** 200-205.

Black, T. A., Cai, Y. & Wolk, C. P. (1993). Spatial expression and autoregulation of *het*R, a gene involved in the control of heterocyst development development in *Anabaena. Mol Microbiol* **9,** 77-84.

Blattner, F. R., Plunkett, G. I. I. I., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, F. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. &

Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277,** 1453-1474.

Boyd, A., Kendall, K. & Simon, M. I. (1983). Structure of the serine chemoreceptor in *Escherichia coli*. *Nature* **301,** 623-626.

Braun, V. & Hantke, K. (1991). Genetics of bacterial iron transpor. In *Handbook of Microbial Iron Chelates* (ed. G. Winkelmann), pp. 107-138. CRC Press, Boca Raton, FL.

Braun, V. (1995). Energy-coupled transport and signal transduction through the Gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiol Rev* **16,** 295-307.

Brendel, V. & Trifonov, E. N. (1984). A computer algorithm for testing potential prokaryotic terminators. *Nucleic Acids Res* **12,** 4411-4427.

Butler, B. A. (1998). Sequence analysis using GCG. *Meth Biochem Anal* **39,** 79-97.

Carlson, C. & Ingraham, J. (1983). Comparison of denitrification by *Pseudomonas stutzeri*, *Pseudomonas aeruginosa* and *Paracoccus denitrificans*. *Appl Environ Microbiol* **45,** 1247-1253.

Chen, N. Y., Jiang, S. Q., Klein, D. A. & Paulus, H. (1993). Organization and nucleotide sequence of the *Bacillus subtilis* diaminopimelate operon, a cluster of genes encoding the first three enzymes of diaminopimelate synthesis and dipicolinate synthase. *J Biol Chem* **268,** 9448-9465.

Corpet, F., J, Gouzy, J. & Kahn, D. (1999). Recent improvements of the ProDom database of protein domain families. *Nuclei Acid Res* **27,** 263-267.

Cridd

Cserz

Daniel

Daniels

Devere

Ditta,

Dybas,

Eiglme

Criddle, C. S., DeWitt, J. T., Grbic-Galic, D. & McCarthy, P. L. (1990). Transfromation of carbon tetrachloride by *Pseudomonas* sp. strain KC under denitrification conditions. *Appl Environ Microbiol* **56,** 3240-3246.

Cserzo, M., Bernassau, J., Simon, I. & Maigret, B. (1994). New alignment strategy for transmembrane proteins. *J Mol Biol* **243,** 388-396.

Daniel, R. A. & Errington, J. (1993). Cloning, DNA sequence, functional analysis and transcriptional regulation of the genes encoding dipicolinic acid acid synthetase required for sporulation in *Bacillus subtilis*. *J Mol Biol* **232,** 468-483.

Daniels, D. L., Plunkett, G. I., Burland, V. D. & Blattner, F. R. (1992). Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. *Science* **257,** 771-778.

Devereaux, J., Haeberli, P. & Smithies, O. (1984). A comparative set of sequence analysis programs for the VAX. *Nucleic Acid Res* **12,** 387-395.

Ditta, G., Schmidhauser, T., Yokobson, E., Lu, P., Liang, X. W., Finlay, D. H., Guiney, D. & Helinski, D. R. (1985). Plasmids related to the broad host vector pRK290 useful for gene cloning and for monitorig gene expression. *Plasmid* **13,** 149-153.

Dybas, M. J., Tatara, G. M. & Criddle, C. S. (1995) Localization and characterization of the carbon tetrachloride transformation activity of *Pseudomonas sp.* strain KC. *Appl Environ Microbiol* **61,** 758-762.

Eiglmeier, K., Boos, W. & Cole, S. (1987). Nucleotide sequence and transcriptional startpoint of the *glp*T gene of *Escherichia coli*: extensive sequence homology of

the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system. *Mol Microbiol* **1,** 251-258.

Fernández-Piñas, F. , Leganés, F. & Wolk, C. P. (1994). A third genetic locus required for the formation of heterocysts in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **176,** 5277-5283.

Gray, G. L., Smith, D. H., Baldridge, J. S., Harkins, R. N., Vasil, M. L., Chen, E. Y. & Heyneker, H. L. (1984). Cloning, nucleotide sequence, and expression in *Escherichia coli* of the exotoxin A structural gene of *Pseudomonas aeruginosa. Proc Nat Acad Sci USA* **81,** 2645-2649.

Gustafsson, C., Lindstroem, P. H. R., Hagervall, T. G., Esberg, K. B. & Bjoerk, G. R. (1991). The *trm*A promoter has regulatory features and sequence elements in common with the rRNA P1 promoter family of *Escherichia coli. J Bacteriol* **173,** 1757-1764.

Gutiérrez, C., Barondess, J., Manoil, C. & Beckwith, J. (1987). The use of transposon Tn*pho*A to detect genes for cell envelope proteins subject to a common regulatory stimulus. *J Mol Miol* **195,** 289-297.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166,** 557-580.

Hansen, F. G., Hansen, E. B. & Atlung, T. (1985). Physical mapping and nucleotide sequence of the *rnp*A gene that encodes the protein component of ribonuclease P in *Escherichia coli. Gene* **38,** 85-93.

Hemmingsen, S. M., Woolford, C., Van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988). Homologous plant

and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333,** 330-334.

Henikoff, J. G., Henkoff, S. & Pietrokovski, S. (1999). New features in the BLOCKS database servers. *Nucleic Acids Res* **27,** 226-228.

Hess, J., Wels, W., Vogel, M. & Goebel, W. (1986). Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with corresponding chromosomal hemolysin sequence. *FEMS Microbiol Lett* **34,** 1-11.

Hirasawa, K., Ichihara, M. & Okanishi, M. (1985). Nucleotide Sequence of Mec^+ gene region of *Streptomyces kasugaensis*. *J Antibiot (Tokyo)* **38,** 1795-1798.

Hirokawa, T., Boon-Chieng, S. & Mataku, S. (1998). Classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14,** 378-379.

Hofmann, K. & Stofell, W. (1993). Tmbase - a database of membrane spanning protein segments. *Biol Chem (Hoppe-Seyler)* **347,** 166.

Hoffmann, K., Bucher, P., Falquet, L. & Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Res* **27,** 215-219.

Javanovic, G., Kostic, T., Jankovic, M. & Savic, D. J. (1994). Nucleotide sequence of the *Escherichia coli* K-12 histidine operon revisited. *J Molec Biol* **239,** 433-435.

Kadner, R. (1990). Vitamin B_{12} transport in *Escherichia coli*: energy coupling between membranes. *Mol Microbiol* **4,** 2027-2033.

Klebba, P. E., Rutz, J. M., Liu, J. & Murphy, C. K. (1993). Mechanism of TonB-catlyzed iron transport through the enteric bacterial cell envelope. *J Bioenerg Biomembr* **25,** 603-611.

Lang, H. P., Cogdell, R. J., Takaichi, S. & Hunter, C. N. (1995). Complete DNA sequence, specific Tn*5* insertion map and gene assignment of the carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*. *J Bacteriol* **177,** 2064-2073.

Larsen, R. A., Foster-Hartnett, D., McIntosh, M. A. & Postle, K. (1997). Regions of E coli TonB and FepA proteins essential for in vivo physical interactions. *J Bacteriol* **179,** 3213-3221.

Lee, C. H., Lewis, T. A., Paszczynski, A. & Crawford, R. L. (1999). Identification of an extracellular catalyst of carbon tetrachloride dehalogenation from *Pseudomonas* stutzeri strain KC as pyridine-2,6-bis(thiocarboxylate). *Biochem Biophys Res Commun* **261,** 562-566.

Lewis, T. A., Cortese, M. S., Sebat, J. L., Green, T. L. & Crawford, R. L. (2000). A *Pseudomonas stutzeri* gene cluster encoding the biosynthesis of the CCl_4 - dechlorinating agent pyridine-2,6-bis(thiocarboxilic acid). *Environ Microbiol* **In Press**.

Link, A. J., Robison, K. & Church, G. M. (1997). Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* K-12. *Electrphoresis* **18,** 1259-1313.

Michal, G. (1999). *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*. John Wiley & Sons, Inc., New York.

Moeck, G. S. & Coulton, J. W. (1998). TonB-dependent iron acquisition: mechanism of siderophore-mediated active transport. *Mol Microbiol* **28,** 675-681.

Nakai, K. & Kaneshisa, M. (1991). Expert systems for predicting protein localization sites in Gram-negative bacteria. *Proteins: Structure, Function and Genetics* **11,** 95-110.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10,** 1-6.

Ockels, W., Römer, A. & Budzikiewicz, H. (1978). An Fe(II) complex pryridine-2,6-di-(monothiocarboxylic acid) - a novel bacterial metabolic product. *Tetrahedron Lett* **36,** 3341-3342.

Ogasawara, N., Nakai, S. & Yoshikawa, H. (1994). Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res* **1,** 1-14.

Paulus, H. (1993). Biosynthesis of the aspartate family of amino acids. In *Bacillus subtilis and other Gram-positive Bacteria: Biochemistry,Physiology and Molecular Genetics* (ed. A. Sonenshein, J. A. Hoch and R. Losick), pp. 237-267. American Society for Microbiology, Washington, D C.

Platt, T. (1986). Transcription terminator and the regulation of gene expression. *Annu Rev Biochem* **55,** 339-372.

Postle, K. (1990). TonB and the Gram-negative dilemma. *Mol Microbiol* **4,** 2019-2025.

Rajagopalan, K. V. (1996). Biosynthesis of the molybdenum cofactor. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (ed. F. C. Neidhardt), pp. 674-679. ASM Press, Washington DC.

Rajagopalan, K. V. (1997). Biosynthesis and processing of the molybdopterin cofactors. *Biochem Soc Trans* **25,** 757-761.

Rakin, A., Saken, E. & Harmsen, D. (1994). The pesticin receptor of *Yersinia enterocolitica*: a novel virulence factor with dual function. *Mol Microbiol* **19,** 253-263.

Rakin, A., Urbitsch, P. & Heesemann, J. (1995). Evidence for two evolutionary lineages of highly pathogenic *Yersenia* species. *J Bacteriol* **177,** 2292-2298.

Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H. & Hopwood, D. A. (1996). A set of ordered cosmids an a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3 (2) chromosome. *Mol Microbiol* **21,** 77-96.

Reese, M. G. & Eeckman, F. H. (1995). Novel neural network algorithms for improved eukaryotic promoter site recognition. In *The 7th International Genome Sequencing and Analysis Conference* (ed. L. Hunter and T. E. Klein), Hilton Head Island, South Carolina.

Rosselló, R. A., García-Valdés, E., Lalucat, J. & Ursing, J. (1991). Genotypic and Phenotyic Diversity of Pseudomonas stutzeri. *Syst Appl Microbiol* **14,** 150-157.

Rosselló-Mora, R. A., Lalucat, J., Timmis, K. N. & Moore, E. R. B. (1996). Strain JM300 represents a new genomovar within *Pseudomonas stutzeri*. *Syst Appl Microbiol* **19,** 596-599.

Rowland, B. M., Grossman, T. H., Osburne, M. S. & Taber, H. W. (1996). Sequence and genetifc organization of a *Bacillus subtilis* operon encoding 2,3-dehydroxybenzoate biosynthetic enzymes. *Gene* **178,** 119-123.

Rusnak, F., Faraci, W. S. & Walsh, C. T. (1989). Subcloning, expression, and purification of the enterobactin biosynthetic enzyme 2,3-dihydroxybenzoate-AMP ligase: demonstration of enzyme-bound (2,3-dihydroxybenzoyl)adenylate product. *Biochemistry* **28,** 6827-6835.

Sambrook, J., Fritsch, E. L. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain - terminating inhibitors. *Proc Nat Acad Sci USA* **74,** 5463-5467.

Sawers, G. & A, B. (1989). Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. *J Bacteriol* **171,** 2485-2498.

Scofield, M. A., Lewis, W. S. & Schuster, S. M. (1990). Nucleotide sequence of *Escherichia coli asn*B and deduced amio acid sequence of asparagine synthetase B. *J Biol Chem* **265,** 12895-12902.

Seeger, C., Poulsen, C. & Dandanell, G. (1995). Identification and characterization of genes (*xap*A, *xap*B and *xap*R) involved in xanthosine catabolism in *Escherichia coli. J Bacteriol* **177,** 5506-5516.

Sepúlveda-Torres, L. del C., Rajendran, N., Dybas, M. J. & Criddle, C. S. (1999). Generation and initial characterization of *Pseudomonas stutzeri* KC mutants with impaired ability to degrade carbon tetrachloride. *Arch Microbiol* **171,** 424-429.

Serino, L., Reimmann, C., Visca, P., Beyeler, M., Della Chiesa, V. & Haas, D. (1997). Biosynthesis pf pyochelin and dihydroaeruginoic acid requires the iron-regulated *pch*DCBA operon in *Pseudomonas aeruginosa. J Bacteriol* **179,** 248-257.

Shine, J. & Delgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. *Proc Nat Acad Sci USA* **71,** 1342-1346.

Simon, R., Priefer, U. & Pühler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Biotechnology* **1,** 784-791.

Som, S., Bhagwat, A. S. & Friedman, S. (1987). Nucleotide sequence and expression of the gene encoding the *Eco*RII modification enzyme. *Nucleic Acids Res* **15,** 313-332.

Sonnhammer, E. L. L., Von Heijne, G. & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. In *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (ed. J. Glasgow), pp. 175-182. AAAI Press.

Stanier, R., Palleroni, N. & Doudoroff, M. (1966). The aerobic pseudomonas: a taxonomic study. *J Gen Microbiol* **43,** 159-271.

Tatara, G. M., Dybas, M. J. & Criddle, C. S. (1993). Effects of medium and trace elements on kinetics of carbon tetrachloride transforamtion by *Pseudomonas* sp. strain KC. *Appl Environ Microbiol* **59,** 2126-2131.

Tatara, G. M., Dybas, M. J. & Criddle, C. S. (1995). Biofactor mediated transformation of carbon tetrachloride by diverse cell types. In *Bioremediation of Chlorinated Solvents*, vol 3(4) (ed. R. E. Hinchee, A. Leeson and L. Semprini), pp. 69-76. Battelle Press, Columbus, Ohio.

Taylor, S. V., Kelleher, N. L., Kindsland, C., Chiu, H.-J., Costello, C. A., Backstrom, A. D., McLafferty, F. W. & Begley, T. P. (1998). Thiamin biosynthesis in *Escherichia coli*. *J Biol Chem* **273,** 16555-16560.

Tosato, V., Albertini, A. M., Zotti, M., Sonda, S. & Bruschi, C. V. (1997). Sequence completion, identification and definition of the fengycin operon in *Bacillus subtilis* 168. *Microbiology* **143,** 3443-3450.

Tusnady, G. E. & Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283,** 489-506.

von Heijne, G. (1992). Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225,** 487-494.

Whitchurch, C. B. & Mattick, J. S. (1994). *Escherichia coli* contains a set of genes homologous to those involved in protein secretion, DNA uptake and assembly of type-IV fimbriae in other bacteria. *Gene* **150,** 9-15.

White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K. S., Aravind, L., Daly, M. J., Minton, K. W., Fleischmann, R. D., Ketchum, K. A., Nelson, K. E., Salzberg, S., Smith, H. O., Venter, J. C. & Fraser, C. M. (1999). Genome sequence of athe radioresistant bacterium *Deinococcus radiodurans* R1. *Science* **286,** 1571-1577.

Wolk, C. P., Cai, Y. & Panoff, J. M. (1991). Use of a tansposon with luciferase as a reporter to identify environmental responsive genes in a cyanobacterium. *Proc Nat Acad Sci USA* **88,** 5355-5359.

Yanish-Perron, C., Vierira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene* **33,** 103-119.

Yazyu, H., Shiota-Niiya, S., Shimamoto, T., Kanazawa, H. & Futai, M. (1984). Nucleotide sequence of the *mel*B gene and characteristics of deduced amino acid sequence of the melibiose carrier in *Escherichia coli. J Biol Chem* **259,** 4320-4326.

Zylstra, G. J., Olsen, R. H. & Paballou, D. P. (1989). Genetic organization and sequence of the *Pseudomonas cepacea* genes for the alpha and beta subunits of protocatechuate 3,4-dioxygenase. *J Bacteriol* **171,** 5915-5921.

CHAPTER 4

RECOMMENDATIONS FOR FUTURE RESEARCH

The main goal at the outset of this project was to characterize genes involved in the degradation of CCl_4 by *Pseudomonas stutzeri* strain KC, when the identity of the agent responsible for CCl_4 dechlorination was still unknown. I was the pioneer in the discovery of genes involved in the production of the compound that was later identified as pyridine-2,6-bis(thiocarboxylic acid) (PDTC) (Figure 1.1). The discovery of a spontaneous *Pseudomonas stutzeri* strain KC mutant unable to degrade CCl_4 by the Crawford group at the University of Idaho, provided both an independent corroboration of my work and additional supplementary information. Although our joint efforts constitute a significant advancement in the understanding of the CCl_4 degradation capacity of *P. stutzeri* strain KC, many questions remain to be answered by future investigations.

The hypothetical PDTC biosynthesis pathway provided in Chapter 3 needs to be empirically corroborated. The information gathered using the DNA and protein similarity searches, as well as the protein motif searches, indicate that open reading frames (ORFs) 3626 and 4099 may be responsible for donating the two sulfurs found in PDTC. A possible interaction with ORF-2435 cannot be discarded. The isolation and characterization of the proteins encoded by these genes will allow the design of experiments to test the interactions between the polypeptides as well as their sulfur-transfer capabilities.

The aforementioned pathway assumes that the carbon and nitrogen atoms in the PDTC ring come from 2,3-dihydrodipicolinic acid, the first intermediate in the branch leading to

the biosynthesis of L-lysine from L-aspartate (Figure 3.7). If this is the case, radiolabeled PDTC should be obtained when ^{14}C-labeled L-aspartate or pyruvate is supplied in the culture medium because these two compounds provide the atoms for 2,3-dihydrodipicolinic acid.

In the proposed pathway, ORF-4460 introduces a double bond into 2,3-dihydrodipicolinic acid to produce dipicolinic acid. The isolation of the protein would allow the *in vitro* characterization of the reaction in the presence of various oxidized cofactors. Computer programs could be used to model the active site of the protein and determine if 2,3-dihydrodipicolinic acid and oxidized cofactors would fit inside the active site. If this protein is involved in this particular step, the accumulation of 2,3-dihydrodipicolinic acid should be observed in strains KC1896, KC2753 and KC3164 because all of them were mutated in this ORF.

ORF-6289 encodes for a putative AMP-ligase probably involved in the activation of the two carboxyl-groups of PDTC by adenylation. The mutants impaired in ORF-6289 generated by the Crawford group should be unable to carry out this particular reaction and must accumulate dipicolinic acid in the cytoplasm. Once this protein is isolated, the reaction could be carried out *in vitro*, allowing the formation of adenylated dipicolinic acid and the isolation of the enzyme-bound adenylated dipicolinic acid. The adenylation of dipicolinic acid by other AMP-ligases could also be studied in similar assays.

Crude cell lysates of wild type strain KC and of mutants impaired in different steps of the PDTC biosynthesis pathway may be used to study the accumulation of biosynthetic intermediates. Such experiments may provide further information about gene function.

Lewis' ORF-K encodes for an outer membrane siderophore receptor that may be the receptor for PDTC. Since siderophore receptors may interact with structurally similar siderophores, it would be worth investigating the possibility of the interactions of ORF-K with PDTC analogues and the interactions of PDTC with the siderophore receptors that showed similarity to ORF-K at the DNA and protein levels. Similar experiments could be designed to determine if metal-loaded PDTC is taken into the cell by ORF-K (or proteins similar to ORF-K); or if PDTC remains in the culture medium while the metal is transported into the cytoplasm.

It is my belief that Lewis' ORF-P is involved in the regulation of PDTC synthesis by changing the state of methylation of ORF-2435, the first protein involved in PDTC biosynthesis. ORF-2435 is an integral membrane protein that has a sequence similar to chemotaxis transducers that are controlled by the level of methylation of some residues near their C-termini. The isolation of the proteins encoded by ORF-P and ORF-2435 may provide the framework for the study of interactions among them. Nevertheless, this task may be difficult because of the membrane topology of ORF-2435.

I was unable to identify mutants in the genes corresponding to Lewis' ORFs L and M and N. On the other hand, mutants in these ORFs did not affect PDTC degradation when the

Crawford group mutated them. I think these proteins are not essential for PDTC production because other proteins can carry out the same functions. Lewis' ORFs O and P appear to be essential, presumably due to their possible regulatory role. The substitutability hypothesis can be tested by the introduction of a plasmid containing the 8.3 kb *Eco*RI fragment harboring the genes directly involved in PDTC biosynthesis, ORF-O and ORF-P into the type strain of *P. stutzeri* and *P. stutzeri* CTN1. If this plasmid can confer the CCl_4 degradation capacity to the aforementioned strains, the substitutability hypothesis would be sustained.

The identification of the genes needed for the synthesis of PDTC also opens the door to investigations regarding the molecular mechanisms of transcriptional control of PDTC production. As indicated in Figure 3.6, a sequence very similar to the consensus sequence of the *E. coli* ferric uptake regulator (Fur, Fur box) was found in the promoter region of the PDTC operon. Fur is believed to bind to the Fur box when ferrous iron is abundant, in order to repress the transcription of genes needed for the biosynthesis and uptake of iron-scavenging compounds. Since PDTC is only produced when *P. stutzeri* strain KC is grown under iron-limiting conditions, the discovery of a Fur box in the promoter region of the PDTC operon is congruent with the mechanism of regulation of iron-scavenging processes. Furthermore, I demonstrated the iron-dependent promoter activation by following the expression of the promoterless reporter luciferase genes when strain KC was grown with different iron concentrations (Figure 2.1). Experiments designed to assess the effects of mutations in the Fur box on PDTC production will provide evidence of the involvement of Fur in the control of PDTC biosynthesis. The

PDTC promoter could be fused to a promoterless reporter gene whose expression is monitored under iron abundance or limitation, in hosts that contain a functional *fur* gene and in *fur*-minus hosts.

The discovery of a Fur box in the promoter region of the PDTC biosynthesis genes also has implications for the utilization of PDTC and *P. stutzeri* KC in CCl_4 remediation projects. If iron-limitation cannot be attained, it is imperative to place the PDTC biosynthesis genes under the control of a promoter that would be activated under the particular circumstances prevailing in the contaminated samples. The transfer of the PDTC biosynthesis genes into hosts already adapted to the environmental conditions that predominate in the contaminated site should also be considered.

Sequence information suggests that the PDTC operon starts in ORF-2435 and may be transcribed until a hairpin structure identified in positions 20,593 to 20,616 between Lewis' ORF-N and ORF-O. ORFs O and P appear to be part of a transcript independent from the larger PDTC operon. There is a possible promoter from position 24,477 to position 24,432 and a hairpin structure from position 20,336 to position 20,316 (between ORFs O and N). This information can be empirically verified by the analysis of mRNA isolated under the conditions leading to PDTC biosynthesis, using the PDTC genes as probes for Northern blots or reverse transcription. The exact transcript beginning could be determined by the primer extension technique or by S1 nuclease analysis. The position of the 3'-end of the transcript could be estimated by reverse transcription,

varying the primer closer to the 3'-end until no cDNA could be generated with the specific primer set.

The control of PDTC synthesis by feedback inhibition should be studied in detail. If PDTC inhibits its own synthesis, *P. stutzeri* strain KC cultures grown under iron limitation should not transcribe the PDTC genes if external PDTC is provided. This hypothesis can be tested with Northern blots or reverse transcription. If feedback inhibition is observed, interactions between PDTC and the enzymes involved in its biosynthesis should be empirically corroborated.

PDTC was originally discovered as a metabolite produced by *Pseudomonas putida* DSM 3601 grown under iron-limiting conditions. Even though *P. putida* DSM 3601 shares this metabolic capability with strain KC, the bacteria apparently use different mechanisms for PDTC production, as indicated by the DNA-based studies described in Chapter 3. The DNA from *P. putida* DSM 3601 and 8 *P. stutzeri* strains incapable of degrading CCl_4 did not amplify with polymerase chain reaction (PCR) primers designed for the amplification of ORF-2435 or ORF-4460, even when the annealing temperature was lowered by 5°C. The same negative result was observed on a Southern blot when a 3.5 kb fragment spanning from ORF-2435 to ORF-4460 was used to hybridize with 10 μg of genomic DNA from *P. putida* DSM 3601 and 8 *P. stutzeri* strains. These results should be verified by the utilization of different PCR primer sets and by low stringency Southern blots. The characterization of the genes involved in PDTC biosynthesis in *P. putida* DSM 3601 would allow a direct comparison of the pathway used by these bacteria for the

production and regulation of PDTC. The information obtained from these comparisons may increase our knowledge about the native role of PDTC.

The physiological experiments described in Chapter 1, along with the genetic studies depicted in Chapters 2 and 3 have greatly enhanced our understanding of the CCl_4 degradation capacity of *P. stutzeri* strain KC. The experiments suggested in Chapter 4 would further extend our comprehension of this process at the molecular level. The combined efforts of past and future experimentation will allow us to fully elucidate the fascinating CCl_4 degradation capacity of *P. stutzeri* strain KC.

APPENDIX A

PHYLOGENY AND TAXONOMY OF *PSEUDOMONAS STUTZERI* STRAIN KC

"*Pseudomonas* strain KC represents a new genomovar within *Pseudomonas stutzeri*"

to be submitted to

The International Journal of Systematic and Evolutionary Microbiology

Pseudomonas strain KC represents a new genomovar within ***Pseudomonas stutzeri***

Lycely del C. Sepúlveda-Torres[1], Jizhong Zhou[2], Caterina Guasp[3], Jorge Lalucat[3], David Knaebel[4*], Jody L. Plank[4†], and Craig S. Criddle[5‡]

[1] Department of Microbiology and National Science Foundation Center for Microbial Ecology, Michigan State University, East Lansing, MI 48823, USA

[2] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

[3] Microbiologia, Departament de Biologia, Universitat de les Illes Balears and Institut Mediterrani d'Estudis Avançats, Palma de Mallorca, Spain

[4] Biology Department, 5805 Clarkson University, Potsdam, NY 13699

[5] Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA

[*] Current address: Environmental Science Center, Syracuse Research Corporation, North Syracuse, NY 13215, USA

† Current address: Department of Biochemistry, Duke University, Durham, NC 27708, USA

‡ Author for correspondence: Craig S. Criddle. Tel: + 650 723 9032. Fax: + 650 725 9474
e-mail: criddle@ce.stanford.edu

Keywords: Pseudomonas KC, phylogeny and taxonomy, genomovar, carbon tetrachloride biodegradation, pyridine-2,6-bis(thiocarboxylate)

Subject category: Evolution, Phylogeny and Biodiversity

Running title: Phylogeny and taxonomy of *Pseudomonas* strain KC

Non-standard abbreviations: DMSO, dimethyl sulfoxide; FAME, fatty acid methyl ester; gv, genomovar; ITS1, 16S – 23S internally-transcribed spacer region; PDTC, pyridine-2,6-bis(thiocarboxylate); TSB, tryptic soy broth; U, units; UPGMA , unweighted pair-group mean analysis

SUMMARY

Pseudomonas sp. strain KC is a denitrifying aquifer isolate that produces and secretes pyridine-2,6-bis(thiocarboxylate) (PDTC), a compound that chelates copper to fortuitously transform carbon tetrachloride without producing chloroform. Although KC has been successfully used for full-scale bioremediation of carbon tetrachloride, its taxonomy has proven difficult to resolve, as it retains properties of both *P. stutzeri* and *P. putida*. In the present work, a polyphasic approach, comprising phenotypic characteristics (carbon substrate utilization patterns, antibiotic resistance profiles and composition of cellular fatty acids) and genotypic information (DNA-DNA hybridization, DNA fingerprinting, 16S rDNA sequencing, ITS sequencing and *gyrB* PCR) is used to establish that strain KC is a member of the species *Pseudomonas stutzeri*. Moreover, we conclude that strain KC represents a new genomovar (genomovar 9) within the species *P. stutzeri*.

INTRODUCTION

Bacterial strain KC (ATCC deposit no 55595, DSM deposit no 7136) is a denitrifying bacterium originally isolated from an aquifer in Seal Beach, California (Criddle *et al.*, 1990). Under iron-limiting conditions, strain KC induces genes for the production and secretion of pyridine-2,6-bis(thiocarboxylate) (PDTC), a molecule that chelates copper and can rapidly dechlorinate CCl_4 yielding CO_2 (~ 50%), and nonvolatile compounds (~ 50%), under anoxic conditions (Criddle *et al.*, 1990; Dybas *et al.*, 1995; Lee *et al.*, 1999; Lewis & Crawford, 1993; Sepúlveda-Torres *et al.*, 1999). This activity is important for bioremediation applications in aquifer sediments because it is rapid, with half-lives of only a few minutes (Tatara *et al.*, 1995) and occurs without accumulation of chloroform. Strain KC attaches to aquifer sediment, but it can also exist in a free-swimming, highly motile form that is chemotactic towards nitrate, and it can sustain dechlorination activity during migration (Witt *et al.*, 1999a; Witt *et al.*, 1999b). Emerson (1999) reported that strain KC reproducibly forms colonies of complex morphology on agar motility plates containing nitrate or nitrite. Five other species of pseudomonads tested under identical conditions were unable to form such complex colonies. Recent developments further underscore the unique environmental significance of this strain. Lewis *et al.* (2000) have reported that a laboratory culture of strain KC had spontaneously lost a 170 kb fragment containing genes necessary for PDTC biosynthesis on a 25 kb fragment of the lost DNA. This fragment was not detected in three other *P. stutzeri* strains. Strain KC has also assumed great significance for biotechnology because of its use in one of the first full-scale field aquifer bioaugmentation applications. Large volumes of strain KC were

grown on-site and injected into a CCl_4-contaminated aquifer in Schoolcraft, Michigan (Hyndman *et al.*, 2000). The resulting biocurtain for CCl_4 degradation has now been maintained for over two years, with efficient removal of CCl_4.

Strain KC was originally classified as a *Pseudomonas stutzeri* – like organism for its ability to reduce nitrate and use maltose, citrate, malonate and glyerol as carbon sources and a preliminary fatty acid profile (Criddle *et al.*, 1990) and some past results have referred to it as a *P. stutzeri*. Nevertheless, no exhaustive studies were performed to elucidate the strain KC taxonomy. The present investigation was performed to conclusively establish the systematic classification of strain KC based on physiological and genotypic studies. The results obtained from DNA-DNA hybridization, DNA fingerprinting, 16S rRNA sequence, ITS1 sequence analysis and *gyrB* PCR studies were combined with substrate utilization, antibiotic resistance and fatty acid methyl ester (FAME) analyses to establish that strain KC should be classified as a type strain for a novel *P. stutzeri* genomovar.

METHODS

Strains and growth conditions: The strains used in this study, their source of isolation and relevant references are provided in Table A.1. The bacteria were grown aerobically in tryptic soy broth (TSB), nutrient broth or LB medium at 30 °C (Sambrook *et al.*, 1989).

Table A.1 Bacterial strains used in this study

Strain	Other designations and origins of isolation	References
Pseudomonas sp. KC	ATCC* 55595, DSM* 7136, aquifer isolate, carbon tetrachloride degrader	(Criddle et al., 1990)
P. stutzeri CCUG* 11256^{T} gv 1†	ATCC 17588, Stanier strain 221, clinical isolate	(Stanier et al., 1966)
P. stutzeri ATCC 17591 gv 2†	Stanier strain 224, clinical isolate	(Stanier et al., 1966)
P. stutzeri DSM 50227 gv 3†	ATCC 11607, clinical isolate	(Van Niel & Allen, 1952)
P. stutzeri 19SMN4 gv 4†	DSM 6084, marine isolate, naphthalene degrader	(Rosselló et al., 1991)
P. stutzeri DNSP21 gv 5†	DSM 6082, wastewater isolate	(Rosselló et al., 1991)
P. balearica DSM 6083^{T} gv 6†	waste water isolate, naphtalene degrader	(Bennasar et al., 1996; Rosselló et al., 1991)
P. stutzeri DSM 50238 gv 7†	ATCC 17832, Stanier strain 419, soil isolate	(Stanier et al., 1966)
P. stutzeri JM300 gv 8†	DSM 10701, soil isolate	(Carlson & Ingraham, 1983)
P. putida ATCC 12633^{T}	DSM 50202, lactate enrichment	(Skerman et al., 1980; Stanier et al., 1966)
P. putida DSM 3601	tomato plant isolate, Produces 2,6-bis(pyridine thiocarboxylate)	(Ockels et al., 1978)

* ATCC, American Type Culture Collection, Rockville, MD, USA; CCM, Czechoslovak Collection of Microorganisms, Brno, Czechoslovakia; CCUG, Culture Collection University of Göteborg, Sweden; CECT, Colección Española de Cultivos Tipo, Valencia, Spain; DSM, Deutsche Sammlung von Mikroorganismen und Zellkulteren, Braunschweig, Germany; LMG, Laboratorium Microbiologie Rijksuniversiteit, Gent, Germany

† type strain of genomovar

T type strain

Determination of carbon source utilization: Three individual colonies of each strain were picked from fresh nutrient agar plates and used to inoculate 50 ml erlenmeryer flasks containing 10 ml of TSB. The cultures were grown at 30 °C and 150 r.p.m. for 20 h. Cultures were centrifuged at 2,500 × g for 10 min, washed twice in PBS and normalized to an optical density of 0•195 – 0•205 in a spectrophotomer at 600 nm. 150 µl aliquots were added to BIOLOG® GN2 plates (Biolog) and the plates were incubated at 30 °C for 48 h. Substrate consumption was followed by reading the plates in a microtiterplate reader at an optical density of 590 nm after 4, 24 and 48 h of incubation.

Antibiotic susceptibility test: Triplicate 20 h cultures in TSB were used to inoculate 150 mm × 15 mm nutrient agar plates, bi-directionally, using a sterile cotton swab. Filter disks (0•7 cm diameter) containing antibiotics were placed on the surface of the plates prior to incubation at 30 °C for 24 h. The following antibiotics were used for the susceptibility test: 10 µg ampicillin ml^{-1}, 10 units (U) penicillin G ml^{-1}, 20 µg amoxillin ml^{-1} and 10 µg ml^{-1}, streptomycin 10 µg ml^{-1}, clavulanic acid ml^{-1}, 30 µg erythromycin ml^{-1}, 30 µg tetracycline ml^{-1}, 1•25 µg trimethropin ml^{-1} and 23•75 µg sulfamethoxazole ml^{-1}, 1 µg oxacillin ml^{-1}, 30 µg cephalothin ml^{-1}, 30 µg vancomycin ml^{-1}, 2 µg clindamycin ml^{-1}, and 300 U polymyxin B ml^{-1}.

Fatty acid analysis: Bacterial isolates were grown overnight on tryptic soy agar at 28 °C and harvested with a sterile loop. Saponification, methylation and extraction were performed by Microbial ID using a previously described procedure (Sasser, 1990a).

Methylated fatty acids were separated by a gas chromatograph equipped with a flame ionization detector, in a 25 m × 0.2 mm phenyl methyl silicone fused silica capillary column using hydrogen as the carrier gas and nitrogen as the makeup gas. Numerical comparisons of the fatty acid profiles were carried out as previously described (Sasser, 1990b).

DNA fingerprinting: DNA to be used for fingerprinting was isolated using the GenomicPrep cells and tissue DNA isolation kit (Amersham Pharmacia Biotech) following the manufacturer's recommendations. PCR for REP, BOX and ERIC fingerprinting were performed in 25 µl reactions containing 50 ng DNA, 2 pmole each primer, 1•25 pmole each dNTP, 2 U *Taq* DNA polymerase, 20% v/v Gitschier buffer (Kogan *et al.*, 1987), 0•8% BSA, 10% v/v dimethyl sulfoxide (DMSO), and 4% v/v Tween 20. Primers REP1R-I and REP2-I (Versalovic *et al.*, 1991) were used for REP fingerprinting, primer BOXA1R (Versalovic *et al.*, 1994) while primers ERIC1R and ERIC2 (Versalovic *et al.*, 1991) were used for ERIC fingerprinting. PCR cycles for REP, ERIC and BOX PCR were as follows: initial denaturation at 95 °C for 7 min, 35 cycles consisting of 94 °C for 1 min, 44 °C for 1 min for REP (52 °C for ERIC or 53 °C for BOX) and 65 °C for 8 min. A single final extension step was performed at 65 °C for 15 min in order to assure chain termination. PCR products were separated by electrophoresis in 2% agarose (w/v) and stained with ethidium bromide. The GelCompar image analysis system (Applied Maths) was used to calculate similarities between all fingerprinting profiles. Dendrograms of relationships were deduced by the unweighted pair-group mean analysis (UPGMA) cluster algorithm (Vauterin & Vauterin, 1992).

DNA-DNA hybridizations: Nucleic acids from bacterial strains were isolated following the method of Marmur (1961). DNA-DNA hybridizations were performed, using a modification of the hydroxyapatite method described previously (Ziemke *et al.*, 1998). Reference DNA's were double-labeled with DIG-11-dUTP and biotin-16-dUTP using the nick-translation kit as recommended by the manufacturer (Boehringer Mannheim).

16S rRNA cloning, sequencing and analysis: The 16S rRNA gene was PCR amplified with modified universal eubacterial primers fD1 (5'-CCA TCG ATG TCG ACA GAG TTT GAT CCT GGC TCA G–3') and rP1 (5'-GAC TAG TGG ATC CAC GGT TAC CTT GTT ACG ACT T-3') (Zhou *et al.*, 1995) as previously described (Weisburg *et al.*, 1991). The sequence of the 16S rRNA gene of strain KC was submitted to GenBank under the accession numbers AF67960 and AF063219. Multiple sequence alignment was done with the PILEUP program in the Genetics Computer Group software package (Devereaux *et al.*, 1984). The alignment was edited for the appropriate analysis by using the SUBALIGN and GDE programs from the Ribosomal Database Project (Maidak *et al.*, 1999). The phylogenetic analyses were performed in the DNA distance program ARB using Neighbor-Joining with Felsenstein correction (Stunk *et al.*, 2000).

Amplification of the 16S rRNA gene with primers specific for *P. stutzeri* strains: A 1160 bp fragment of the 16S rRNA gene of strain KC and several *P. stutzeri* strains was amplified with the *P. stutzeri*- specific PCR primers fps150 (5'-GTG GGG GAC AAC GTT TC-3') and rps1271 (5'-CTA CGA TCG GTT TTA TGG-3') as previously

described (Bennasar *et al.*, 1998a). The PCR products were purified using the Wizard™ PCR preps DNA purification kit as recommended by the manufacturer (Promega). 10 µl of purified PCR products were restricted with 10 U of the endonuclease *Bam*HI, for 1 h at 37 °C in a total volume of 30 µl. The restricted fragments were separated by electrophoresis in a 1.5 % (w/v) agarose gel and stained with ethidium bromide.

Amplification and sequencing of the 16S-23S internally transcribed spacer region (ITS1): ITS1 was amplified by PCR with oligonucleotide primers 16F945 and 23R458 (Lane *et al.*, 1985) designed to anneal to conserved positions in the 3' and 5' regions of the bacterial 16S rRNA and 23S rRNA genes, respectively. Primers 16F945 (5'-GGG CCC GCA CAA GCG GTG G-3') and 23R458 (5'-CTT TCC CTC ACG GTA C-3') targeted positions 927-945 of the *Escherichia coli* 16S rRNA gene (Brosius *et al.*, 1978) and positions 458-473 of the *E. coli* 23S rRNA gene (Brosius *et al.*, 1980), respectively. PCR amplification cycles were performed as per Guasp *et al.* (2000). The sequence of the ITS1 region was determined by *Taq* cycle sequencing using fluorescent dye-labeled dideoxynucleotides. The primers used for sequencing were rrn16S (5'-GAA GTC GTA ACA AGG-3') and rrn23S (5'-CAA GGC ATC CAC C-3') (Jensen *et al.*, 1993) designed to anneal to conserved positions in the 3' and 5' regions of the bacterial 16S rRNA and 23S rRNA genes, respectively. Primers rrn16S (5'-GAA GTC GTA ACA AGG-3') and rrn23S (5'-CAA GGC ATC CAC CTG-3') targeted positions 1491-1505 of the *Escherichia coli* 16S rRNA gene (Brosius *et al.*, 1978) and positions 21-35 of the *E. coli* 23S rRNA gene (Brosius *et al.*, 1980), respectively. The ITS1 sequence of strain KC was submitted to GenBank under accession number (to be submitted). ITS1

sequences were aligned using the computer program CLUSTAL W (Thompson *et al.*, 1994), with a final manual adjustment (Rabaut, 1996). Evolutionary distances were calculated from pairwise sequence similarities (Jukes & Cantor, 1969) and estimations of relationships were generated using the Fitch program within the Phylogeny Inference Package (PHYLIP) (Felsenstein, 1989).

***gyrB*-based PCR amplification:** PCR amplification using primers specific for the *Pseudomonas putida gyr*B gene was performed as previously described (Yamamoto & Harayama, 1995).

RESULTS

Substrate utilization, antibiotic resistance and fatty acid analysis. Fourty of the 95 carbon sources tested were used differently by the bacterial strains, as seen in Table A.2. This result is consistent with previous reports that indicate the high degree of physiological heterogeneity within *P. stutzeri* strains (Palleroni *et al.*, 1970; Rosselló *et al.*, 1994a; Stanier *et al.*, 1966). Strain KC is similar to *P. stutzeri* strains in its ability to grow on dextrin, glycogen, maltose and α-ketobutyric acid and its incapacity of utilizing D-arabinose, D-sorbitol and phenyl ethylamine, three carbon sources used only by the *P. putida* strains. Strain KC was the only organism, of the eleven tested, capable of growing on m-inositol.

Even though significant differences in carbon source utilization were observed, the behavior of strain KC and *P. stutzeri* strains in the antibiotic resistance test and the fatty acid analysis was more homogeneous. As seen in Table A.3, the antibiotic susceptibility results obtained for strain KC coincided with the consensus for the majority of *P. stutzeri* strains and diverged from the *P. putida* pattern for any of the 12 antibiotic tested, except trimethroprin/sulfamethoxazole, an antibiotic combination that showed variability among the strains tested.

* *Pseudomonas stutzeri* strains tested: CCUG 1126^T (gv 1), ATCC 17591 (gv 2), DSM 50227 (gv 3), 19SMN4 (gv 4), DNSP21 (gv 5), DSM 50238 (gv 7), JM300 (gv 8): *Pseudomonas balearica* DSM 6083^T (gv 6); *Pseudomonas putida* strain tested: ATCC 12633^T (9), DSM 3601 (10)

† +, positive for substrate utilization; —, negative for substrate utilization; w, weak positive

Table A.2 Substrate utilization by strain KC and various *Pseudomonas stutzeri*, *P. balearica* and *P. putida* strains

BIOLOG	*Pseudomonas stutzeri* strains *								*P. balearica* *	*P. putida* *	
carbon source	KC	1	2	3	4	5	7	8	6	9	10
dextrin	+ †	+	+	+	+	+	+	+	+	— †	—
glycogen	+	+	+	+	+	+	+	+	+	—	—
L-arabinose	—	—	—	—	—	—	—	—	—	—	+
D- arabitol	+	—	—	—	—	—	—	+	—	—	—
D-fructose	+	—	+	—	—	—	—	+	—	+	—
m-inositol	+	—	—	—	—	—	—	—	—	—	—
maltose	+	+	+	+	+	+	+	+	+	—	—
D-mannitol	+	—	+	w †	—	+	—	+	—	—	—
D-mannose	—	—	—	—	—	—	—	+	—	—	+
D-sorbitol	—	—	—	—	—	—	—	—	—	+	—
trehalose	+	—	w	—	—	—	—	w	—	—	—
formic acid	w	—	w	w	—	—	w	+	w	+	+
α-hydroxybutiric acid	+	+	w	w	+	—	+	+	w	+	—
γ-hydroxybutiric acid	—	—	w	w	—	—	w	—	—	—	—
p-hydroxyphenylacetic acid	—	+	—	—	—	w	w	w	w	+	w
itaconic acid	+	+	+	+	+	+	+	+	+	+	—
α-keto butyric acid	+	+	w	w	+	w	+	+	w	—	—
α-keto valeric acid	w	—	—	—	+	—	—	+	—	+	—
propionic acid	+	+	+	+	+	—	+	+	+	+	+
D-saccharic acid	+	+	w	+	—	w	+	w	w	+	+
sabacic acid	—	+	—	+	+	—	—	+	—	—	—
succinamic acid	+	—	—	—	—	—	—	+	—	—	—
glucuronamide	+	+	+	w	—	—	—	+	—	+	+
alaninamide	+	+	+	+	+	—	w	+	+	+	+
L-alanyl-glycine	w	+	—	w	+	—	w	w	—	—	—
L-histidine	—	+	—	w	w	—	+	w	w	+	+
hydroxy-L-proline	—	+	w	w	w	w	+	w	w	+	+
L-leucine	+	+	+	+	+	+	+	+	+	+	—
L-ornithine	—	—	—	—	—	—	—	w	—	+	—
L-phenylalanine	—	—	—	—	—	—	—	+	—	—	—
L-pyroglutamic acid	+	+	+	+	+	—	—	—	w	—	—
D-serine	—	+	—	—	—	—	—	—	w	+	—
L-serine	—	+	—	—	—	—	+	+	w	+	+
L-threonine	+	+	w	—	—	—	—	+	—	+	—
D,L-carnitine	—	—	—	w	—	w	+	w	w	+	+
γ-amino butyric acid	+	+	—	+	+	+	+	w	w	+	+
phenyl ethylamine	—	—	—	—	—	—	—	—	—	+	—
putrescine	w	+	w	w	—	w	+	w	w	+	+
2-amino ethanol	+	—	+	—	+	—	+	+	—	+	—
2,3butanediol	—	—	—	—	—	—	—	—	w	+	—

Table A.3 Antibiotic susceptibility test for strain KC and various *Pseudomonas stutzeri*, *P. balearica* and *P. putida* strains

Antibiotic	*Pseudomonas stutzeri* strains*								*P. balearica**	*P. putida*
	KC	1	2	3	4	5	7	8	6	ATCC† 12633T
ampicillin	MS‡	MS	MS	MS	MS	S‡	MS	MS	MS	R‡
penicillin G	R	R	R	R	R	MS	R	R	R	R
amoxicillin / clavulanic acid	S	S	S	S	S	S	S	S	MS	R
streptomycin	MS	MS	MS	MS	MS	S	MS	MS	MS	R
erythromycin	S	MS	S	S	S	S	MS	MS	MS	R
tetracycline	S	S	S	S	S	S	S	S	MS	MS
trimethroprin / sulfamethoxazole	MS	R	R	R	MS	S	MS	S	R	R

* *Pseudomonas stutzeri* strains tested: CCUG1126 (gv , type strain), ATCC17591 (gv 2), DSM50227 (gv 3), 19SMN4 (gv 4) DNSP21 (gv 5), DSM50238 (gv 7), JM300 (gv 8); *Pseudomonas balearica* DSM6083 (gv 6, type strain)

† ATCC, American Type Culture Collection, Rockville, MD, USA

‡ R, resistant, diameter of inhibition < 1 cm; MS, moderate sensitiveness, diameter of inhibition 1•0 - 1•7 cm; S, sensitive, diameter of inhibition > 1•7 cm

T type strain

The fatty acid distribution of strain KC is very similar in composition and abundance to the profiles observed for *P. stutzeri* and *P. balearica* strains (Table A.4). This result is congruent with previous observations (Rosselló *et al.*, 1994a; Stead, 1992; Veys *et al.*, 1989) indicating that a differentiation of *P. stutzeri* strains on the basis of fatty acid patterns is not possible. Even though the fatty acid profile of *P. putida* strains were similar to the profiles of strain KC and the *P. stutzeri* strains, a significant difference in the abundance of two fatty acids was observed. The abundance of fatty acids 16:0 and 18:1 ω7c were 14 – 18% and 27 – 40%, respectively, for all *P. stutzeri* strains (including

strain KC) and *P. balearica*. The abundance of these two fatty acids was 25% and 15%, respectively, in the case of the two *P. putida* strains tested.

Table A.4 Cellular fatty acid composition of strain KC and several *Pseudomonas stutzeri, P. balearica* and *P. putida* strains

Fatty acid	% of total fatty acids										
		Pseudomonas stutzeri strains*									
	KC	gv 1	gv 2	gv 3	gv 4	gv 5	gv 7	gv 8	gv 6	9	10
10:0	ND†	0•25	0•18	0•28	0•18	0•27	ND	ND	0•29	ND	0•25
10:0 3OH	3•36	3•70	2•42	2•50	3•45	3•65	2•86	3•93	3•46	3•52	3•89
11:0 iso 3OH	ND	ND	0•09	ND	0•10	ND	ND	ND	ND	ND	0•20
12:0	8•12	11•5	7•43	10•25	7•44	11•13	10•26	11•59	10•29	3•24	2•67
12:0 2OH	ND	0•16	0•10	ND	0•11	ND	ND	ND	ND	6•16	7•03
12:0 3OH	3•90	4•13	2•42	2•89	2•99	3•52	3•30	3•31	3•53	4•64	4•82
12:1 3OH	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0•21
13:0 iso	ND	ND	0•11	ND	0•08	ND	ND	ND	ND	ND	ND
14:0	0•79	1•25	0•61	1•18	0•90	1•30	1•06	1•71	0•87	0•44	0•34
15:0	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0•25
15 iso 2OH/ 16:1 ω7c	33•56	32•56	28•33	38•33	32•28	37•33	35•49	34•97	30•44	38•10	31•62
16:0	14•73	14•69	16•84	16•29	17•84	13•95	17•81	15•62	16•61	25•20	26•56
17:0 iso	ND	ND	0•49	0•41	0•43	ND	ND	ND	1•28	ND	0•81
17:0 cyclo	ND	ND	ND	0•35	0•20	ND	ND	0•97	1•00	3•11	4•29
18:1 ω7c	35•28	31•57	39•56	27•23	33•18	28•85	29•21	27•13	30•79	15•59	16•48
18:0	ND	ND	0•55	ND	0•36	ND	ND	0•78	ND	ND	0•37
19:0 cyclo ω8c	ND	ND	0•15	ND	0•14	ND	ND	ND	0•51	ND	ND

* *Pseudomonas stutzeri* strains tested: CCUG 1126 (gv 1, type strain), ATCC 17591 (gv 2), DSM 50227 (gv 3), 19SMN4 (gv 4) DNSP21 (gv 5), DSM 50238 (gv 7), JM300 (gv 8); *Pseudomonas balearica* DSM 6083 (gv 6, type strain); *Pseudomonas putida* ATCC 12633 (9, type strain), *Pseudomonas putida* DSM 3601 (10)

† ND, not detected

DNA-DNA similarity studies. Strain KC did not show similarity values higher than 70% with any of the type strains of the 7 genomovars of *P. stutzeri* described so far (Table A.5). DNA-DNA similarity values are usually higher than 70% for members of the same gv, between 40 and 60% for members of different gv and under 20% when a *P. stutzeri* strain is compared with other *Pseudomonas* species (Rosselló *et al.*, 1991). DNA hybridizations between strain KC and most *P. stutzeri* genomovars were in the 40 to 60% range, as expected for inter-genomovar hybridizations, and low similarity indices were observed when KC was hybridized to *P. balearica*, *P. putida* and *P. aeruginosa* type strains.

Table A.5 DNA – DNA similarity results for strain KC and several *Pseudomonas* strains

Strain		% DNA – DNA similarity
KC		100
P. stutzeri	ATCC* 17589 (gv‡ 1)	42•1
	ATCC 17951 (gv 2)	41•8
	DSM* 50227 (gv 3)	49•6
	19SMN4 (gv 4)	31•7
	DNSP21 (gv 5)	65•9
	DSM 50238 (gv 7)	41•1
	JM300 (gv 8)	18•7
P. balearica	DSM 6083^T (gv 6)	24•9
P. putida	DSM 50202^T	32•7
P. aeruginosa	CCM* 1960^T	25•8
Pooled standard deviation		2•31

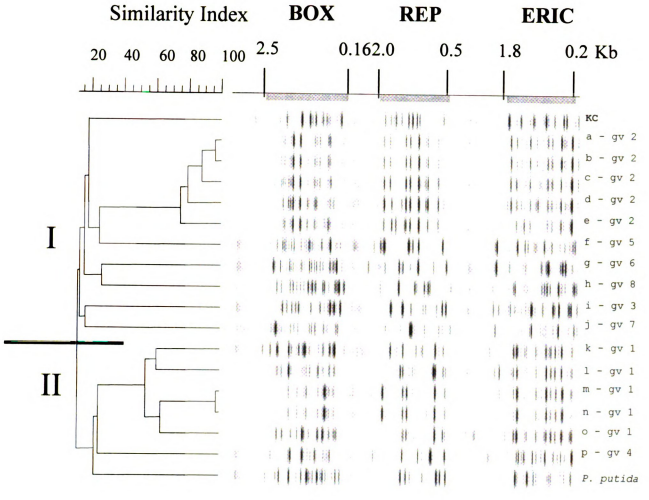
† ATCC, American Type Culture Collection, Rockville, MD, USA; CCM, Czechoslovak Collection of Microorganisms, Brno, Czechoslovakia; DSM, Deutsche Sammlung von Mikroorganismen und Zellkulteren, Braunschweig, Germany

^{T} type strain

DNA fingerprinting studies. When the results of REP, BOX and ERIC fingerprinting were combined to generate a profile, comparisons once again revealed low similarity indices between strains, as seen in Figure A.1. Two major clusters were observed, one containing strains belonging to genomovar 1 and 4 and another cluster with the other six genomovars. Strain KC distantly linked to the second cluster. Higher similarity indices would be expected for strains of the same species. The differences in the results of genomic typing between strains of the eight genomovars confirms the genomic groups within the species *P. stutzeri*, as described previously (Bennasar *et al.*, 1998a). Strain KC cannot be assigned to a previously described genomovar by genomic typing analysis. The low similarity indices observed in the DNA similarity assays may be explained by the large chromosomal plasticity detected in *P. stutzeri* (Ginard *et al.*, 1997) since chromosomal re-arrangements would interfere with experiments that depend on DNA sequence homogeneity. This observation is also consistent with a recent report describing a spontaneous deletion of a 170 kb fragment of chromosomal DNA from strain KC (Lewis *et al.*, 2000).

Figure A.1 Simplified phenogram based on the UPGMA analysis or normalized BOX, REP and ERIC fingerprinting of strain KC and *Pseudomonas stutzeri* strains. The following strains were used for the analysis: a, *P. stutzeri* ATCC 17591; b, *P. stutzeri* ATCC 17587; c, *P. stutzeri* ATCC 17592; d, *P. stutzeri* ATCC 17595; e, *P. stutzeri* ZoBell; f, *P. stutzeri* DNSP21; g, *P. balearic*a DSM 6083 (type); h, *P. stutzeri* JM300; i, *P. stutzeri* DSM 50227; j, *P. stutzeri* DSM 50238; k, *P. stutzeri* CCUG 11256 (type); l, *P. stutzeri* ATCC 17589; m, *P. stutzeri* ATCC 17593; n, *P. stutzeri* ATCC 17594; o, *P. stutzeri* ATCC 27951; *P. stutzeri* 19SMN4; *P. putida* ATCC 12633 (type).

Strain ATCC 27951 was deposited to the ATCC as *Flavobacterium lutescens* and was later re-classified as *P. stutzeri* (Bennasar *et al.*, 1998b).


Similarity Index
BOX
REP
ERIC
20 40 60 80 100
2.5
0.16
2.0
0.5
1.8
0.2 Kb
I
II
KC
a - gv 2
b - gv 2
c - gv 2
d - gv 2
e - gv 2
f - gv 5
g - gv 6
h - gv 8
i - gv 3
j - gv 7
k - gv 1
l - gv 1
m - gv 1
n - gv 1
o - gv 1
p - gv 4
P. putida

16S rDNA, *gyr*B PCR and ITS1 region sequencing. 16S rDNA sequence comparisons with sequences available in the Ribosomal Database Project demonstrate that the 16S rDNA gene of strain KC clustered with *P. stutzeri* in the Neighbor-Joining analysis, with a bootstrap confidence of 86% (Figure A.2). When the sequence was compared to 16S rDNA genes from strains assigned to different *P. stutzeri* genomovars, strain KC clustered within the *P. stutzeri* phylogenetic branch, being gv 3 the closest group with 99•38%, followed by gv 4 with a 99•24% sequence identity. The 16S rDNA of strain KC can be amplified using the *P. stutzeri* specific primers fps158 and rps1271 (Bennasar *et al.*, 1998a). The 16S rDNA of some *P. putida* strains can also be amplified with these primers but the 1,160 bp PCR product cannot be restricted by *Bam*HI to yield a 645 bp fragment and 465 bp fragment (Bennasar *et al.*, 1998a). As seen on Figure A.3, only *P. stutzeri* strains and strain KC had PCR products that could be digested with *Bam*HI. No PCR products were obtained from strain KC with the *gyr*B primers specific to *P. putida*, indicating that strain KC is not likely to be a *P. putida* strain.

Figure A.2 Bootstrap parsimony tree obtained when the 16S rDNA gene of strain KC is compared to some sequences available in the Ribosomal Database Project. The percentages indicate the bootstrap confidence estimates on the branches obtained with 100 replicates.

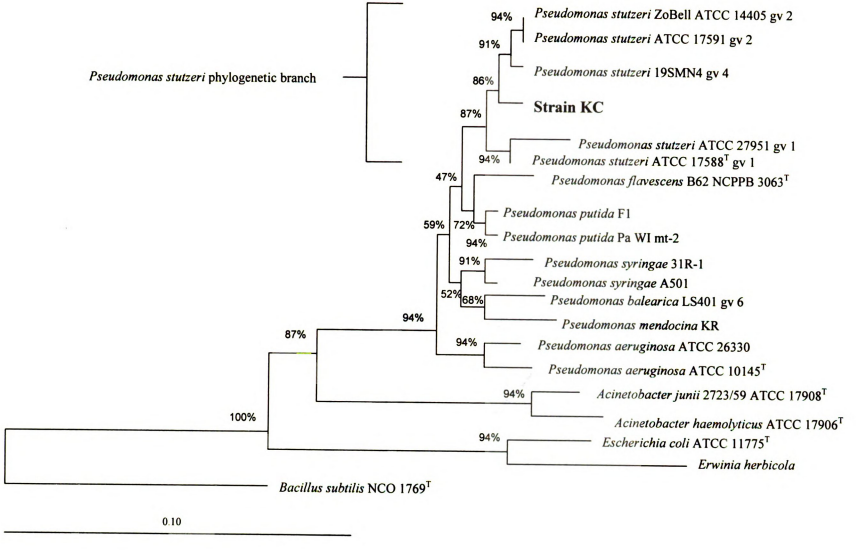
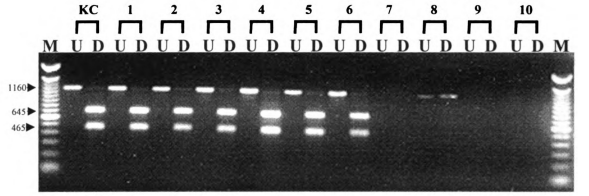

Pseudomonas stutzeri phylogenetic branch
94%
Pseudomonas stutzeri ZoBell ATCC 14405 gv 2
91%
Pseudomonas stutzeri ATCC 17591 gv 2
Pseudomonas stutzeri 19SMN4 gv 4
86%
Strain KC
87%
Pseudomonas stutzeri ATCC 27951 gv 1
94%
Pseudomonas stutzeri ATCC 17588T gv 1
47%
Pseudomonas flavescens B62 NCPPB 3063T
Pseudomonas putida F1
59%
72%
Pseudomonas putida Pa WI mt-2
94%
91%
Pseudomonas syringae 31R-1
Pseudomonas syringae A501
52%
68%
Pseudomonas balearica LS401 gv 6
94%
Pseudomonas mendocina KR
87%
94%
Pseudomonas aeruginosa ATCC 26330
Pseudomonas aeruginosa ATCC 10145T
94%
Acinetobacter junii 2723/59 ATCC 17908T
100%
Acinetobacter haemolyticus ATCC 17906T
94%
Escherichia coli ATCC 11775T
Erwinia herbicola
Bacillus subtilis NCO 1769T
0.10

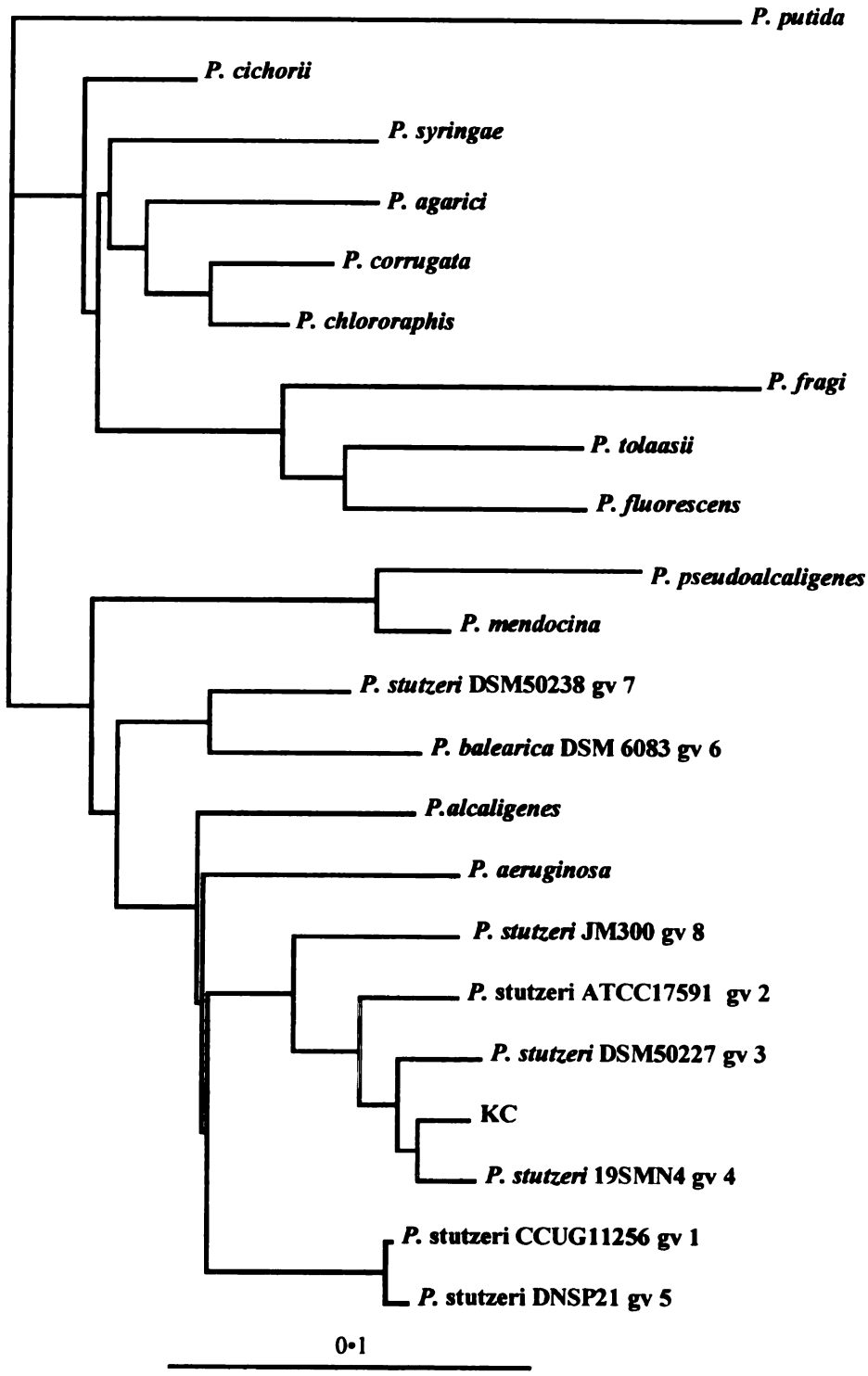
Figure A.3 *Bam*HI restriction digests of the 16S rRNA gene amplified by PCR using the *Pseudomonas stutzeri*- specific primers fps158 and rps1271.



KC, *Pseudomonas* sp. strain KC; M, 100 bp DNA ladder; U, undigested sample; D, *Bam*HI restricted sample; 1160, length of undigested fragment (bp); 465 and 695, length of restricted fragments (bp); 1, *P. stutzeri* CCUG 11256 (type, gv 1); 2, *P. stuzeri* ATCC 17591 (gv 2); 3, *P. stutzeri* DSM 30227 (gv 3); 4, *P. stutzeri* 19SMN4 (gv 4); 5, *P. stutzeri* DNSP21 (gv 5); 6, *P. stutzeri* DSM50238 (gv 7); 7, *P. balearica* DSM 6083 (type, gv 6); 8, *P. putida* ATCC 12633 (type); 9, *P. aeruginosa* ATCC 10145 (type); 10, *E. coli* K-12 ATCC 10798

ITS1 sequence analysis results agree with 16S rDNA sequence comparisons. Once again, strain KC clustered with the *P. stutzeri* phylogenetic group, as seen in Figure A.4. ITS1 sequences are identical within all the strains of a *P. stutzeri* genomovar, and deletions or insertions in this portion of DNA can be used as a taxonomic tool to differentiate strains at the gv level (Guasp *et al.*, 2000). Strain KC is closely linked to the gv 3 – gv 4 cluster. When the ITS1 sequences of strain KC and strain 19SMN4 (gv 4) are compared, differences are only observed towards the end of the sequence. The ITS1 region of strain KC has an 11 bp and a 2 bp insertion separated by 4 bases. Nine mismatches are also observed in the vicinity of these insertions. The sequence differences observed in the ITS1 region indicate that strain KC could be a new genomovar of *P. stutzeri*.

Figure **A.4** Dendrogram depicting phylogenetic relationships among strain KC, several *P. stutzeri* strains and type strains of other *Pseudomonas* species, as estimated by comparing the ITS1 sequence. Scale bar 0•1 nucleotide substitutions per nucleotide position. The EMBL accession numbers of the sequences used for this analysis are provided in parenthesis: *P. putida* (need #), *P. cichorii* (AJ279242), *P. syringae* (D86356), *P. agarici* (AJ279243), *P. corrugata* (need #), *P. chlororaphis* (AJ279240), *P. fragi* (AJ279241), *P. tolaasii* (AJ279244), *P. fluorescens* (need #), *P. pseudoalcaligenes* (AJ27945), *P. mendocina* (L28159), *P. stutzeri* DSM 50238 (AJ51909), *P. alcaligenes* (need #), *P. aeruginosa* (L28148), *P. stutzeri* JM300 (AJ390581), *P. stutzeri* ATCC 17591 (AJ251901), *P. stutzeri* DSM 50227 (AJ251903), *Pseudomonas* sp. strain KC (to be submitted), *P. stutzeri* 19SMN4 (AJ241906), *P. stutzeri* CCUG 11256 (AJ251910), *P. stutzeri* DNSP21 (AJ21908).


P. putida
P. cichorii
P. syringae
P. agarici
P. corrugata
P. chlororaphis
P. fragi
P. tolaasii
P. fluorescens
P. pseudoalcaligenes
P. mendocina
P. stutzeri DSM50238 gv 7
P. balearica DSM 6083 gv 6
P.alcaligenes
P. aeruginosa
P. stutzeri JM300 gv 8
P. stutzeri ATCC17591 gv 2
P. stutzeri DSM50227 gv 3
KC
P. stutzeri 19SMN4 gv 4
P. stutzeri CCUG11256 gv 1
P. stutzeri DNSP21 gv 5
0•1

DISCUSSION

Based on the overall evidence from phylogenetic and genetic studies, we conclude that strain KC should be classified as a member of a new genomovar within the species *P. stutzeri*. *Pseudomonas stutzeri* is a motile, non-fluorescent, denitrifying, Gram-negative, rod-shaped bacterium that is widely distributed in nature. It was first described in 1895 (Burri & Stutzer, 1895) as *Bacillus denitrificans* II and re-classified fifty-seven years later as *P. stutzeri* (Van Niel & Allen, 1952). Members of the newly described species were distinguishable from other non-fluorescent pseudomonas by the dry, wrinkled colonies of fresh isolates and the ability to use maltose and starch as sole carbon sources, as well as the ability to produce large amounts of molecular nitrogen from nitrate. Many strains of this species are of special interest for their ability to degrade environmental pollutants (Baggi *et al.*, 1987; Roselló-Mora *et al.*, 1994b). It has not been possible to assign each of the *P. stutzeri* genomovars to new species due to their highly diverse phenotypes which do not allow the identification of phenotypic traits unique to each genomovar (Rosselló *et al.*, 1991; Rosselló *et al.*, 1994a). To date, only one of the *P. stutzeri* genomovars has been re-classified as a new species based on distinct phenotypic characteristics (Bennasar *et al.*, 1996). Strain KC has the unique ability to produce PDTC enabling carbon tetrachloride degradation, and it can grow on m-inositol, unlike any *P. stutzeri* strain studied in this work. On the other hand, other phenotypic characteristics like resistance to antibiotics and the composition of cellular fatty acids, correlate with the patterns observed in other *P. stutzeri* strains.

DNA similarity, as measured by DNA-DNA re-association studies (Johnson, 1973; Johnson & Palleroni, 1989) has been the standard genotypic method for assigning a strain to a given genomovar (Rosselló *et al.*, 1991). When strain KC is compared to *P. stutzeri* strains, the similarity indices are below the 70% threshold; the value used classically to differentiate between species and members of the same genomovar. On the other hand, the similarity indices remained between the 40 to 60% observed when members of different genomovars are compared (Rosselló *et al.*, 1991). It should be noted that unlike many species that show homogeneity at the genotypic levels, *P. stutzeri* have proven to be highly diverse. Comparison of genomic maps for different *P. stutzeri* strains revealed a high degree of genomic plasticity as chromosomal re-arrangements can occur without apparent consequences in strain fitness (Ginard *et al.*, 1997). This genomic plasticity may have played an important role in *P. stutzeri*'s ability to colonize and persist in diverse environments. Of interest in this regard is the recent report of Lewis *et al.* (2000) indicating that strain KC can spontaneously lose large chromosomal fragments (~ 170 kb) without loss of viability. This trait is consistent with the characteristics of the *P. stutzeri* species.

When smaller pieces of DNA with phylogenetic relevance like 16S rDNA, ITS1 and *gyrB* are analyzed to deduce phylogenetic relationships, strain KC clusters in the *P. stutzeri* phylogenetic branch. Strain KC exhibited over a 99% similarity index with the 16S rDNa gene of members of gv 3 and 4. 16S rRNA gene sequence comparisons support the natural relationship among the genomovars and have further sustained the genomovar concept because similarities of 16S rRNA genes is 99•9 to 100% for

members of the same genomovar and 98•0 to 99•7% for members of different genomovars (Bennasar *et al.*, 1996). Strain KC shows more than 98% 16S rDNA sequence similarity with the type strains of *P. stutzeri* genomovars and similarities of 96%, with *P. putida*, *P. aeruginosa* and *P. balearica*. These results are consistent with previous observations reporting 16S rDNA similarity indices of less than 97% for strains of different species (Stackebrant & Goebel, 1994). PCR amplification of *gyr*B also excluded the possibility that strain KC belongs to the species *P.* putida because the *gyr*B gene of strain KC did not amplify with *P. putida*-specific primers.

The intragenic, 16S-23S internally-transcribed spacer region (ITS1) has also been used as a tool to confirm genomovar assignments (Guasp *et al.*, 2000). The sequence of ITS1 is assumed to be less susceptible to selective pressures, due to its non-coding function, and should have accumulated a higher percentage of mutations than the rRNA genes (Tyrrell *et al.*, 1997). Comparison of ITS1 sequences indicate that the considerable variation in length and sequence make these regions good candidates for discriminating among closely related taxa (Gürtler & Stanisich, 1996). Strain KC clustered with *P. stutzeri* strains (Figure A.4) showing more than 80% identity at the sequence level with *P. stutzeri* strains and less than 70% sequence identity with other closely related species such as *P.putida*, *P. aeruginosa* and *P. mendocina*.

The results reported in this publication demonstrate that strain KC is a member of the *P. stutzeri* species. Its phenotype fits the description of the overall phenotype of the species, except for the ability to grow on m-inositol and the capability to degrade carbon

tetrachloride. We therefore propose that strain KC be classified the sole representative of a new genomovar, genomovar 9, following the enumeration of Rosselló *et al.* (1991, 1996). The isolation and characterization of new strains belonging to the same gv as strain KC may help to clarify if the unique phenotypic characteristics of strain KC are sufficient to propose its reclassification as a new species within the genus *Pseudomonas*; or if these attributes simply reflect the unusual physiological traits within the diverse *P. sutzeri* species.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Mrs. Carmen M. Medina-Ferrer for technical assistance in DNA fingerprinting analysis. This work was supported, in part, by grants from the National Science Foundation Center for Microbial Ecology (BIR-9120006) and by the NIEHS Superfund Basic Research Program of the Institute for Environmental Toxicology (ES04911) at Michigan State University.

REFERENCES

Baggi, G., Barbieri, P., Galli, E. & Tollari, S. (1987). Isolation of a *Pseudomonas stutzeri* strain that degrades o-xylanes. *Appl Environ Microbiol* **53,** 2129-2132.

Bennasar, A., Rosselló-Mora, R. A., Lalucat, J. & Moore, E. R. (1996). 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stutzeri* and proposal of *Pseudomonas balearica* sp. nov. *Int J Syst Bacteriol* **46,** 200-205.

Bennasar, A., Guasp, C., Tesar, M. & Lalucat, J. (1998a). Genetic relationships among *Pseudomonas stutzeri* strains based on molecular typing methods. *J Appl Microbiol* **85,** 643-656.

Bennasar, A., Guasp, C. & Lalucat, J. (1998b). Molecular methods for the detection and identification of *Pseudomonas stutzeri* in pure culture and environmental samples. *Microb Ecol* **35,** 22-33.

Brosius, J., Palmer, M. L., Kennedy, P. J. & Noller, H. F. (1978). Complete nucloetide sequence of a 16S ribosomal RNA gene form *Escherichia coli. Proc Natl Acad Sci USA* **75,** 4801-4805.

Brosius, J., Dull, T. J. & Noller, H. F. (1980). Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli. Proc Natl Acad Sci USA* **77,** 201-204.

Burri, R. & Stutzer, A. (1895). Ueber Nitrat zerstörende Bakterien und den durch dieselben bedingten Stickstoffverlust. *Zbl Bakt* **II,** 257-265, 350-364, 392-398, 422-432.

Carlson, C. & Ingraham, J. (1983). Comparison of denitrification by *Pseudomonas stutzeri*, *Pseudomonas aeruginosa* and *Paracoccus denitrificans*. *Appl Environ Microbiol* **45,** 1247-1253.

Criddle, C. S., DeWitt, J. T., Grbic-Galic, D. & McCarthy, P. L. (1990). Transfromation of carbon tetrachloride by *Pseudomonas* sp. strain KC under denitrification conditions. *Appl Environ Microbiol* **56,** 3240-3246.

Devereaux, J., Haeberli, P. & Smithies, O. (1984). A comparative set of sequence analysis programs for the VAX. *Nucleic Acid Res* **12,** 387-395.

Dybas, M. J., Tatara, G. M. & Criddle, C. S. (1995). Localization and chracterization of the carbon tetrachloride transformation activity of *Pseudomonas* sp. strain KC. *Appl Envrion Microbiol* **61,** 758-762.

Emerson, D. (1999). Complex pattern formation by *Pseudomonas* strain KC in response to nitrate and nitrite. *Microbiology* **145,** 633-641.

Felsenstein, J. (1989). PHYLIP-phylogeny inference package (version 3.2). *Cladistics* **5,** 164-166.

Ginard, M., Lalucat, J., Tümmler, B. & Römling, U. (1997). Genome organization of *Pseudomonas stutzeri* and resulting taxonomic and evolutionary considerations. **Int J Syst Bacteriol 47,** 132-143.

Guasp, C., Moore, E. R. B., Lalucat, J. & Bennasar, A. (2000). Utility of internally-transcribed 16S-23S rDNA spacer regions for the definition of *Pseudomonas stutzeri* genomovars and other *Pseudomonas* species. *Int J Sys Bacteriol* **50,** (part4).

Gürtler, V. & Stanisich, V. A. (1996). New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology* **142,** 3-16.

Hyndman, D. W., Dybas, M. J., Forney, L., Heine, R., Mayotte, T., Phanikumar, M. S., Tatara, G., Tiedje, J., Voice, T., Wallace, R., Wiggert, D., Zhao, X. & Criddle, C. S. (2000). Hydraulic Characterization and design of a full scale biocurtain. *Ground Water* **In Press.**

Jensen, M. A., Webster, J. A. & Straus, N. (1993). Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl Environ Microbiol* **59,** 945-952.

Johnson, J. L. (1973). Use of nucleic acid homologies in the taxonomy of anaerobic bacteria. *Int J Syst Bacteriol* **23,** 308-315.

Johnson, J. L. & Palleroni, N. J. (1989). Deoxiribonucleic acid similarities among *Pseudomonas* species. *Int J Sys Bacteriol* **39,** 230-235.

Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mammalian protein metabolism* (ed. H. N. Munro), pp. 21-132. Academic Press, New York.

Kogan, S., Doherty, M. & Gitschier, J. (1987). An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. *N Engl J Med* **317,** 985-990.

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L. & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* **82,** 6955-6959.

Lee, C.-H., Lewis, T. A., Paszczynski, A. & Crawford, R. L. (1999). Identification of an extracellular catalyst of carbon tetrachloride dehalogenation from

Pseudomonas stutzeri strain KC as pyridine-2,6-bis(thiocarboxylate). *Biochem Biophys Res Commun* **261,** 562-566.

Lewis, T. A. & Crawford, R. L. (1993). Physiological factors affecting carbon tetrachloride dehalogenation by the denitrifying bacterium *Pseudomonas* sp. strain KC. *Appl Environ Microbiol* **59,** 1635-1641.

Lewis, T. A., Cortese, M. S., Sebat, J. L., Green, T. L. & Crawford, R. L. (2000). A *Pseudomonas stutzeri* gene cluster encoding the biosynthesis of the CCl_4 - dechlorinating agent pyridine-2,6-bis(thiocarboxilic acid). *Environ Microbiol* **In Press.**

Maidak, B. L., Cole, J. R., Parker, C. T., Garrity, G. M., Larsen, N., Li, B., Lilburn, T. G., McCaughey, M. J., Olsen, G. J., Overbeek, R., Pramanik, S., Schmidt, T. M., Tiedje, J. M. & Woese, C. R. (1999). A new version of RDP (Ribosomal Database Project). *Nucleic Acids Res* **27,** 171-173.

Marmur, J. (1961). A procedure for the isolation of DNA from microorganisms. *J Mol Biol* **3,** 208-218.

Ockels, W., Römer, A. & Budzeikiewicz, H. (1978). An Fe(II) complex of pyridine-2,6-di-(monothiocarboxilic acid) - a novel bacterial metabolic product. *Tetrahedron Lett* **36,** 3341-3342.

Palleroni, N., Doudoroff, M., Stanier, R., Solanes, R. & Mandel, M. (1970). Taxonomy of the aerobic pseudomonas: the properties of the *Pseudomonas stutzeri* group. *J Gen Microbiol* **60,** 215-31.

Rabaut, A. (1996). *Se-Al: sequence alignment editor. Version 1.0 alpha 1. Department of Zoology. University of Oxford. South Parks Roads. Oxford OX1 4JD. U.K.*

Rosselló, R. A., García-Valdés, E., Lalucat, J. & Ursing, J. (1991). Genotypic and Phenotyic Diversity of *Pseudomonas stutzeri*. *Syst Appl Microbiol* **14,** 150-157.

Rosselló, R. A., Lalucat, J., Dott, W. & Kämpfer, P. (1994a). Biochemical and chemotaxomic characterization of *Pseudomonas stutzeri* genomovars. *J Appl Bacteriol* **76,** 226-233.

Roselló-Mora, R. A., Lalucat, J. & García-Valdés, E. (1994b). Comparative biochemical and genetic analysis of naphthalene degradation among *Pseudomonas stutzeri* strains. *Appl Environ Microbiol* **60,** 966-972.

Rosselló-Mora, R. A., Lalucat, J., Timmis, K. N. & Moore, E. R. B. (1996). Strain JM300 represents a new genomovar within *Pseudomonas stutzeri*. *Syst Appl Microbiol* **19,** 596-599.

Sambrook, J., Fritsch, E. L. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sasser, M. (1990a). Technical note # 101: Identification of bacteria by gas chromatography of cellular fatty acids. *MIDI, Inc., Newark, DE* .

Sasser, M. (1990b). Technical note # 102: "Tracking" a strain using the Sherlock Microbial Identification System. *MIDI, Inc., Newark, DE* .

Sepúlveda-Torres, L. del C., Rajendran, N., Dybas, M. J. & Criddle, C. S. (1999). Generation and initial characterization of *Pseudomonas stutzeri* KC mutants with Impaired ability to degrade carbon tetrachloride. *Archiv Microbiol* **171,** 424-429.

Skerman, V., McGowan, V. & Sneath, P. (1980). Approved list of bacterial names. *Int J Syst Bacteriol* **30,** 225-420.

Stackebrant, E. & Goebel, B. (1994). Taxonimic note: a place for DNA-DNA reassociation and 16S sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44,** 846-849.

Stanier, R., Palleroni, N. & Doudoroff, M. (1966). The aerobic pseudomonas: a taxonomic study. *J Gen Microbiol* **43,** 159-271.

Stead, D. (1992). Grouping of plant-pathogenic and some other *Pseudomonas* spp. by using cellular fatty acids profiles. *Int J Sys Bacteriol* **42,** 281-295.

Stunk, O., Gross, O., Reichel, B., May, M., Hermann, S., Struckmann, N., Nonhoff, B., Lenke, M., Vilbig, A., Ludwig, T., Bode, A., Schleifer, K. H. & Luwig, W. (2000). ARB: a sofware environment for sequence data. *Nucleic Acids Res* **In Press.**

Tatara, G. M., Dybas, M. J. & Criddle, C. S. (1995). Effect of medium and trace metals on kinetics of carbon tetrachloride transformation by Pseudomonas stutzeri KC. *Appl Environ Microbiol* **59,** 2126-2131.

Thompson, J. D., Higgins, D. G. & Gibson, T. G. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* **22,** 4673-4680.

Tyrrell, G. J., Bethune, R. N., Willey, B. & Low, D. E. (1997). Species identification of enterococci via intergenic ribosomal PCR. *J Clin Microbiol* 35, 1054-1060.

Van Niel, C. & Allen, M. (1952). A note on *Pseudomonas stutzeri. J Bacteriol* **64,** 413-422.

Vauterin, L. & Vauterin, P. (1992). Computer aided comparison of electrophoresis patterns for grouping and identification of microorganisms. *European Microbiology* **1,** 37-41.

Versalovic, J., Koeuth, T. & Lupski, J. (1991). Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nuclei Acids Res* **19,** 6823-6831.

Versalovic, J., Schneider, M., de Bruijn, F. J. & Lupski, J. R. (1994). Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Molec Cell Biol* **5,** 25-40.

Veys, A., Callewaert, W., Waelkens, E. & Van den Abbeele, K. (1989). Application of gas-liquid chromatography to the routine identification of nonfermenting gram-negative bacteria in clinical specimens. *J Clin Microbiol* **27,** 1538-1542.

Weisburg, W. W., Barns, S. M., Palletier, D. A. & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173,** 697-703.

Witt, M. E., Dybas, M. J., Worden, R. M. & Criddle, C. S. (1999a). Motiltity-enhanced bioremediation of carbon tetrachloride-contaminated aquifer sediments. *Environ Sci Technol* **33,** 2958-2964.

Witt, M. E., Dybas, M. J., C, W. D. & Criddle, C. S. (1999b). Use of bioaugmentation for continuous removal of carbon tetrachloride in model aquifer columns. *J Environ Eng Sci* 16, 475-485.

Yamamoto, S. & Harayama, S. (1995). PCR amplification and direct sequencing of *gyr*B genes with universal primers and their application to the detection and

taxonomic analysis of *Pseudomonas putida* strains. *Appl Environ Microbiol* 61, 1104-1109.

Zhou, J., Fries, M. R., Chee-Sanford, J. C. & Tiedje, J. M. (1995). Phylogenetic analyses of a new group of denitrifiers capable of anaerobic growth on toluene and description of *Azoarcus tolulyticus* sp. nov. *Int J Syst Bacteriol* **45,** 500-506.

Ziemke, F., Höfle, M. G., Lalucat, J. & Rosselló-Mora, R. (1998). Reclassification of *Shewanella putrefaciens* Owen's genomic group II as *Shewanella baltica* sp. nov. *Int J Syst Bacteriol* **48,** 179-86.

APPENDIX B

SEQUENCE OF A 8,274 BASE PAIR *ECO*RI FRAGMENT MUTATED IN FOUR *PSEUDOMONAS STUTZERI* STRAIN KC TRANSPOSITIONAL MUTANTS WITH IMPAIRED ABILITY TO DEGRADE CARBON TETRACHLORIDE

RBS 437 - 439
CDS 446 - 1012
Translation

MSYRPGTARPRGTGRLPTKPNPVETLPFPSLLARPHALQSCWPT
QLTEPLRDGRPCPVSPAPAGRRQGWADTARVKMAQVQAVRADVS
CEILPRQPVFSPIGRTLNMDGMPGLVAVALAKPKLEVSGVGFEC
PAAAVVAQQPIESVGLPDRLSARSGDLVQEVEQRDDVCFPFASV
AGLLARKEENLS

RBS 994 - 999
CDS 1009 - 2262
miscelaneous 12 transmembrane helices indicated by underlined residues
Translation

MTKSRVVGLQLLFGWMNLVLAVPSIYLMLGMPLVMRQHGWSGAE
IGLFQLAALPAIFKFLLAVPVQRVRLGRGHFVHWLLLLCALLLA
LYWLIGRHNLIGDRIMLFALTFAISIAATWADIPLNALAVQWLP
RSEQLRAGSIRSAALFVGAIVGGGVMIMVQARVGWQAPFWLLGV
GLLIGALPFLLLRRHAALPEQAEPRETTDPPPGVMADWASFFHQ
PGARQWTLLLLTSFPFLGATWLYLKPLLLDMGMQLERVAFIVGI
VGGTAGALFSLLGGQLVQMLGIARAIAWYLLAALGALALLTFSV
WAQLGAAWLIASALCVAASMGAISALMFGLTMFFTRNRRNASDY
ALQTTMFTVARLAVPIAAGVLLDRVGYTGMLLAMTLALLLSFAL
ACRVREKVESSAQSILEHERV

Miscellaneous Rho-independent terminator 2295 - 2331

-35 seq 2346 - 2351
-10 seq 2372 - 2376
Fur box 2391 - 2409
RBS 2423 - 2426
CDS 2435 - 3610
Miscellaneous 2 transmembrane helices indicated by single underlines, sequence similar to chemotaxis transducer signature indicated by double underlines
Translation

MPLSALVAPAGELSRAEINRYSRHLLIPDVGMIGQRRLKNAKVL
VIGAGGLGSPTLLYLAAAGVGTIGIIDFDRVDDSNLQRQVIHGV
DTVGELKVDSAKKAIARLNPFVQVETYTDRLERDMAIELFSRYD
LIMDGTDNFATRYLVNDACVLANKPYVWGSIFRFEGQASVFWEN
APNDLGLNYRDLYPEPPPPEMAPSCSEGGVFGILCASIASIMAT
EAVKLITGIGEPLLGRLVVYDALDMRYRELPVRRLPNRQPITDL
AEDYQVFCGLGLPKGDTADAVPGISVTELKKRMDQDEVPVLIDV
REPTEWDIVRIPGAILVTKSPTAAQTLRERYGADANLVIVCKSG
RRSADVTAELLNLGMRNVRNLEGGVLAWVKDVDSSLPSY

RBS 3612 - 3615
CDS 3626 - 4036
Translation

MALLIKRQALGQVLAQARRDHPLETCGIVASSLEAQLATRVIPM
RNQAASQTFFRLDSQEQFQVFRSLDDRNEFQRVIYHSHTASEAY
PSREDIEYAGYPEAHHLIVSTWENAREPARCFRILRGKVIEESI
SIVE

RBS 4088 - 4091
CDS 4099 - 4371
Miscelaneous perfect TonB dependent recognition sequence indicated by a single underline, glycine-glycine C-terminus domain indicated by a double underline
Translation

MSISVIVPTLLRPLTNGEKTVFTQGNSVAEAIENLEHQFPGLKA
RLVSAEHVHRFVNIYVNEDDIRFSDGLNTPLKAGDSLTVLPAVA
GG

RBS 4450 - 4451
CDS 4460 - 6291
Translation

MPTLLNEFSLLHSSTSFPPNWNELQLSLTEQARLLGICPLAISP
PVDMEGAAFQLQHPAISPIQAHFASPAGWLPNRHLSELLLQAGS
GLMSVHGRASGRAQPLGVDYLSTLTAVMTLHGTLAAAVGQLRGG
AFDQVQLSPLGCGLLSIGQYLAGATAPEDREAFLPGGSDPHLRP
PFRSADGITFELETLDSTPWRSFWTAVGIESELAGTAWKGFLLR
YARAVSPLPAACLTALARLRYAKIQQLAAQAGVAVVPVRTDAQR
REDPDYRQSLATPWQFESFPPSPERHRDTAFPSLLPLQGMRVIE
SCRRIQGPLAGHLLASLGAEVIRLEPPGGDPLRAMPPCAEGCSV
RFDALNHLKSVHEVDIKSAHGRQLVYELARDADVFLHNWAPGKA
HEMQLDAEHLRRVQPHLVYAYAGGWGRAPVNAPGTDFTVQAWSG
VSAAIARQSGIRGGSLFTVLDVLGGAIAALGVTAALLNRAVTGT
GTYVESSLLGAADLLMHSSGKASRGILSGVYPTLSGLIAIDCQH
PDQFQSLAMLLDIPATADTCQQTLAERLRKRPASEWETVLNERG
IGACVVIEDLKQLAADTRISECLTRKSYFSVNAPWRFL

RBS 6279 - 6283
CDS 6289 - 7983
Miscelaneous AMP-binding domain indicated by underlined residues

Translation

MNNAGIIDLVPAEERQRWVQDGTYPNQPVFTLFAAKAEAHPDKK
AVLSPQGDVTYGELLDAALRMAHSLRDSGIVAGDVVAYQLTNHW
LCCAIDLAVAALGAIVAPFPPGRGKLDIQSLVRRCDARAVIVPQ
AYEGIDLCEVIESLRPTLLSMRRLIVQGKPREGWITLDELMSTE
PLDLASLPRVCPNSPVRLLVSSGTESEPKLVAYSHNALVGGRGR
FLQRIASDGEDFRGMYLVPLGSSFGSTATFGVLCWLGGSLVVLP
KFDVDEAIKAIAAFRPGFILGVPTMLQRIAAQPALESIDKSSLR
GLIVGGSVIDEATVRKCRDAFGCGFISLYGSADGVNCHNTLDDP
IEVVLTSVGKPNPAVCAIRLVDDEGREVRQGEVGEITARGPLTP
MQYVNAPELDERYRDPQGWVKTGDLGYINDKGYLVLAGRKKDVI
IRGGANISPTQIEGLVMAHPDVVTVACIPVPDDDLGQRVCLSVT
LREGAAKFSLKAITDFLRELGLEVNKLPEYLRFYRALPLTPAGK
IDKKALTEEARELGTSGICPAGPGQSTPERSLREYA

RBS 7970 - 7974
CDS 7985 > 8274
Translatation

MRGQPMMMATALICAFVPGPQLAFAAPGSAASPDSTTLPEITVT
AEKIERPLERVPASVAVIDGWDAEQSGITSLKQLEGRIPGLSFQ
PFGQAGMN...

```
1    GAATTCCAGG CTGGTGACAT CCTGCCCGTC GTAGCAGGAG CGGCCTTTCA
51   TCCCAACGGT GATGACCATG ACCCGACCTT CATGCGGTCC GCTCGTTTCC
101  TCGATCAGCG GCACAGTCGG ACGATAGTGC AGCCTGCCAA GACCGAGGCC
151  TGGTTCAAGT TCGACAAAAT CCTGGAAGCA TTGGCCAATC TCGTCGGGCA
201  GCTGCAGCCG AGTCTTGCCA TCGCTCAGGA CGCTGCGGAT GCGTGGTTCT
251  TCCATTGTCA GCCGGTAGCA GCTCTTGTGT GCGCTCACCT CTCACCTCCA
301  GTCCGTTAGC CCCCCGAATG TCGCGCATAT CAATCTGTGT TTCTCGCATA
351  CAAAAAATGC TAATGGTTCT CATTACTATA ATGCATCATT TTACCGATGG
401  CCAGTCATTG GCGACCAAAC ACCCAACTAA AAATACGGAA AATCAATGTC
451  TTACAGGCCA GGCACTGCAC GCCCGAGAGG AACTGGACGA CTACCCACGA
501  AGCCAAACCC TGTCGAAACA CTGCCGTTCC CTTCTCTCCT TGCTCGACCG
551  CACGCGCTGC AATCGTGTTG GCCGACGCAG CTGACCGAGC CGCTGCGCGA
601  TGGCAGGCCG TGTCCGGTCT CACCCGCTCC TGCTGGACGG CGGCAGGGAT
651  GGGCCGATAC CGCCCGGGTG AAGATGGCTC AGGTGCAGGC GGTCAGGGCC
701  GACGTGTCCT GTGAGATTCT TCCTCGACAG CCCGTGTTTT CGCCGATAGG
751  ACGGACACTG AACATGGACG GCATGCCCGG CTTGGTGGCG GTCGCACTGG
801  CGAAGCCGAA GCTGGAAGTG TCCGGTGTAG GGTTCGAGTG CCCGGCCGCC
851  GCCGTCGTGG CGCAGCAGCC CATCGAAAGC GTCGGACTTC CCGACCGTCT
901  GAGTGCGCGC AGCGGCGATC TGGTACAGGA GGTCGAGCAA CGGGATGACG
951  TCTGTTTTCC GTTCGCTTCC GTCGCTGGCC TGCTCGCCCG CAAAGAGGAA
1001 AACCTCTCAT GACCAAGTCG CGAGTGGTAG GTCTGCAACT GTTGTTCGGC
1051 TGGATGAATC TGGTGCTGGC GGTACCCAGC ATCTACTTGA TGCTCGGCAT
1101 GCCACTTGTA ATGCGCCAGC ATGGCTGGAG CGGCGCAGAG ATCGGGTTGT
1151 TCCAGCTTGC CGCGCTGCCG GCGATATTCA AATTCCTGTT GGCTGTGCCG
```

1201 GTGCAGCGTG TGCGCCTCGG GCGCGGACAT TTCGTGCACT GGTTGCTGTT
1251 GCTCTGTGCG CTACTACTGG CGCTGTACTG GCTAATCGGA CGGCATAATC
1301 TGATCGGCGA TCGCATAATG CTGTTCGCGC TGACCTTCGC CATCAGCATT
1351 GCCGCCACGT GGGCCGACAT TCCGCTAAAT GCGCTAGCGG TGCAGTGGTT
1401 GCCGCGTAGT GAACAGTTGC GCGCCGGCAG CATCCGTTCC GCAGCGCTGT
1451 TCGTAGGCGC CATTGTTGGC GGCGGCGTCA TGATCATGGT GCAGGCGCGC
1501 GTGGGCTGGC AGGCCCCCTT CTGGCTGCTA GGGGTCGGAC TGCTGATTGG
1551 CGCCCTGCCC TTCCTGCTGT TGCGTAGACA CGCCGCACTG CCCGAGCAGG
1601 CCGAGCCGCG CGAGACTACA GATCCTCCAC CGGGCGTGAT GGCGGACTGG
1651 GCAAGCTTCT TCCACCAGCC AGGGGCGCGG CAATGGACAT TGCTGCTGCT
1701 GACCAGTTTC CCCTTCCTCG GCGCGACGTG GCTGTACCTC AAACCTTTAT
1751 TGTTGGACAT GGGCATGCAG CTAGAGCGCG TGGCCTTCAT CGTTGGCATC
1801 GTCGGCGGCA CCGCAGGCGC ACTGTTCAGC CTGCTCGGCG GACAGCTAGT
1851 GCAAATGTTG GGCATAGCAC GGGCCATTGC CTGGTACCTG CTGGCGGCGC
1901 TGGGCGCGCT GGCACTTTTG ACGTTCAGCG TCTGGGCCCA ACTGGGGGCG
1951 GCATGGCTGA TTGCCAGCGC CCTCTGCGTG GCAGCCAGCA TGGGCGCCAT
2001 CTCGGCGCTG ATGTTCGGGT TGACCATGTT CTTCACCCGA AATCGGCGCA
2051 ACGCGTCGGA CTATGCCCTG CAAACCACCA TGTTCACCGT TGCGCGACTG
2101 GCGGTGCCGA TCGCCGCCGG GGTGTTGCTC GACCGGGTGG GCTACACCGG
2151 CATGCTCCTG GCAATGACCC TGGCGCTGCT GCTTTCCTTC GCGCTCGCCT
2201 GTCGGGTGCG GGAAAAGGTG GAATCTTCGG CACAGTCGAT ACTCGAGCAC
2251 GAGCGGGTTT GAAGGCTGAA GTGACCGGCC ATGCCCCTTC GGACAATGGC
2301 CTGAAATGCG CGGTCCTTTT GCATAGTTTT TCATGCTCAC GTCATATGAA
2351 AGAACAGCCA ACGGCAATTG CTATAGTCAT CACCACGAAC GATAATGATT
2401 ATCGTTACCA TTGAAATCAA ACAGGATAAG CGATATGCCA CTATCAGCGC
2451 TGGTGGCGCC GGCTGGGGAA CTGAGCCGCG CCGAGATCAA CCGTTACAGC
2501 CGCCACCTAC TGATACCCGA TGTGGGCATG ATCGGGCAGC GTCGGTTGAA
2551 GAACGCCAAG GTGTTGGTCA TCGGCGCCGG CGGTCTTGGC TCTCCGACTC
2601 TGCTCTATCT AGCTGCAGCA GGTGTGGGCA CCATCGGGAT AATCGACTTT
2651 GATCGGGTTG ACGACTCCAA CCTTCAGCGC CAGGTCATCC ACGGGGTGGA
2701 TACCGTGGGC GAGCTCAAGG TGGACAGTGC AAAGAAAGCC ATTGCGCGAC
2751 TAAATCCCTT TGTCCAGGTC GAAACCTATA CCGATCGCCT GGAACGGGAC
2801 ATGGCGATCG AGCTGTTTTC GCGCTACGAC CTGATCATGG ACGGTACCGA
2851 CAACTTCGCA ACCCGTTACT TGGTCAACGA CGCCTGCGTG CTGGCCAACA
2901 AACCCTATGT GTGGGGCTCG ATATTCCGTT TCGAAGGGCA GGCGTCCGTG
2951 TTCTGGGAAA ACGCCCCGAA CGACCTGGGC CTGAACTACC GCGACCTGTA
3001 TCCGGAGCCT CCGCCGCCCG AGATGGCCCC CTCGTGCTCC GAGGGCGGTG
3051 TGTTCGGCAT TCTTTGCGCA TCCATCGCAT CGATCATGGC CACCGAGGCG
3101 GTCAAGCTGA TCACGGGCAT CGGCGAGCCA CTACTGGGTC GGCTGGTGGT
3151 GTACGACGCT TTGGATATGC GCTATCGGGA GCTTCCTGTG CGCCGCCTGC
3201 CAAATCGACA ACCGATCACC GACCTGGCCG AGGACTATCA GGTGTTCTGC
3251 GGTCTGGGGT TGCCCAAAGG TGACACGGCG GACGCCGTGC CAGGGATCAG
3301 CGTAACGGAA CTCAAGAAGC GGATGGATCA GGACGAGGTG CCTGTGCTGA
3351 TAGACGTGCG CGAGCCCACC GAGTGGGACA TCGTCCGTAT TCCGGGCGCA
3401 ATCTTGGTGA CCAAATCGCC CACCGCAGCG CAGACACTGC GCGAGCGATA
3451 CGGGGCAGAT GCCAACCTGG TGATCGTCTG CAAGTCCGGG CGGCGCTCTG
3501 CCGACGTGAC CGCCGAGTTG CTAAACCTGG GCATGCGCAA TGTTCGCAAC

3551 CTCGAAGGTG GCGTTTTGGC CTGGGTGAAG GACGTGGACT CTTCTCTGCC
3601 TAGCTACTGA TAGGAGGCCT AGAAGATGGC ATTGCTTATC AAGCGTCAGG
3651 CGCTGGGGCA GGTTCTGGCT CAAGCACGTC GCGATCACCC ACTTGAAACC
3701 TGTGGAATCG TGGCGTCTTC ACTGGAAGCC CAGTTAGCGA CAAGAGTAAT
3751 CCCAATGCGC AACCAGGCGG CATCACAAAC CTTCTTTCGG CTCGACTCGC
3801 AGGAGCAATT CCAGGTGTTC CGATCTCTGG ATGATCGCAA CGAGTTCCAA
3851 CGGGTCATCT ACCACTCTCA TACCGCGAGT GAAGCCTATC CGAGCAGGGA
3901 GGACATCGAG TATGCGGGCT ATCCGGAAGC GCATCACCTG ATTGTGTCCA
3951 CATGGGAGAA CGCCCGAGAG CCCGCCCGTT GTTTCCGGAT ACTTCGTGGA
4001 AAAGTCATCG AAGAAAGTAT CTCCATTGTG GAATAGCGAC TTTCCAATAT
4051 TTCAATCAGC AATGCCTCAG CCACAGCTAG AGGCAAAAGG AGTTCTACAT
4101 GTCGATTTCA GTGATCGTTC CCACATTGCT GCGCCCGCTG ACCAATGGGG
4151 AAAAGACAGT TTTTACCCAA GGCAACTCGG TGGCAGAGGC CATCGAGAAC
4201 CTTGAACACC AGTTCCCTGG CCTTAAGGCC CGGCTGGTCA GTGCGGAACA
4251 TGTGCATCGT TTCGTCAATA TCTACGTCAA CGAAGACGAC ATCCGCTTCT
4301 CAGATGGGCT CAACACGCCA CTCAAGGCCG GTGACAGTTT GACCGTGCTG
4351 CCTGCCGTCG CCGGTGGCTG ACTCGCACCT CCGGACACCG CTGAAAGAAT
4401 GACCCCTGGC GATTCAAATC CAGGGGCAAG TGCAACGCTT TTGTCTCTCG
4451 GTTTTGAATA TGCCGACACT CTTAAATGAA TTTTCCCTGC TGCATTCATC
4501 CACTTCGTTT CCGCCGAATT GGAATGAACT GCAACTTAGC CTGACGGAAC
4551 AGGCCAGATT ATTGGGCATT TGCCCGCTCG CAATCTCGCC GCCTGTGGAT
4601 ATGGAAGGAG CCGCATTCCA GCTGCAGCAT CCGGCTATTT CTCCTATTCA
4651 GGCCCACTTC GCCTCACCAG CCGGCTGGCT GCCAAATCGA CACCTCTCGG
4701 AGCTGCTGCT GCAGGCGGGC AGCGGTCTTA TGTCGGTGCA CGGCCGTGCT
4751 AGCGGTAGGG CCCAACCGCT GGGCGTGGAT TATCTTTCGA CACTTACCGC
4801 CGTCATGACG CTGCACGGAA CGCTGGCCGC AGCCGTGGGG CAGCTGCGTG
4851 GCGGTGCATT TGATCAGGTT CAGCTTTCTC CACTGGGATG CGGGCTGCTC
4901 AGTATCGGGC AGTATCTGGC AGGCGCCACG GCACCAGAAG ATCGTGAGGC
4951 GTTCCTGCCG GGCGGCTCCG ATCCGCATTT GAGGCCGCCA TTTCGTTCCG
5001 CTGACGGCAT CACATTCGAG CTGGAAACGC TCGACAGCAC ACCGTGGCGA
5051 AGCTTTTGGA CCGCCGTCGG CATTGAATCG GAATTGGCCG GTACGGCCTG
5101 GAAAGGTTTT CTGCTTCGCT ACGCGAGGGC CGTGTCGCCT CTACCTGCCG
5151 CCTGTCTCAC GGCGCTCGCC CGCCTGCGTT ACGCAAAGAT CCAACAATTG
5201 GCAGCGCAAG CGGGTGTTGC GGTCGTGCCC GTCCGCACCG ATGCGCAACG
5251 CCGCGAGGAC CCCGATTACC GGCAGTCACT GGCTACGCCA TGGCAGTTCG
5301 AGTCTTTCCC GCCGTCCCCC GAAAGGCATC GAGACACCGC ATTTCCGTCA
5351 CTGCTGCCGC TACAGGGGAT GCGCGTCATC GAATCCTGTC GACGCATTCA
5401 GGGACCGCTG GCCGGGCATC TGCTGGCATC GCTGGGCGCC GAAGTCATTC
5451 GGCTGGAGCC GCCGGGTGGC GATCCGTTGC GAGCCATGCC GCCCTGCGCC
5501 GAAGGCTGTT CGGTGCGCTT TGACGCGCTG AACCACCTCA AATCCGTTCA
5551 CGAAGTCGAT ATCAAATCCG CCCATGGGCG GCAGTTGGTC TACGAGCTCG
5601 CCCGCGATGC GGATGTCTTT CTGCACAACT GGGCGCCCGG CAAGGCCCAT
5651 GAAATGCAAC TGGATGCTGA ACATCTGCGC AGGGTTCAGC CACATCTCGT
5701 TTACGCCTAT GCGGGAGGCT GGGGCCGGGC TCCCGTCAAT GCCCCGGGTA
5751 CCGACTTCAC CGTCCAGGCC TGGTCGGGTG TGTCCGCCGC CATTGCACGT
5801 CAATCCGGCA TCCGCGGCGG CTCGCTGTTC ACCGTGTTGG ATGTGCTGGG
5851 CGGCGCGATC GCGGCACTGG GTGTGACGGC CGCGTTGCTC AATCGAGCAG

5901 TCACGGGCAC GGGTACTTAT GTCGAGAGCT CATTGCTGGG CGCCGCCGAT
5951 CTGCTGATGC ACAGCAGCGG CAAGGCGTCG AGGGGCATCT TGTCCGGCGT
6001 GTATCCCACG CTATCGGGAC TGATCGCCAT CGACTGCCAA CACCCAGATC
6051 AGTTCCAGTC GCTGGCCATG TTGCTGGACA TTCCTGCCAC TGCGGATACC
6101 TGCCAGCAGA CGCTGGCGGA GCGCTTACGC AAGCGACCCG CTTCGGAATG
6151 GGAAACGGTG CTGAACGAAC GGGGCATCGG CGCCTGTGTA GTCATCGAAG
6201 ACCTCAAGCA GCTCGCCGCC GACACCCGCA TCTCTGAATG CCTCACTCGC
6251 AAGTCTTACT TCTCTGTCAA CGCCCCCTGG AGGTTCCTAT GAACAACGCT
6301 GGCATCATCG ACCTGGTTCC TGCTGAGGAA CGCCAACGTT GGGTGCAGGA
6351 CGGTACCTAC CCGAACCAGC CCGTATTCAC GCTGTTTGCC GCCAAAGCCG
6401 AAGCGCATCC CGACAAGAAG GCCGTGCTGT CGCCGCAAGG TGACGTGACC
6451 TACGGCGAGC TCCTCGATGC AGCCCTGCGG ATGGCTCACA GCCTGCGTGA
6501 TTCGGGGATC GTGGCCGGCG ACGTGGTGGC TTACCAGCTC ACCAACCACT
6551 GGTTGTGCTG CGCAATCGAC CTGGCAGTGG CAGCGCTCGG TGCCATCGTC
6601 GCCCCCTTCC CTCCGGGACG CGGCAAGCTG GATATCCAAT CGCTGGTTCG
6651 CCGCTGCGAC GCGCGAGCGG TGATCGTCCC GCAAGCGTAC GAAGGCATCG
6701 ATCTGTGCGA GGTTATCGAG TCACTGCGCC CCACCCTGCT ATCCATGCGC
6751 CGCCTGATTG TTCAGGGCAA GCCTCGCGAA GGATGGATTA CGCTCGATGA
6801 GCTGATGAGC ACCGAGCCGC TGGATCTCGC CAGCCTACCC AGGGTGTGCC
6851 CGAACTCGCC GGTGCGTCTG CTGGTGTCTT CAGGCACCGA GTCGGAGCCC
6901 AAGCTGGTGG CGTACTCGCA CAATGCGTTG GTTGGTGGTC GCGGGCGCTT
6951 CCTGCAGCGC ATCGCGTCCG ATGGCGAAGA TTTTCGCGGC ATGTACCTCG
7001 TTCCGCTGGG TTCGTCCTTC GGCTCCACTG CCACCTTCGG TGTGTTGTGC
7051 TGGCTGGGTG GTTCGCTGGT CGTATTGCCC AAGTTCGACG TGGATGAAGC
7101 CATCAAGGCG ATTGCGGCAT TTCGGCCGGG CTTCATTCTC GGCGTACCCA
7151 CCATGCTGCA ACGCATCGCC GCTCAACCGG CGTTGGAGAG CATCGACAAA
7201 TCCAGCCTGC GTGGTTTGAT CGTCGGCGGC TCGGTCATCG ACGAGGCCAC
7251 CGTGCGCAAA TGCCGTGATG CGTTTGGCTG CGGCTTCATC AGCCTCTACG
7301 GTTCCGCCGA CGGCGTGAAC TGCCATAACA CCCTGGACGA CCCCATCGAA
7351 GTTGTGCTGA CCAGCGTCGG CAAGCCCAAT CCGGCGGTCT GCGCGATTCG
7401 TCTGGTGGAC GACGAAGGCC GGGAGGTCCG GCAAGGCGAG GTTGGCGAAA
7451 TCACCGCCCG CGGGCCATTG ACTCCAATGC AGTACGTCAA CGCGCCGGAG
7501 CTGGACGAGC GTTACCGCGA CCCGCAAGGC TGGGTGAAGA CCGGAGATCT
7551 GGGCTACATC AACGACAAGG GTTATCTGGT CCTAGCCGGT CGCAAGAAGG
7601 ACGTCATCAT CCGTGGGGGC GCCAATATCA GCCCGACCCA GATTGAAGGC
7651 CTGGTCATGG CGCATCCCGA TGTCGTGACC GTTGCGTGCA TTCCTGTTCC
7701 CGATGATGAT CTCGGGCAGC GGGTGTGCCT TTCCGTCACC TTGCGCGAGG
7751 GTGCAGCGAA GTTTTCCCTG AAAGCGATCA CCGACTTCCT GCGCGAACTG
7801 GGACTGGAGG TGAACAAGCT CCCCGAGTAC CTACGCTTCT ACCGCGCTCT
7851 GCCTCTGACA CCGGCGGGAA AGATCGATAA AAAAGCGCTA ACCGAGGAAG
7901 CCCGCGAGCT GGGCACCTCG GGCATTTGTC CCGCTGGGCC GGGGCAGTCG
7951 ACTCCCGAGC GCAGCTTACG GGAGTACGCA TGATATGCGC GGTCAACCGA
8001 TGATGATGGC TACAGCTTTG ATCTGTGCCT TTGTACCAGG GCCACAGTTG
8051 GCGTTTGCTG CGCCAGGCTC CGCGGCTTCG CCTGACTCCA CGACGCTACC
8101 GGAAATCACC GTCACAGCCG AGAAAATCGA GCGGCCGCTG GAAAGGGTGC
8151 CCGCCAGCGT GGCGGTGATC GATGGCTGGG ACGCCGAGCA GTCAGGCATC
8201 ACTAGCCTCA AACAACTGGA AGGACGCATT CCTGGTCTGT CATTCCAGCC

8251 GTTCGGGCAA GCAGGTATGA ATTC

APPENDIX C

MOTIF INFORMATION ABOUT PROTEINS ASSOCIATED TO THE CARBON TETRACHLORIDE DEGRADATION CAPACITY OF *PSEUDOMONAS STUTZERI* STRAIN KC

Table C.1 Transmembrane domains for proteins identified by Sepúlveda and by Lewis, as predicted by TMHMM, HMMTOP, SOSUI and TMPred.

ORF [a]	TMHMM [b]			HMMTOP [c]			SOSUI [d]			TMPred [e]		
	# tmh[f]	from aa[g]	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa
446	0			0			0			0		
1009	12	11	33	12	11	35	12	8	30	11	11	29
		45	63		45	65		43	65		42	60
		76	94		75	94		71	93		75	93
		106	128		104	128		106	128		106	122
		139	161		139	159		140	162		145	163
		170	188		169	187		167	188		169	188
		226	244		228	248		227	249		225	241
		259	277		258	278		256	278		260	276
		290	308		289	309		285	307		291	307
		317	339		318	342		318	340		316	336
		352	374		352	372		353	375		377	398
		378	396		382	400		378	399			
2435	1	43	65	2	46	70	0			2	43	61
					206	230					206	224
3626	0			0			0			0		
4099	0			0			0			0		
4460	0			0			0			6	109	127
											139	155
											221	239
											428	447
											456	474
											509	527
6289	0			3	88	107	1	87	108	4	89	107
					234	258					243	264
					271	290					274	292
											328	346
K	0			0			0			4	22	41
											124	141
											470	488
											525	543
L	0			0			0			5	226	242
											264	281
											303	32o
											444	460
											640	659
M	0			2	49	73	2	56	76	7	23	41
					82	103		84	105		57	75
											85	101
											194	211
											219	239
											552	569
											591	607

Table C.1 (continuation)

ORF	TMHMM			HMMTOP			SOSUI			TMPred		
	# tmh	from aa	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa
N	12	13	35	12	11		11	16	38	10	11	28
		50	72		50			51	73		52	72
		79	97		79			84	106		81	98
		101	119		103			144	166		170	186
		144	166		144			168	190		208	227
		170	188		169			208	230		237	255
		208	230		207			241	263		274	290
		239	261		237			271	293		304	321
		272	290		273			303	325		338	354
		302	324		300			333	355		368	386
		337	359		329			365	386			
		365	387		364							
O	0			0			0			2	361	377
											454	470
P	0			0			0			5	21	37
											191	210
											247	267
											275	295
											325	343

[a] ORF, open reading frame

[b] TMHMM, transmembrane helix on a hidden Markov model (Sonnhammer *et al.*, 1998)
http://www.cbs.dtu.dk/services/TMHMM-1.0/

[c] HMMTOP, prediction of transmembrane helices and topology of proteins (Tusnady & Simon, 1998)
http://www.enzim.hu/hmmtop/server/submit.html

[d] SOSUI, classification and secondary structure prediction system for membrane proteins (Hirokawa *et al.*, 1998)
http:/azusa.proteome.bio.tuat.ac.jp/sosui/submit.html

[e] TMPred, transmembrane prediction (Hofmann & Stofell, 1993)
http://www.ch.embnet.org/software/TMPRED_form.html

[f] thm, transmembrane helices

[g] aa, amino acid

Table C.2 Transmembrane domains for proteins identified by Sepúlveda and by Lewis, as predicted by PSort, DAS and TopPred2.

ORF [a]	Psort [b]			DAS [c]			TopPred2 [d]		
	# tmh [e]	from aa [f]	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa
446	0			2	112	118	0		
					173	178			
1009	11	47	63	11	10	31	12	15	35
		78	94		50	63		45	65
		103	119		76	93		74	94
		145	161		103	115		102	122
		172	182		146	159		143	163
		228	244		171	187		168	188
		260	276		228	238		225	245
		294	310		259	277		258	278
		322	338		287	309		290	310
		358	374		314	339		323	343
		382	398		381	397		354	374
								378	398
2435	2	56	72	2	44	68	2	50	70
		204	220		206	218		203	223
3626	0			0			0		
4099	0			0			0		
4460	1	456	478	7	35	41	5	56	76
					109	124		108	128
					143	150		139	159
					230	235		458	478
					321	325		506	526
					456	477			
					514	521			
6289	2	89	105	5	91	105	4	88	108
		247	263		248	263		244	264
					281	287		271	291
					310	314		452	472
					462	465			
K	0			5	25	36	3	20	40
					128	139		126	146
					176	178		524	544
					478	482			
					531	538			
L	0			2	444	451	0		
					537	544			
M	0			4	29	33	4	55	75
					63	71		84	104
					88	103		551	571
					560	565		586	606

Table C.2 (continuation)

ORF	PSort			DAS			TopPred2		
	# tmh	from aa	to aa	# tmh	from aa	to aa	# tmh	from aa	to aa
N	8	81	97	11	12	37	12	14	34
		98	114		58	65		50	70
		148	164		80	112		79	99
		168	184		151	163		102	122
		207	223		171	186		145	165
		241	257		209	234		168	188
		274	290		241	262		215	235
		369	385		274	288		238	258
					306	308		271	291
					339	345		298	318
					367	387		332	352
								368	388
O	0			1	362	369	0		
P	0			5	170	175	0		
					196	202			
					254	262			
					281	284			
					299	307			

[a] ORF, open reading frame

[b] PSort, prediction of protein sorting signals ad localization sites in amino acid sequences (Nakai & Kaneshisa, 1991)
http://psort.nibb.ac.jp/form.html

[c] DAS, dense alignment surface method for the prediction of membrane α-helices in prokaryotic membrane proteins (Cserzo *et al.*, 1994; von Heijne, 1992)
http://www.Biokemi.su.se/~server/DAS

[d] TopPred2, topology prediction of membrane proteins (von Heijne, 1992)
http://www.Biokemi.su.se/~server/toppredServer.cgi

[e] tmh, transmembrane helices

[f] aa, amino acid

Table C.3 Leader peptides for proteins identified by Sepúlveda and by Lewis, as predicted by SignalP, PSort, and SPScan.

ORF [a]	SignalP [b]	PSORT [c]	SPScan [d]
446	No signal peptide	No signal peptide	No
1009	446	No signal peptide	Cleave 118 – 119 TW-AD
2435	Cleave 10 – 11 APA-GE	No signal peptide	No
3626	Cleave 11 – 12 ALG-QV	No signal peptide	No
4099	No signal peptide	Uncleavable N-terminal sequence	No
4460	Cleave 124 – 125 LAA-AV	No signal peptide	No
6289	No signal peptide	No signal peptide	No signal peptide
K	Cleave 48 – 49 SAA-SP	No signal peptide	Cleave 42 – 43 FA-AP
L	Cleave 96 – 97 ATA-RQ	No signal peptide	No
M	No signal peptide	No signal peptide	No signal peptide
N	Cleave 118 – 119 SLA-NP	Cleave 30-31 FY-AI	Cleave 70 – 71 GA-LV Cleave 118 – 119 LA-NP
O	No signal peptide	No signal peptide	No signal peptide
P	No signal peptide	No signal peptide	No signal peptide

[a] ORF, open reading frame

[b] SignalP, identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage site (Nielsen *et al.*, 1997)
http://www.cbs.dtu.dk/services/SignalP

[c] PSort, prediction of protein sorting signals and the localization sites in amino acid sequences (Nakai & Kaneshisa, 1991)
http://psort.nibb.ac.jp/form.html

[d] SPScan, sport protein scan (Butler, 1998)

Table C.4 Information about some motifs found by BLOCKS [a] in ORF-2435.

ID [b]	Description	Sequence in protein [c] / Comments
BL00538D	Bacterial chemotaxis sensory transducer	333 - 369 QtlreeygAdanlvivcksGRrsAdVtaEllNLgmRn integral membrane proteins, respond to changes in concentration of attractants and repellents in the environment by changing levels of methylation in C-terminus
BL00886D	Dihydroxy-acid and 6-phosphogluconate dehydratase	36 - 82 rrlkNAkVLVIgaGGlgspTLyLAAAgvgtIgIiDfDrvDdSNLqr introduces a C=O bond in a former C-OH bond
BL00074D	Amino acid dehydrogenases	40 - 74 NAKVLViGaGglGSpTlLYLAAAGvgtIGIIDFDR introduces a C=O bond in a former C-NH_2 bond
BL00573A	pyridine nucleotide-disulfide oxidoreductase	45 - 62 VIGAGglGsptlLYLAaA sequence located ~ 110 amino acids upstream of C-X-X-C redox active disulfide bond
BL01316A	ATP-phosphoribosyl-transferase	344 - 368 nLVIVcKSGRrSaDVTaeLlNlGmR involved in the first step of histidine biosynthesis
BL00110E	Pyruvate kinase	254 - 309 PLpNRQpITDLAEdyQVFcGLglpkGDtadaVPgIsVteLkKRMd QdevpV involved of final stage of glycolysis, activated by AMP and sugar phosphates
BL00557D	FMN-dependent α-hydroxyacid dehydrogenase	26 - 67 IlpDvGmigqrrlknAkvLvigAgGLGsPTLlyLAAAGvgtI introduces a C=O bond in a former C-OH bond
BL01305	MoaA/nifB family	228 - 237 GIGEPLLgRl involved in the synthesis of molybdopterin precursor Z from guanosine
BL00443D	Glutamine aminotransferase	32 - 47 MIGqRRLknaKVLVIG involved in asparagine biosynthesis

[a] BLOCKS, motif finder (Henikoff *et al.*, 1999)
http://motif.genome.ad.jp

[b] ID, BLOCKS identification number for the specific motif

[c] upper case residues indicate perfect matches

Table C.5 Information about a motif found by BLOCKS in ORF-3626. [a]

ID	Description	Sequence in protein / Comments
BL00648A	Bacterial ribonuclease P protein component	62 - 83 FqVfRSLdDRnEFQRVIyHshT Endonucleolitic cleavage of RNA, removing the 5'-extra nucleotide from tRNA precursor

[a] See legend provided in Table C.4.

Table C.6 Information about a motif found by PROSITE[a] Pattern in ORF-4099.

ID [b]	Description	Sequence in protein / Comments
PS00430	TonB-dependent receptor proteins signature 1	**perfect match** X78 79 – DSLTVXPA -86 pattern seen in many TonB-dependent outer membrane receptors

[a] http://motif.genome.ad.jp (Hoffmann *et al.*, 1999)
[b] ID, PROSITE identification number for the particular motif

Table C.7 Information about some motifs found by BLOCKS [a] and PRINTS [b] in ORF-4460.

ID	Description	Sequence in protein / Comments
BL00836G	Alanine dehydrogenase and pyridine nucleotide transhydrogenase	465 - 496 GAiAALGVTaALLnravTgTGtyVessllGAa alanine dehydrogenase-involved in the assimilation of L-alanine as energy source through tricarboxylic acid cycle, pyridine nucleotide transhydrogenase – couples transhydrogenation of NaDH and NADP to respiration
BL00462E	γ-glutamyltranspeptidase	335 - 347 PggdPLraMpPca transfer of glutamyl group from polypeptide to an amino acid
FADPNR3 (PRINTS)	AD-dependent pyridine nucleotide reductase signature	59 - 484 VLdVLGGaIAALgVtAALLNRAVTGT signature found in many reductases and dioxigenases

[a] See legend provided in Table C.4.
[b] (Attwood *et al.*, 1999)

Table C.8 Information about some motifs found by BLOCKS and PROSITE pattern in ORF-6289. [a]

ID	Description	Sequence in protein / Comments
BL00455 (BLOCKS)	AMP-binding domain	196 - 211 vSSGTESEPKLVAYSH
PS00455 (PROSITE)	AMP-binding domain	194 - 205 LLVSSGTESEPK perfect signature for AMP-binding

[a] See legends provided in Table C.4 and Table C.6.

Table C.9 Information about a motif found by Pfam[a] in ORF-K.

ID	Description	Sequence in protein / Comments
TonB_boxC	TonB-dependent receptor C-terminal region	585 - 687 signature found in a variety of TonB-dependent receptors

[a] http://motif.genome.ad.jp (Bateman *et al.*, 1999)

Table C.10 Information about some motifs found by BLOCKS in ORF-L. [a]

ID	Description	Sequence in protein / Comments
BL00296F	Chaperonins cpn60 proteins	271 - 308 GWqlYGRAYViAMTSGmIDeARgTlanLKRAAapGIIV prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions
BL00662F	Bacterial type II secretion protein system protein E	65 - 110 qiFAQAcRyLrsQprDklpalfsRcfhSmatArQAlsaGdwtLSpL involved in translocation of the type IV pilin, DNA uptake and protein export

[a] See legend provided in Table C.4.

Table C.11 Information about some motifs found by PROSITE pattern and BLOCKS in ORF-M. [a]

ID	Description	Sequence in protein / Comments
PS00600 (PROSITE)	Aminotransferases class-III pyridoxal-phosphate attachment site	524 - 561 WILDEIQTGLGRTGKMFACEWEDVSPDIIVLSKSL SGG perfect match with PROSITE pattern, pyridoxal attachment in one of the lysine residues
BL00600B	Aminotransferases class-III pyridoxal-phosphate attachment site	404 - 429 GLERVFLSNSGTAEVEAALKLALAAAS pyridoxal attachment site = lysine residue
BL00600F	Aminotransferases class-III pyridoxal-phosphate attachment site	549 - 561 PDIIVLSKSLSGG pyridoxal attachment site = lysine residue
BL00600C	Aminotransferases class-III pyridoxal-phosphate attachment site	434 - 449 LLyCTNGYHGKTLGAL pyridoxal attachment site = lysine residue

[a] See legends provided in Table C.4 and Table C.6.

Table C.12 Information about some motifs found by BLOCKS in ORF-N. [a]

ID	Description	Sequence in protein / Comments
BL00543D	HlyD family secretion proteins	13 -35 GLSaLIIIamGMPmmIfYAigIL inner membrane bound proteins, secrete hemolysin and alkaline protease
BL00216B	Sugar transport proteins	57 - 106 tFGLaALLSPwAGALvQRMGTRAGIICmFILvGLsFS LMAVLpGFGGLVT 104 - 153 LVTALLLCGTAqSLANPATnQAIAhsvPVARKAGV VGLkQSGVQASALLA
BL01022E	PTR2 family proton/oligopeptide symporter	82 - 117 iCMFLIvGLsFSLmAVIPGfgGLVTALLLcGTAqSI
BL00872A	Sodium:galactoside symporter	53 - 101 LtastfgLAAALLsPWAGaLVqRaGLicmFLLVGIsFSI MAvLpgF
BL01005C	Formate and nitrate transporter	20 - 43 iAmGMpMmIFyaIGIlgpHIVADL
BL00942B	*glp*T family of transporters	239 - 281 vSvStIGamvSCFgamgilSrvLltpiADKlkdetiLlgvLFi glycerol-phosphate uptake

[a] See legends provided in Table C.4.

Table C.13 Information about some motifs found in ORF-O by PROSITE pattern and BLOCKS. [a]

ID	Description	Sequence in protein / Comments
PS00072 (PROSITE)	Acyl-CoA dehydrogenases signature 1	256 - 268 AMSEPEAGSDANG involved in ß-oxidation of fatty acids
PS00073 (PROSITE)	Acyl-CoA dehydrogenases signature 2	464 - 483 QIFGGMGYCTELPIERYYRD
BL00072E	Acyl-CoA dehydrogenase	461 - 503 SAVQIFGGMGYcTElPIERYYRDARVFRlYdGTSEI HRImIAR
BL00072D	Acyl-CoA dehydrogenase	374 - 424 VGArAVGMAsKlLEmSVDfAKQRsQFGAPIGsFQm VQKMLADMqcElYgAR
BL00072B	Acyl-CoA dehydrogenase	280 - 292 WILNGSKhFISdA
BL00072C	Acyl-CoA dehydrogenase	319 - 359 GLeLGPIQEMMGhGthQhGLFFtDCRIapqqLLGEPG RGm
BL00072A	Acyl-CoA dehydrogenase	184 - 194 LGLwAMHmPqE

[a] See legends provided in Table C.4 and Table C.6.

Table C.14 Information about some motifs found in ORF-P by BLOCKS. [a]

ID	Description	Sequence in protein / Comments
BL01230B	NA methyltransferase *trm*A family proteins	187 - 199 FLDLgCGpGmvaI involved in tRNA maturation
BL00094A	C-5 cytosine-specific DNA methylase	185 - 205 rRFLDLgCGPGmvAIALARAI
BL01183B	*ubi*E/CoQ5 methyltransferase	245 - 289 deigsgYDLIwcasvLhfVPDlaqtLRklRaaLaPGGvFVsIh ae involved in biquinone biosynthesis
BL00533B	Porphobilinogen deaminase cofactor binding site	101 - 151 aKrFfakqssDycGdAwafrlrSLrDfgTrLPEyLqVrAIeEap vPTDqcw
BL00379	CDP-alcohol phospatidyltransferase	111 - 147 DYCgdaWAfRLRslrDFGtrLpeylqVrAIeeAPVPt involved in biosynthesis of acidic phospholipids

[a] See legends provided in Table C.4 and Table C.6.

REFERENCES

Attwood, T. K., Flower, D. R., Lewis, A. P., Mabey, J. E., Morgan, S. R., Scordis, P., Selley, J. & Wright, W. (1999). PRINTS prepares for the new millenium. *Nucleic Acid Res* **27,** 220-225.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D. & Sonnhammer, E. L. (1999). Pfam3.1: 1313 multiple alignments match the majority of proteins. *Nucleid Acids Res* **27,** 260-262.

Butler, B. A. (1998). Sequence analysis using GCG. *Meth Biochem Anal* **39,** 79-97.

Cserzo, M., Bernassau, J., Simon, I. & Maigret, B. (1994). New alignment strategy for transmembrane proteins. *J Mol Biol* **243,** 388-396.

Henikoff, J. G., Henkoff, S. & Pietrokovski, S. (1999). New features in the BLOCKS database servers. *Nucleic Acids Res* **27,** 226-228.

Hirokawa, T., Boon-Chieng, S. & Mataku, S. (1998). Classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14,** 378-379.

Hoffmann, K., Bucher, P., Falquet, L. & Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Res* **27,** 215-219.

Hofmann, K. & Stofell, W. (1993). Tmbase - a database of membrane spanning protein segments. *Biol Chem (Hoppe-Seyler)* **347,** 166.

Nakai, K. & Kaneshisa, M. (1991). Expert systems for predicting protein localization sites in Gram-negative bacteria. *Proteins: Structure, Function and Genetics* **11,** 95-110.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10,** 1-6.

Sonnhammer, E. L. L., Von Heijne, G. & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. In *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (ed. J. Glasgow), pp. 175-182. AAAI Press.

Tusnady, G. E. & Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283,** 489-506.

von Heijne, G. (1992). Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225,** 487-494.