

5. Identification of genes involved in Fe(II) oxidation by *Rhodopseudomonas palustris* TIE-1 and *Rhodobacter* sp. SW2

ABSTRACT

Oxidation of Fe(II) by anoxygenic photoautotrophic bacteria is thought to be one of the most ancient forms of metabolism and it is hypothesized that these bacteria catalyzed the deposition of a class of Precambrian sedimentary rocks known as Banded Iron Formations. Testing this hypothesis requires knowledge of the molecular mechanism and components of this metabolism. To begin to identify these components and elucidate the mechanism of photoautotrophic Fe(II) oxidation, we have taken two approaches: 1) we have performed a transposon mutagenesis screen of the Fe(II)-oxidizing strain *Rhodopseudomonas palustris* TIE-1 and 2) we have expressed a genomic cosmid library of the genetically intractable, Fe(II)-oxidizing strain *Rhodobacter* sp. SW2 in *Rhodobacter capsulatus* SB1003 - a strain unable to grow photoautotrophically on Fe(II). In TIE-1, two genes, one predicted to encode an integral membrane protein that appears to be part of an ABC transport system and the other a homolog of CobS, an enzyme involved in cobalamin (vitamin B₁₂) biosynthesis, were identified. This suggests that components of the Fe(II) oxidation system of this bacterium may reside at least momentarily in the

periplasm and that a protein involved in Fe(II) oxidation may require cobalamin as cofactor. In the heterologous expression approach, four cosmids that confer Fe(II) oxidation activity to *Rhodobacter capsulatus* SB1003 were identified. Sequence analysis suggests that the gene(s) responsible for this phenotype encode a permease or a protein with binding domain for the redox cofactor pyrroloquinoline quinone (PQQ).

INTRODUCTION

Bacteria able to oxidize Fe(II) photoautotrophically are phylogenetically diverse and isolated strains include members of the purple sulfur (*Thiodictyon sp.* strain F4), purple non-sulfur (*Rhodobacter* strain SW2, *Rhodovulum sp.* strains N1 and N2, and *Rhodomicrobium vannielii BS-1*) and green sulfur bacteria (*Chlorobium ferrooxidans* KoFox) [41, 52, 69, 70, 163, 182]. These bacteria carry out a form of metabolism that likely represents one of the first to have evolved [20, 41, 185], and it has been proposed that direct photoautotrophic Fe(II) oxidation may have catalyzed the deposition of Banded Iron Formations (BIFs), a class of ancient sedimentary iron ore deposits [67, 101, 182]. In this respect, the effect these bacteria may have had on the Fe cycle of an ancient anaerobic Earth is of particular interest.

Indirect Fe(II) oxidation mediated by cyanobacteria, however, is also likely to have played a role in BIF deposition once the concentrations of O₂ produced by these organisms reached sufficient levels in the atmosphere and oceanic systems of the ancient Earth [37]. To distinguish these two biological processes

from each other in the rock record, as well as from other proposed abiotic mechanisms of Fe(II) oxidation [32], biological signatures that uniquely represent the activity of Fe(II)-oxidizing organisms and that are capable of being preserved in the rock record must be identified.

A first step towards identifying such biosignatures and quantifying the contribution of photoautotrophic Fe(II)-oxidizing bacteria to BIF deposition over time, is understanding the molecular mechanisms of Fe(II) oxidation by extant relatives of these ancient bacteria. Although first reported over a decade ago [182], very little is known about the molecular mechanism of Fe(II) oxidation in phototrophic bacteria. Thus, at present, there are no unique organic biomarkers associated with this physiology, nor are there clear inorganic biosignatures [41].

Here, in an effort to characterize the mechanism of photoautotrophic Fe(II) oxidation, we performed a transposon mutagenesis screen to identify genes involved in Fe(II) oxidation in the newly isolated, genetically tractable, *Rhodopseudomonas palustris* TIE-1 [83]. Two genes were identified; one predicted to encode an integral membrane protein that appears to be part of an ABC transport system and the other a homolog of CobS, an enzyme involved in cobalamin (vitamin B₁₂) biosynthesis. The role these gene products may play in Fe(II) oxidation by this phototroph is discussed.

In addition, to identify genes involved in Fe(II) oxidation in the genetically intractable Fe(II)-oxidizing photoautotroph, *Rhodobacter* strain SW2, a genomic cosmid library of this strain was expressed in *Rhodobacter capsulatus* SB1003 - a closely related strain unable to grow photoautotrophically on Fe(II). Four

cosmid clones that confer Fe(II) activity to *Rhodobacter capsulatus* SB1003 were identified and the specific gene(s) responsible for this phenotype was localized to an approximately 9.4 kilobase (kb) region on the insert of one of these cosmids. The identity of this gene(s) remains to be determined, but potential candidates identified through sequence analysis include genes predicted to encode a permease or a protein with binding domain for the redox cofactor pyrroloquinoline quinone (PQQ).

EXPERIMENTAL PROCEDURES

Bacterial strains, cosmids, and plasmids

Strains used are listed in Table 5-1. *Rhodobacter sp.* strain SW2 (SW2), *Rhodobacter capsulatus* strain SB1003 (1003), *Rhodobacter sphaeroides* CM06 (CM06) and *Rhodopseudomonas palustris* CGA009 (CGA009) were gifts from F. Widdel (MPI, Bremen, Germany), R. Haselkorn (U. Chicago), S. Kaplan (UT-Houston Medical School), and C. Harwood (UW-Seattle), respectively. *Rhodopseudomonas palustris* TIE-1 was isolated in our lab by Yongqin Jiao [83]. 1003, TIE-1 and CGA009 were routinely grown in YP medium (3 g Bacto Yeast Extract and 3 g Bacto Peptone per Liter (BD, Franklin Lakes, New Jersey) and incubated at 30°C. CM06 and all *E. coli* strains were grown on Luria-Bertani (LB) medium and were incubated at 30°C and 37°C, respectively. When testing CM06 for complementation of the *trpB* mutation it carries, this strain was replica plated to Siström's minimal medium agar plates [154]. SW2 was maintained in a previously described anoxic minimal salts medium for freshwater cultures (basal

phototrophic medium) [52] and incubated at 16°C. This basal phototrophic medium was used for phototropic growth of all strains. Drug and other supplements were added as necessary. For *Rhodobacter* strains SW2, 1003 and CM06, concentrations of 1 µg/ml tetracycline, 5 µg/ml kanamycin, and 20 µg/ml gentamicin were used. For gentamicin in liquid, the concentration needed to be dropped to 5 µg/ml. Kanamycin and tetracycline were used at 200 and 75 µg/ml, respectively, for TIE-1. Drug and supplement concentrations for *E. coli* strains were: 15 µg/ml tetracycline, 50 µg/ml kanamycin, 50 µg/ml gentamicin, 0.2 mM diaminopimelic acid (DAP), 8.5 mg/ml isopropyl-β-D-thiogalactopyranoside (IPTG) and 30 mg/ml 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal). For phototrophic growth, strains were incubated approximately 30 cm from a 34 W tungsten, incandescent light source. All cultures containing tetracycline that were incubated in the light were incubated behind UV light filters to minimize the light mediated degradation of tetracycline [45]. Electron donors for photosynthetic growth were added to the basal phototrophic medium as follows: H₂ was provided as a headspace of 80% H₂: 20% CO₂, acetate was added from a 1 M, filter sterilized, anoxic solution at pH 7 to a final concentration of 10 mM and Fe(II) was added from a filter sterilized, anoxic 1 M Fe(II)Cl₂·H₂O stock solution to a final concentration of 1 mM. The precipitate that forms after the addition of the Fe(II)Cl₂·H₂O to the medium is likely a mixture of the ferrous minerals vivianite (Fe₃(PO₄)₂·8H₂O) and siderite (FeCO₃) [41]. For cultivation of TIE-1 cultures, this precipitate was removed by filtration as previously described [41]. *E. coli* strains UQ950 and DH10β were

used for routine cloning and *E. coli* strains WM3064 or β 2155 were used for transferring mobilizable plasmids and cosmids to *Rhodobacter* and *Rhodopseudomonas* strains.

Table 5-1: Bacterial strains and plasmids used in this study.

Strain or plasmid	Genotype, markers, characteristics and uses	Source and/or reference(s)
<i>Bacterial Strains</i>		
<i>E. coli</i> β 2155	Donor strain for conjugation; <i>thrB1004 pro thi strA hsdS</i> <i>lacZΔM15</i> (F' <i>lacZΔM15 lacIq</i> <i>trajD36proA+ proB+</i>) Δ <i>dapA::erm</i> (Erm ^R) <i>pir::RP4 (::kan (Km^R)</i> from SM10)	[48]
<i>E. coli</i> WM3064	Donor strain for conjugation; <i>thrB1004 pro thi rpsL hsdS</i> <i>lacZΔM15 RP4–1360</i> Δ (<i>araBAD</i>)567 Δ <i>dapA1341::[erm</i> <i>pir(wt)]</i>	W. Metcalf (UI-Urbana- Champaign)
<i>E. coli</i> UQ950	<i>E. coli</i> DH5 α λ (<i>pir</i>) host for cloning; F- Δ (<i>argF-lac</i>)169 Φ 80d <i>lacZ58</i> (Δ M15) <i>glnV44</i> (AS) <i>rfbD1 gyrA96</i> (Nal ^R) <i>recA1 endA1</i> <i>spoT1 thi-1 hsdR17 deoR</i> λ <i>pir+</i>	D. Lies (Caltech)
<i>E. coli</i> DH10 β	Host for <i>E. coli</i> cloning; F- <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Δ 80d <i>lacZΔM15</i> Δ <i>lacX74 deoR</i> <i>recA1 endA1 araD139</i> Δ (<i>ara</i> <i>leu</i>)7697 <i>galU galK rpsL nupG</i> (Str ^R)	D. Lies (Caltech)
<i>Rhodobacter capsulatus</i> SB1003	<i>rif-10</i>	R. Haselkorn (U. Chicago), [192]
<i>Rhodobacter sphaeroides</i> CM06	2.4.1 Δ S CII <i>trpB::Tn5TpMCS</i> (Tp ^R Trp ⁻)	S. Kaplan (UT-Houston Medical School), [112]
<i>Rhodopseudomonas</i> <i>palustris</i> CGA009	Wild type (ATCC BAA-98)	Caroline Harwood (UW- Seattle), [93]
<i>Rhodopseudomonas</i> <i>palustris</i> TIE-1	Wild type	Y. Jiao (Caltech), [83]

<i>Rhodobacter sp.</i> SW2	Wild type	F. Widdel (MPI, Bremen, Germany), [52]
----------------------------	-----------	--

Plasmids

pRK415	10.5 kb incP-1 (pK2) Tc ^R , <i>lacZα</i>	[92]
pT198	T198 PCR fragment, including the promoter region, cloned into the <i>Xba</i> I site of pRK415	This work
pT498	T498 PCR fragment, including the promoter region, cloned into the <i>Xba</i> I site of pRK415	This work
pLAFR5	21.5-kb broad-host-range cosmid cloning vector derivative of pLAFR3, <i>ori</i> RK2 (Tc ^R , <i>lacZα</i>)	[92]
pBBR1MCS2	Derivative of pBBR1, (Km ^R)	[102]
pBBR1MCS3	Derivative of pBBR1, (Tc ^R)	[102]
pBBR1MCS5	Derivative of pBBR1, (Gm ^R)	[102]
p2B3	Contains SW2 genomic DNA cloned into the <i>Bam</i> HI site of pLAFR5 that confers Fe(II) oxidation activity and Tc ^R to 1003	This work
p9E12	Contains SW2 genomic DNA cloned into the <i>Bam</i> HI site of pLAFR5 that confers Fe(II) oxidation activity and Tc ^R to 1003	This work
p11B3	Contains SW2 genomic DNA cloned into the <i>Bam</i> HI site of pLAFR5 that confers Fe(II) oxidation activity and Tc ^R to 1003	This work
p12D4	Contains SW2 genomic DNA cloned into the <i>Bam</i> HI site of pLAFR5 that confers Fe(II) oxidation activity and Tc ^R to 1003	This work
pP1	Sub-clone of an ~24 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pP2	Sub-clone of an ~13 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pP3	Sub-clone of an ~9.4 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3	This work

	(Tc ^R)	
pP4	Sub-clone of an ~6.5 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pP5	Sub-clone of an ~4 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pP6	Sub-clone of an ~1.8 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pP7	Sub-clone of an ~1.7 kb <i>Pst</i> I fragment of p9E12 in pBBR1MCS3 (Tc ^R)	This work
pH5	Sub-clone of an ~4 kb <i>Hind</i> III fragment of p9E12 pBBR1MCS2 (Km ^R)	This work
pH6	Sub-clone of an ~1.5 kb <i>Hind</i> III fragment of p9E12 in pBBR1MCS2 (Km ^R)	This work
pP3-gm1	Sub-clone of an ~9.4 kb <i>Pst</i> I fragment of p9E12 pBBR1MCS5 (Gm ^R)	This work
pP3-gm2	Sub-clone of an ~9.4 kb <i>Pst</i> I fragment of p9E12 pBBR1MCS5 (Gm ^R)	This work

Transposon mutagenesis of Rhodospseudomonas palustris TIE-1

Genetic screen for mutants defective in Fe(II) oxidation

To generate a library of transposon mutants to screen for Fe(II)-oxidation defects, the plasmid pSC189, carrying the kanamycin resistant hyperactive mariner transposon [35], was moved via conjugation from the donor strain, *E. coli* β 2155, to TIE-1. A deletion of the *dapA* gene of *E. coli* β 2155 renders it unable to grow without the exogenous addition of DAP to the growth medium [48]. Thus, TIE-1 exconjugants with transposon insertions can be selected on YP agar plates

containing kanamycin but no DAP. Transposon containing TIE-1 exconjugants were picked to 96 well microtiter plates containing YP plus kanamycin and incubated aerobically at 30°C overnight with shaking. Transposon containing isolates were tested for Fe(II) oxidation activity by a cell suspension assay. 20 µl of each clone grown in YP was transferred to conical bottom 96 well microtiter plates containing 200 µl phototrophic basal medium. These plates were incubated anaerobically in the light under an atmosphere of 80% N₂:15% CO₂:5% H₂ in an anaerobic chamber (Coy Laboratory Products, Grasslake, MI). After 3 days of incubation, the plates were centrifuged (3500 rpm for 7 min, JS 5.9 rotor Beckman rotor) and the supernatant was removed. In an anaerobic chamber, cell pellets were washed in 100 µl of anoxic buffer containing 50 mM HEPES, 20 mM NaCl, 20 mM NaHCO₃ and 200-300 µM of Fe(II)Cl₂·H₂O at pH 7. After 5-hour incubation in the light, the amount of remaining Fe(II) was determined by the *Ferrozine* assay [159]. 100 µl of *Ferrozine* solution (1g of *Ferrozine* and 500 g of ammonia acetate in 1 L of dH₂O) was added into each well and the OD₅₇₀ was read after a 10 min incubation. Putative mutants were identified in instances where the total Fe(II) removed from the system was less than ~50% relative to the wild type. At least three independent checks were performed for each mutant.

Southern blot

To verify that the mariner transposon inserted in a random fashion, we performed southern blot on 10 randomly selected mutants from different mating

events. *SmaI* and *SphI* digested genomic DNA from the mutants was separated on a 1% agarose gel and transferred to nylon membrane using a positive pressure blotting apparatus (Stratagene, CA) according to the manufacturer's instructions. Probe DNA was prepared from a gel-purified *MluI* restriction fragment of pSC189 that contained an internal part of the mariner transposon including the kanamycin resistance gene. Approximately 25 ng of probe DNA was labeled with 50 μ Curies of alpha-P32-dCTP using the Ready-To-Go labeling beads (Amersham Pharmacia Biotech). Prior to hybridization, unincorporated radioactive nucleotides were removed from the reaction by centrifugation through sephadex columns (ProbeQuant G-50 Microspin columns, Amersham Pharmacia Biotech) according to the manufacturer's instructions. Nylon membranes were hybridized overnight at 65°C. Hybridized membranes were washed 3 times for 5 minutes each in 2x SSC buffer (20 x SSC: 175.3 g/L NaCl plus 88.2 g/L of trisodium citrate) plus 0.1% SDS (sodium dodecyl sulfate) at room temperature, then twice for 15 minutes each with 0.1x SSC plus 0.1% SDS at 65°C. The membrane was exposed to X-ray film at -80°C for 48 hr prior to development.

Cloning of mariner-containing fragments

To identify the DNA sequence flanking the transposon in the mutants, genomic DNA was digested with restriction enzyme *SacII* followed by ligation at a DNA concentration (2-3 μ g/ml) that favored intramolecular ligation [136]. Ligated DNA was washed and concentrated using a DNA purification kit (Qiagen) and transformed into *E. coli* UQ950 cells. Plasmid DNA was extracted from overnight

cultures of kanamycin resistant clones. The sites of transposon insertions of these mutants were determined by sequencing with primers Mar3 (5'-CTTCTTGACGAGTTCTTCTGAGC-3') and Mar4 (5'-TAGGGTTGAGTGTTGTTCCAGTT -3') that anneal near the ends of the mariner transposon in opposite directions.

Complementation of Fe(II) oxidation mutants

Plasmid pT198 and pT498 were constructed to complement the genetic defect in mutants 76H3 and A2, respectively. Primers were designed based on the corresponding gene sequences in *R. palustris* CGA009 that were analogous to the disrupted genes in the mutants. For mutant 76H3, a 1.4 kb gene fragment was amplified through PCR from wild type TIE-1 with primers T198L (5'-GGCTCTAGATCAACCAGAAACCAGCTTCC-3') and T198R (5'-GGCTCTAGATGTGAGCCACTCTGTCATCC-3'). For mutant A2, a 1.3 kb gene fragment was generated with primers T498L (5'-GGCTCTAGACAATTGCGACAGCTTACGAC-3') and T498R (5'-GGCTCTAGAAGAACCGCCTTCTTGGTCT-3'). The purified PCR products were digested and ligated to the *Xba*I cloning site of the broad host plasmid pRK415 vector to generate the vectors pT198 and pT498 for complementation. pT198 and pT498 were introduced into *E. coli* UQ950 by transformation. Transformants with the inserts were isolated through a blue/white screen on LB plates with tetracycline (15 µg/ml). The plasmids pT198 and pT498 were purified from *E. coli* UQ950, transformed by heat shock into the donor strain *E. coli* WM3064 and moved via

conjugation into the mutant TIE-1 strains. Similar to *E. coli* β 2155, *E. coli* WM3064 not only contains the genes required for plasmid transfer on its chromosome, but also requires DAP for growth. TIE-1 exconjugants containing vector pT198 and pT498 were selected on YP agar plates supplemented with 75 μ g/ml of tetracycline. Colonies were picked and grown up in YP liquid medium with tetracycline (75 μ g/ml). YP cultures were washed and sub-cultured in the basal medium plus tetracycline (75 μ g/ml) with H₂ as the electron donor. Cells were then collected by centrifugation and tested for complementation of Fe(II)-oxidation activity by the cell suspension assay as described above.

Heterologous expression of a genomic cosmid library of Rhodobacter sp.

SW2 in Rhodobacter capsulatus SB1003

Preparation of a genomic cosmid library of Rhodobacter SW2 in E. coli WM3064

Genomic DNA from SW2 was isolated according to standard protocols [43]. 25 mls of SW2 grown phototrophically on acetate was harvested for 10 minutes at 5000 rpm on a JA 25.50 Beckman rotor. The pellet was resuspended in 25 mls of lysis buffer (0.25 mls 1 M Tris·HCL pH7.4, 0.5 mls 0.5 M EDTA pH 7.4, 2.5 mls 1 mg/ml proteinase K, 21.75 mls dH₂O), after which 1.25 mls of 10% SDS was added. After a 4 hour incubation at 37°C, 0.75 ml 5 M NaCl was added. Following sequential phenol and chloroform extractions, genomic DNA was precipitated with 100% EtOH, harvested by centrifugation, washed with 70% EtOH and resuspended in 1 ml of TE buffer (10 mM Tris·HCL pH 8, 1 mM EDTA

pH 8). pLAFR5 and all other cosmid DNA was purified using the Qiagen HiSpeed Plasmid Midi Kit (Qiagen, Valencia, CA).

After optimization of partial digestion conditions on a small scale, 13.2 μg of SW2 genomic DNA was digested at 37°C for 3-4 min with 0.2 Units of *Sau3A*I. After incubation at 65°C for 20 min, the genomic DNA fragments were dephosphorylated for 45 min with HK Thermolabile Phosphatase (Epicentre, Madison, WI) according to the manufacturer's instructions. 26.9 μg of the pLAFR5 cosmid vector [92] was digested sequentially with *Sca*I and *Bam*HI at 37°C.

Ligations containing 7.5 μg total DNA were carried out at a 9:1 molar ratio of insert to vector. Appropriate volumes of vector and insert were combined, brought up to 500 μl with dH₂O and the mixture was concentrated to ~14.5 μl using a Pall Nanosep 30K Omega filter spin column (Pall, East Hills, New York). T4 DNA ligase (Roche, Basel, Switzerland), ligase buffer and ATP (5 mM final concentration to inhibit blunt ligations) were added, the reactions were brought up to 20 μl with dH₂O and incubated at 16°C overnight.

The Stratagene Gigapack III XL packaging extract (Stratagene, La Jolla, CA) was used to package the ligation reaction (0.375 μg total DNA) into recombinant λ phage. An undiluted portion of the packaging reaction was used to infect *E. coli* DH10 β according to the manufacturer's instructions. *E. coli* containing cosmids with inserts were selected on LB + tetracycline, DAP, IPTG and X-gal. Cosmids were purified from 10 random white colonies for restriction digest with *Eco*RI to verify that the library was random and determine the

average insert size. After this diagnostic, the rest of the packaging reaction was used to infect *E. coli* WM3064. The library in this strain was stored at -80°C in LB containing tetracycline, DAP, and 10% glycerol.

Introduction of the SW2 cosmid library into Rhodobacter capsulatus SB1003

The SW2 cosmid library was introduced into 1003, a strain unable to grow phototrophically on Fe(II), via conjugation from the WM3064 hosts. Each cosmid containing WM3064 strain was mated to 1003 independently to ensure complete representation of the cosmid library. WM3064 clones were grown in 1 ml of LB + tetracycline and DAP in 96 deep well plates overnight at 37°C with shaking. The cells were harvested by centrifugation (3500 rpm for 15 min using a JS 5.9 rotor Beckman rotor) and washed once with LB to remove traces of tetracycline. To the each of the pellets, 1 ml 1003 was added and the plate was centrifuged again. Cell pellets were resuspended in 20 µl of YP and spotted individually to a 245 mm by 245 mm YP agar plate. The mating plate was incubated at 30°C for 18-22 hours after which the mating spots were resuspended in 200 µl YP + 1 µg/ml tetracycline. The entire mating was plated to a 60 mm by 15 mm agar pate of YP + 1 µg/ml. Exconjugants were inoculated into 96 well microtiter plates containing YP + 1 mg/ml tetracycline and incubated overnight at 30°C with shaking. After growth, sterile glycerol was added to a final concentration of 10% and the library was preserved at -80°C.

Identification of cosmid clones conferring Fe(II) oxidation activity by cell suspension assay

Cosmid containing 1003 exconjugants were tested for Fe(II) oxidation activity by a cell suspension assay similar to that described above. Here, 10 µl of each clone grown in YP was transferred to conical bottom 96 well microtiter plates containing 200 µl phototrophic basal medium. These plates were incubated anaerobically in the light in a GasPak 150 large anaerobic jar under a headspace of 80% H₂:20%CO₂. Residual O₂ was eliminated from the headspace using a BBL GasPak Brand Disposable H₂ and CO₂ Envelope (BD, Franklin Lakes, NJ). After growth, the plates were centrifuged and the supernatant was removed. In an anaerobic chamber (Coy Laboratory Products, Grasslake, MI), cell pellets were washed in 200 µl of assay buffer (an anoxic buffer containing 50 mM Hepes, 20 mM NaCl at pH 7) to remove traces of the medium and centrifuged. The pellets were then resuspended in 100 µl of assay buffer containing 0.1 mM of FeCl₂ and 20 mM NaHCO₃. After an approximately 20-hour incubation in the light, the amount of remaining Fe(II) was determined by the *Ferrozine* assay [159]. Here, 100 µl of *Ferrozine* solution was added into each well. The plates were then centrifuged, 100 µl of the supernatant was transferred to a flat bottom 96 well plate and the OD₅₇₀ was read after a 10 min incubation. Putative Fe(II) oxidizing clones were identified as those clones that had less remaining Fe(II) than the negative control, 1003 + pLAFR5. SW2 served as a positive control. Putative clones were retested in triplicate. Clones that came

through as positive after the retesting were reconstructed and retested in triplicate.

For cell suspension assays where the total concentration of Fe as well as Fe(II) was followed, exponential phase cultures of cosmid containing 1003 cells grown phototrophically on H_2 + 1 μ g/ml tetracycline were diluted to the same OD_{600} . 5 mls of these cultures were harvested by centrifugation (10,000 rpm on a Beckman JLA 10.5 rotor for 20 min) and then, under anaerobic conditions, the cell pellets were washed once with an equal volume of assay buffer and resuspended in 2.5 mls of assay buffer containing 0.1 mM $Fe(II)Cl_2 \cdot H_2O$ and 20 mM $NaHCO_3$. Cell suspensions were incubated at 30°C in 12 ml stoppered serum bottles 30 cm from a 34 W tungsten incandescent light bulb under a headspace of 80% N_2 :20% CO_2 . Fe(II) and Fe(III) concentrations were measured in triplicate by the *Ferrozine* assay where 50 μ l of cell suspension was added to either 50 μ l of 1N HCl (Fe(II) measurement) or 50 μ l of a 10% hydroxylamine hydrochloride solution in 1N HCL (Fe total measurement). 100 μ l of *Ferrozine* solution was then added and after 10 minutes the absorbance at OD_{570} was read. Fe concentrations were determined by comparison to Fe(II) standards.

Sub-cloning of cosmid clones

To identify a smaller fragment of cosmid 9E12 that conferred Fe(II)-oxidation activity, restriction fragments of this cosmid were cloned into broad-host-range vectors of the pBBR1 series [102]. p9E12 was digested with *Pst*I and

HindIII, pBBR1MCS-2 was digested with *HindIII* and pBBR1MCS-3 and pBBR1MCS-5 were digested with *PstI*. All vectors were dephosphorylated with Antarctic Phosphatase (New England Biolabs, Beverly, MA) according to the manufacturer's instructions. The restriction fragments were gel purified using Qiagen kits (Qiagen, Valencia, CA). Fragments less than 10 kb were purified using the QIAquick Gel Extraction kit and fragments greater than 10 kb were purified using the QIAEX II Gel Extraction kit. Ligation reactions contained an approximately 6:1 volume ratio of insert to vector and were transformed into *E. coli* DH10 β . Clones containing inserts were selected on the appropriate drug containing plates supplemented with DAP, IPTG, and X-gal. Plasmids were purified from putative clones using a QIAprep Spin Miniprep Kit and digested with the appropriate restriction enzyme to verify the insert size. Clones with the correct size insert were introduced into *E. coli* WM3064 by transformation and subsequently introduced into 1003 via conjugation. Exconjugants were tested for Fe(II) oxidation activity in the 96 well plate cell suspension assay format described above.

Sequencing and analysis

All sequencing was performed by Laragen (Los Angeles, CA). Putative ORFs were identified using ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Proteins in the database similar to the translated ORFs were identified by BlastP (<http://www.ncbi.nlm.nih.gov/blast/>), and conserved domains were identified

using the NCBI Conserved Domain Database (CDD)

(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Additional protein

analyses (e.g., sub-cellular localization and motif identification) were performed

using the tools on the ExPASy proteomics server (<http://us.expasy.org/>).

Predicted promoters were identified using BPROM

(<http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup>

[=gfindb](#)), a bacterial σ^{70} promoter recognition program with ~80% accuracy and specificity.

RESULTS

Genes involved in photoautotrophic Fe(II)-oxidation by

Rhodopseudomonas palustris strain TIE-1

To identify genes involved in phototrophic Fe(II) oxidation in TIE-1 we performed transposon mutagenesis on this strain using a hyperactive mariner transposon. We chose this transposon because it exhibits a relatively high transposition frequency in diverse gram-negative bacteria and integrates into the host chromosome with little sequence specificity [35]. The frequency of transposon insertion obtained for TIE-1 was $\sim 10^{-5}$ with this transposon. Southern blot analysis of 10 randomly selected isolates derived from independent transposition events indicated that the transposon integrates as a single event in random locations (data not shown).

We performed a limited screen of $\sim 12,000$ transposon insertion mutants for defects in phototrophic Fe(II) oxidation using a cell suspension assay. Based

on the assumption that strain TIE-1 has the same number of genes as strain CGA009 and that the transposition is purely random, this screen is ~88% saturated assuming a Poisson distribution [36]. Fourteen mutants were identified as being defective in Fe(II) oxidation: eight mutants had general photosynthetic growth defects; the other six were specifically defective in Fe(II) oxidation. BLAST analysis performed on DNA sequences flanking the mariner insertions revealed that the sequence flanking the transposon has significant similarity to sequences from the genome of *R. palustris* strain CGA009 [105] in all cases.

The eight mutants exhibiting general growth defects grew at least 50% less on acetate and H₂ compared to the wild type (data not shown). Two of these mutants were disrupted in genes that are homologs of *bchZ* and *bchX*, known to encode proteins involved in bacteriochlorophyll synthesis [29]. These mutants will not be discussed further here, however, it is not surprising that our screen picked up components of the general photosynthetic electron transport system given the large variance in cell density in the step prior to the cell suspension assay. Two mutants, however, were identified that are specifically defective in Fe(II) oxidation: 76H3 and A2. 76H3 is a representative of 5 mutants that have transposon insertions at different locations in the same gene, whereas A2 was only isolated once. Both mutants exhibit normal photosynthetic growth in minimal medium with H₂ as the electron donor, but their ability to oxidize Fe(II) is less than 10% of the wild type (Figure 5-1A and 1B). Complementation of the disrupted genes indicates that their expression is necessary and sufficient to

restore nearly wild-type levels of activity, suggesting that Fe(II) oxidation defects were not caused by the downstream genes (Figure 5-1C and 1D).

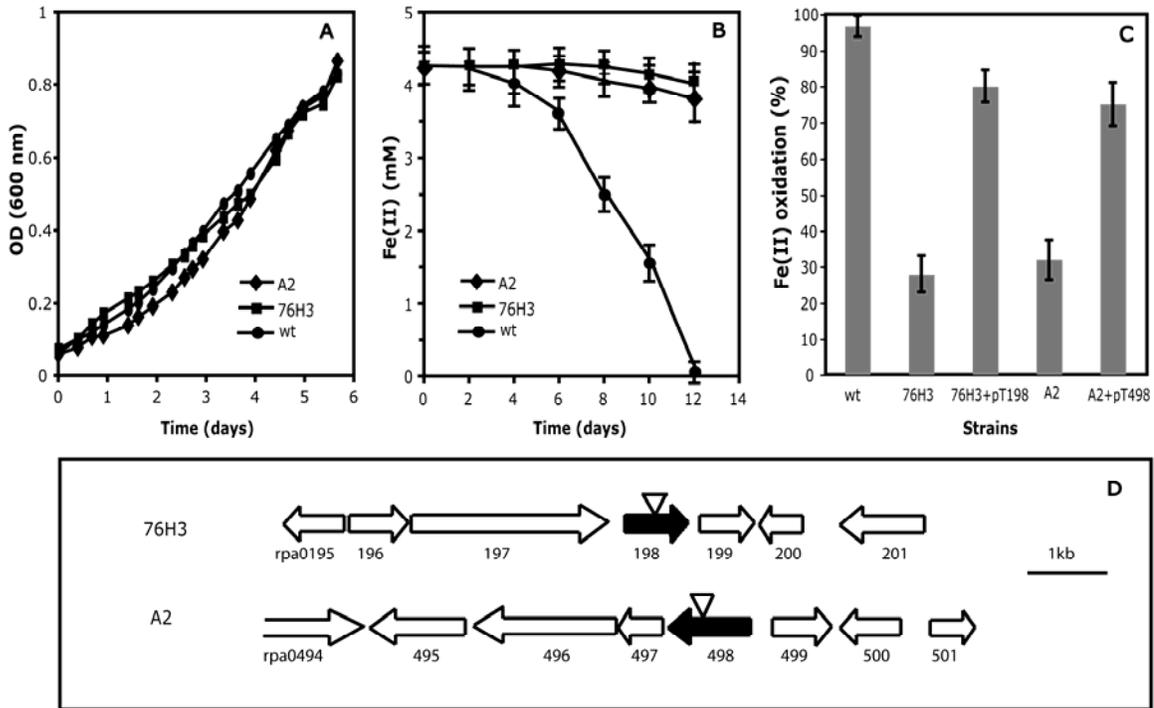


Figure 5-1: Mutants 76H3 and A2 are specifically defective in Fe(II) oxidation.

A. Normal growth of mutant 76H3 and A2 with H₂ as the electron donor. Data are representative of two independent cultures. B. Defects in phototrophic growth on Fe(II) for mutants 76H3 and A2 compared to wild type. Growth was stimulated with H₂ present in the headspace initially. Data are representative of duplicate cultures. C. Mutant 76H3 and A2 carrying plasmids pT198 and pT498, respectively, show 80% of Fe(II) oxidation compared to the wild type in the cell suspension assay. D. Organization of the genomic regions surrounding the mutated genes in mutants 76H3 and A2. The black arrows indicate the disrupted genes and the transposon insertion sites are marked by the open triangles. The

numbers provided below the open reading frames (all arrows) are consistent with the numbers given for the identical regions from the CGA009 genome.

Because the sequence fragments from TIE-1 flanking the transposon insertions were highly similar to sequences from strain *R. palustris* CGA009, we designed primers based on the CGA009 genome to sequence the regions surrounding the transposon insertions in 76H3 and A2 (Figure 5-1D). Both regions contained homologs of genes found in the same order in CGA009. Mutant 76H3 has a transposon insertion in a gene that shares 99% identity over the entire gene sequence (791 base-pairs) to gene RPA0198 in CGA009 that encodes a putative integral membrane protein. BLAST search predicts that the protein encoded by this gene shares 100% identity to a possible transport protein in *R. palustris* CGA009, 85% identity to a probable ABC transport permease in *Bradyrhizobium japonicum*, and 60% identity to a hypothetical transmembrane protein from *Magnetospirillum* sp. MS-1. It is predicted to encode a cytoplasmic membrane protein with 6 internal helices based on sequence analysis with the Psort program (<http://www.psort.org/>) and no known motifs could be identified in this protein by the Motifscan program (http://myhits.isb-sib.ch/cgi-bin/motif_scan). Based on the annotation of the CGA009 genome, the upstream genes encode a putative ABC transporter permease (RPA0197) and a putative ABC transporter ATP-binding protein (RPA0196). The downstream gene (RPA0199) encodes a putative phosphinothricin acetyltransferase.

Mutant A2 has a transposon insertion in a gene (995 base-pairs) with 99% identity to gene RPA0498 in *R. palustris* CGA009 that is annotated as a *cobS* gene. The translated protein sequence is 100% identical to a putative CobS in strain CGA009, 93% identical to a putative CobS from *Bradyrhizobium japonicum*, 80% identical to a well studied CobS from *Pseudomonas denitrificans* and 76% and 71% identical to MoxR-like ATPases from *Rhodospirillum rubrum* and *Rhodobacter sphaeroides*, respectively. Studies of CobS function in *P. denitrificans* have shown that CobS is a cobaltochelatase: a cytoplasmic protein involved in cobalt insertion into porphyrin rings [47]. MoxR-like ATPases belong to a superfamily of proteins with associated ATPase activity (AAA) [78]. Not surprisingly, members of the MoxR family function as chaperons/chelatasers in the assembly of specific metal-containing enzymatic complexes. Based on the annotation of the CGA009 genome, the genes downstream appear to encode a GCN5-related N-acetyltransferase (RPA0497), a CobT homolog (RPA0496), and a conserved hypothetical protein (RPA0495).

Preparation of a Genomic Cosmid Library of Rhodobacter SW2 in E. coli WM3064

The goal of this work was to identify genes involved in photoautotrophic Fe(II) oxidation in *Rhodobacter* sp. SW2. Because SW2 is not amenable to genetic analysis, our approach was to express a genomic cosmid library of this strain in *Rhodobacter capsulatus* SB1003, a strain that is unable to grow

phototrophically on Fe(II), and screen for clones that confer Fe(II) oxidation activity.

The SW2 genomic library was constructed in the broad host range cosmid vector pLAFR5 [92]. This cosmid vector has tandem *cos* cassettes (the recognition sequence for λ phage) separated by a unique *ScaI* site. Digestion of pLAFR5 with *ScaI* and *BamHI* generates two vector arms, each containing a *cos* cassette. During ligation, these arms are joined via a *Sau3AI* digested insert at the *BamHI* site of pLAFR5. Dephosphorylation of the insert DNA prior to ligation decreases the potential for insert-insert ligation. All DNA between the *cos* cassettes is subsequently packaged into λ phage heads and used to transduce the genomic library into the *E. coli* host of choice. The nature of the packaging reaction used here is such that 47-51 kb cosmids are preferentially packaged, thus, inserts of approximately 25.5 - 29.5 kb are expected. Additionally, because the multiple cloning site of pLAFR5 lies within the *lacZ α* gene, clones containing inserts can easily be identified as white colonies on plates containing IPTG and X-gal.

To verify that the cosmid library was representative of the SW2 genome and determine the average insert size, DH10 β was infected with a portion of the library-containing phage lysate. 94% of the clones obtained were white on LB plates containing X-Gal and IPTG, indicating that they contained cosmids with inserts. Restriction digest analysis of ten randomly selected cosmids revealed that these cosmids contained different inserts with an average size of approximately 23.5 kb (Figure 5-2). Although the size of the SW2 genome is

unknown, by comparison to other *Rhodobacter* strains, (*Rhodobacter sphaeroides* 2.4.1: 4500 kb and *Rhodobacter capsulatus*: 3600 kb), an average insert size of 23.5 kb suggests that a library of approximately 700 - 880 clones would be sufficient to guarantee 99% representation of the SW2 genome [36].

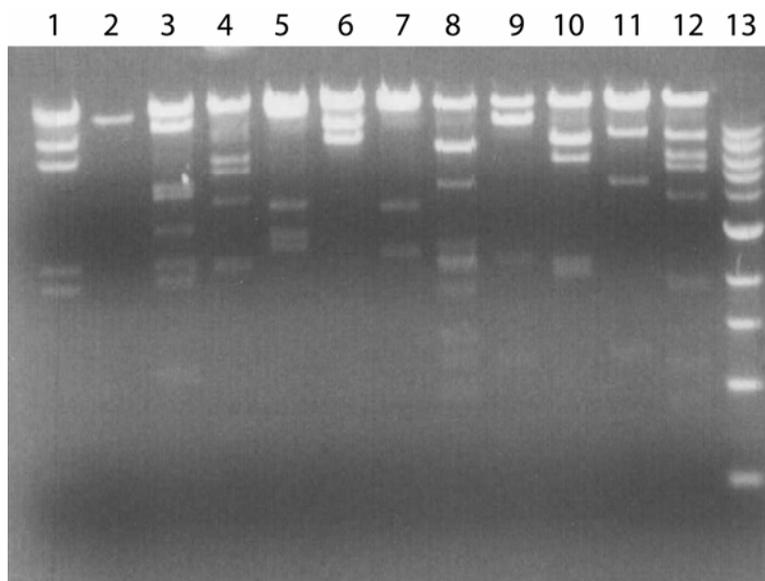


Figure 5-2: Restriction digests of ten randomly selected cosmids with *EcoRI* reveal that the cosmid genomic library of SW2 is representative of the SW2 genome and that the average insert size is approximately 23.5 kb. Lane 1 – λ *HindIII* molecular weight marker; lanes 2-12 – ten randomly selected cosmids digested with *EcoRI*; Lane 13 – 1 kb molecular weight marker (Bio-Rad).

Heterologous expression of the SW2 genomic library in Rhodobacter capsulatus SB1003 and identification of four cosmids that confer Fe(II) oxidation activity

After moving the SW2 genomic library into 1003 via conjugation from WM3064, we verified that DNA from SW2 could be expressed in other *Rhodobacter* species by successfully using the library to complement a *trpB* mutant of *Rhodobacter sphaeroides* 2.4.1 (CM06, data not shown). After this demonstration, 1536 cosmid containing WM3064 strains were mated independently to 1003. All exconjugants were pre-screened for Fe(II)-oxidation activity using a cell suspension assay in 96 well format. Four cosmids, designated p2B3, p9E12, p11B3, and p12D4, showed a decrease in Fe(II) that was at least 99% greater than the negative control, 1003 + pLAFR5 (Figure 5-3).

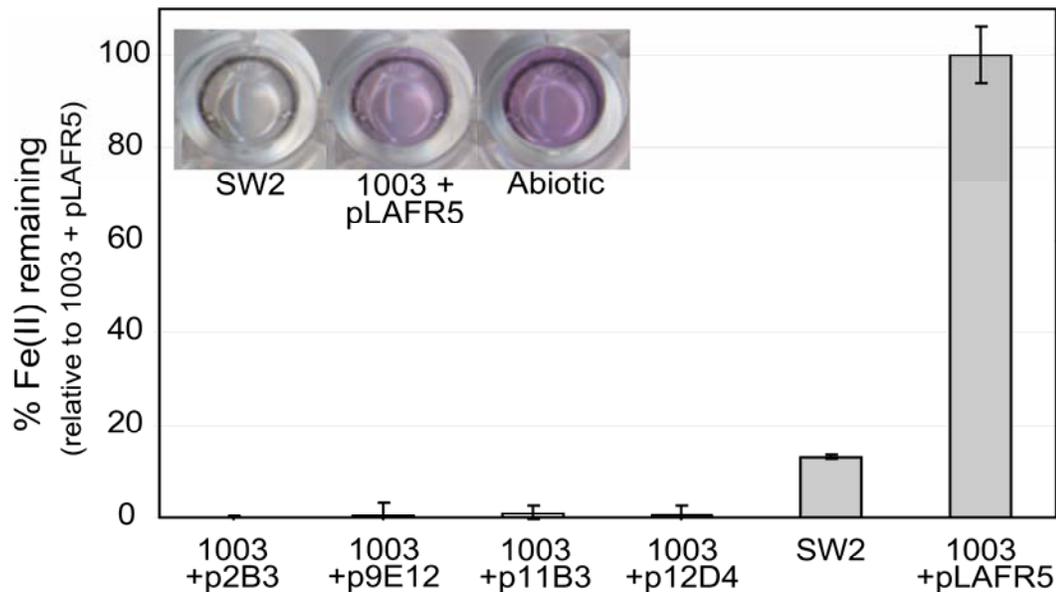


Figure 5-3: Strains of *Rhodobacter capsulatus* SB1003 containing cosmid clones p2B3, p9E12, p11D3 and p12D4 show a decrease in Fe(II) that is 99% greater than the negative control, 1003 + pLAFR5. Error bars represent the error for cell suspension assays of 24, 14, 24, 10, and 9 independent colonies of 1003 containing p2B3, p9E12, p11D3, p12D4 and pLAFR5 respectively, and 4 independent cultures of SW2. The inset shows the color difference between the positive control, SW2 and the abiotic control after the addition of *Ferrozine* during a 96 well plate format cell suspension assay.

Interestingly, we found that 1003 seemed to have Fe(II) oxidation activity in our assay, albeit less than that conferred to the p2B3, p9E12, p11B3, and p12D4 containing 1003 strains. Here a decrease in Fe(II), equivalent to approximately 73% of the total Fe(II) added, was observed (Figure 5-4).

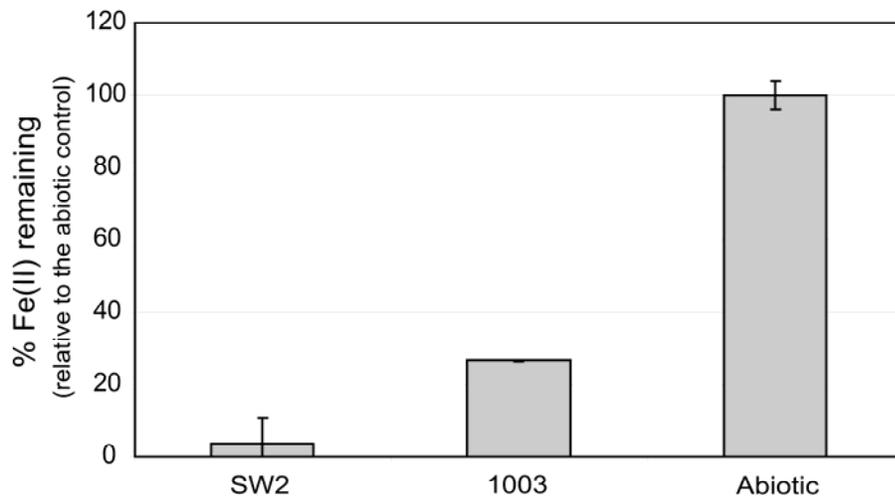


Figure 5-4: In our 96 well format cell suspension assay, *Rhodobacter capsulatus* SB1003 shows a decrease in Fe(II), equivalent to approximately 73%

of the total Fe(II) added. This decrease, however is less than that observed for the p2B3, p9E12, p11B3, and p12D4 containing 1003 strains.

Because only Fe(II) was followed in the plate assay, we could not conclude that the decrease in Fe(II) was due to oxidation as opposed to Fe(II) uptake by the cells or strong Fe(II) chelation by cell surface components or other cell produced molecules. Therefore, to verify that these cosmids were in fact conferring Fe(II) oxidation activity, reconstructed strains containing these four cosmids were retested in a cell suspension assay where the concentration of both Fe(II) and total Fe were followed over time. During these assays, the total amount of Fe stayed constant while Fe(II) decreased. This indicated that the decrease in Fe(II) observed in the cosmids containing strains was attributable to oxidation of Fe(II) (Figure 5-5). Although these cosmids did confer Fe(II)-oxidation activity to 1003, they were not sufficient to confer photoautotrophic growth on Fe(II) (Figure 5-6).

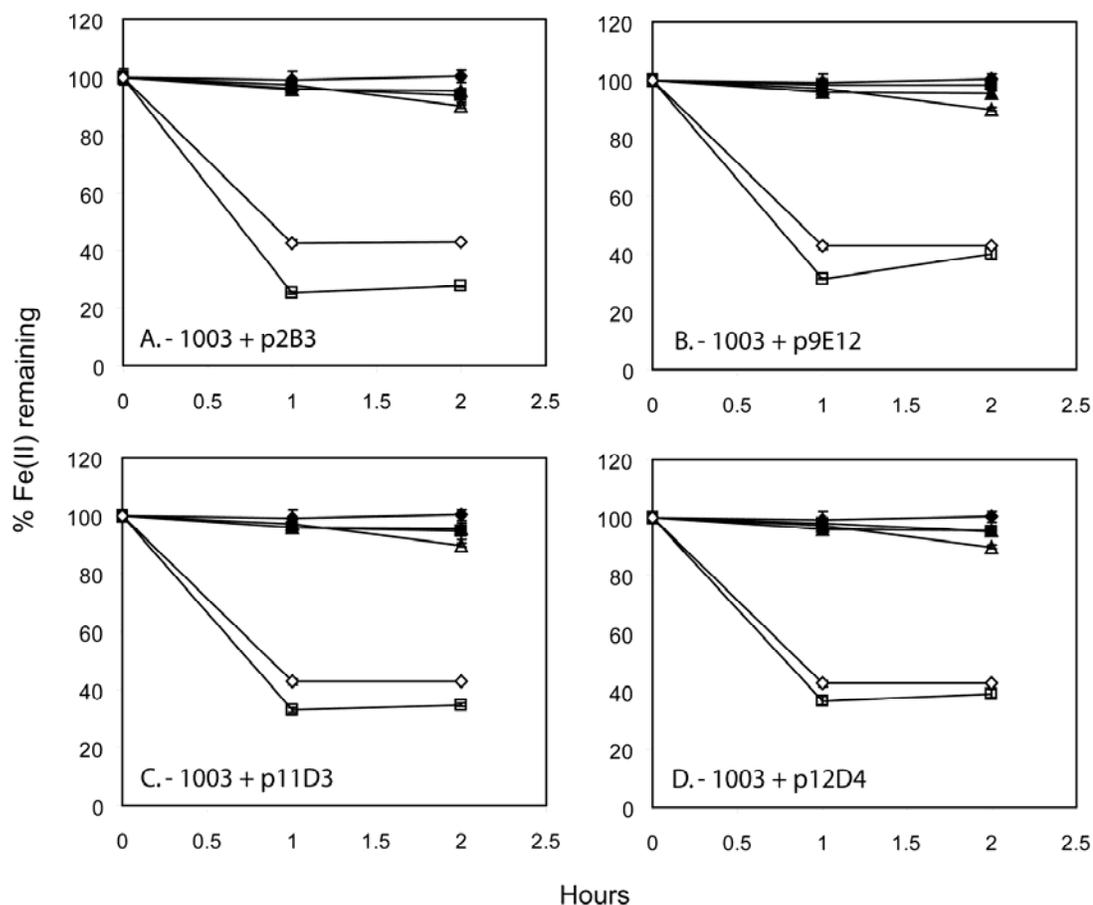


Figure 5-5: Fe(II) oxidation by 1003 + p2B3 (A), p9E12 (B), p11D3 (C), and p12D4 (D) in comparison to SW2 and 1003. On all graphs, ■ – Fe total for 1003 + cosmid; ▲ – Fe total for SW2; ◆ – Fe total for 1003; □ – Fe(II) for 1003 + cosmid strain; ◇ – Fe(II) for SW2; Δ – Fe(II) for 1003. When both the concentration of Fe(II) and total Fe were followed in cell suspension assays over time, a decrease in the amount of Fe(II) was observed while the total amount of Fe stayed constant. This shows that Fe(II) is being oxidized by these cosmid containing strains rather than being sequestered within the cells or chelated by cell surface components or other cell produced molecules. Assays were normalized for cell number and the error bars represent the error on triplicate

Ferrozine measurements. In these assays, it seem as though Fe(II) is not oxidized to completion by SW2 or the cosmid strains. This, however, is due to OD₅₇₀ absorbance by the high number of cells in the sample taken for the [Fe] measurement.

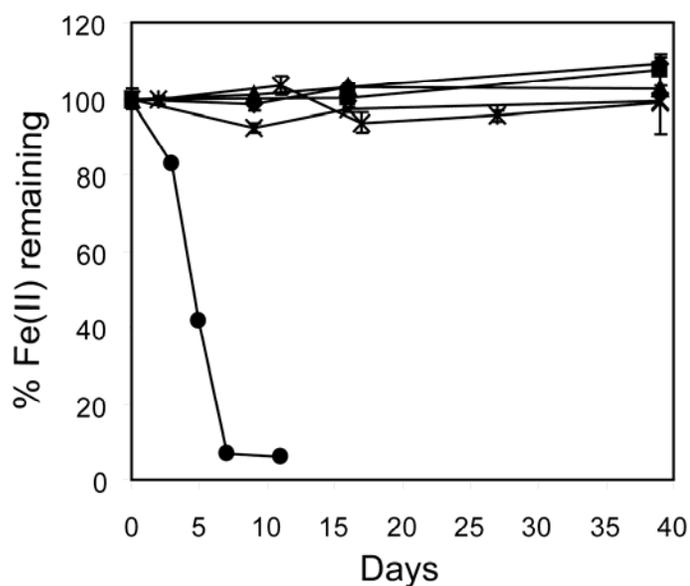


Figure 5-6: Fe(II) oxidation conferring cosmids p2B3, p9E12, p11D3 and p12D4 do not confer the ability to grow photoautotrophically on Fe(II) to 1003. ◆ – p2B3; ■ – p9E12; ▲ - p11D3; X – p12D4; ✕ – 1003; ● – SW2.

Identification of genes on p9E12 that confer Fe(II) oxidation activity

In an initial approach to identify the specific genes on these cosmids responsible for the observed Fe(II) oxidation phenotype, p9E12 was mutagenized *in vitro* using the Epicentre EZ-Tn5™ <T7/KAN-2> Promoter Insertion Kit. After mutagenesis, the transposon derivative cosmid library was

transformed into WM3064, introduced into 1003 via conjugation and exconjugants were screened for loss of Fe(II) oxidation activity. This approach, however, was unsuccessful. Here we found that all of the transposon containing cosmids that no longer conferred the ability to oxidize Fe(II) contained the transposon in different sites in the pLAFR5 backbone, rather than the SW2 DNA insert. Genes in the sequence flanking the transposon insertions of these cosmids included *trfA*, *tetA*, *kfrA*, *trbD/E*, and *korF/G*, indicating that the disruptions may affect plasmid stability, replication and/or maintenance [71, 79].

Moving on to a sub-cloning strategy, the four cosmids were digested with *PstI*, *HindIII* and *EcoRI* (Figure 5-7) so that the inserts of these cosmids could be mapped and common regions among them identified and cloned.

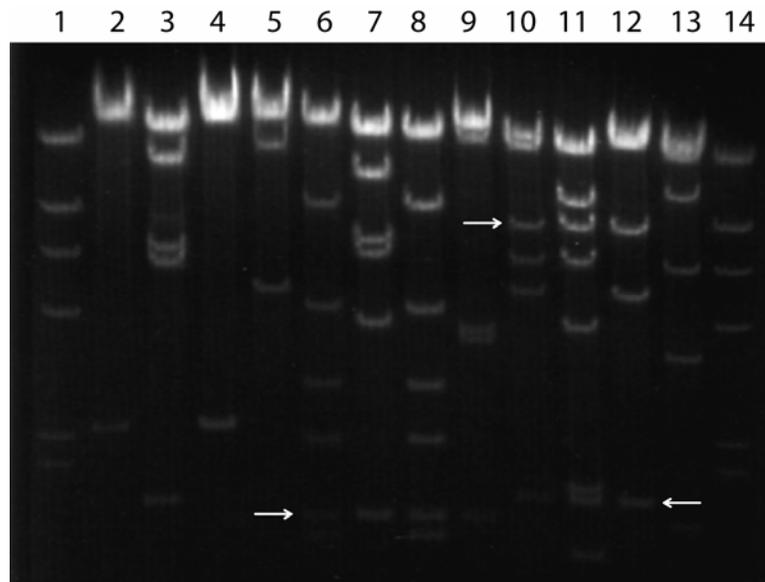


Figure 5-7: Digestion of p2B3, p9E12, p11D3 and p12D4 with *EcoRI*, *HindIII* and *PstI* reveals common restriction fragments among three or more of these cosmids. The common fragments are highlighted by the white arrows. Lanes 1

and 14 – λ *HindIII* molecular weight marker; lanes 2-5 – p2B3, p9E12, p11D3 and p12D4 digested with *EcoRI*, respectively; lanes 6-9 – p2B3, p9E12, p11D3 and p12D4 digested with *HindIII*, respectively; lanes 10-13 – p2B3, p9E12, p11D3 and p12D4 digested with *PstI*, respectively. Some small pieces are missing from this gel. These include *HindIII* fragments of p2B3 (~0.5 kb) and p11D3 (~0.5 kb) and *PstI* fragments of p2B3 (~0.6 and 0.2 kb), p9E12 (~0.2 kb), p11D3 (~0.2 kb) and p12D4 (~0.6 and 0.7 kb).

A restriction map of the four cosmids digested with *PstI* is shown in Figure 5-8. Map for the *EcoRI* and *HindIII* digests were not constructed. Here, discrepancies in the data seemed to suggest that the inserts among these cosmids may be from different regions of the SW2 genome. For example, the only *PstI* restriction fragment that p12D4 shares with any of the other cosmids is an approximately 0.75 kb fragment, present only in p2B3 (fragments 7 of p2B3 and p12D4). However, although the insert of p9E12 spans this region and should also contain this piece, it does not. Interestingly, the *PstI* restriction fragments 6, 4, 5, and 2 of p12D4 are very similar in size to restriction fragments 8, 5, 6, and 2 of the 9E12 *PstI* digest (Figure 5-7 and 5-8).

It is possible that genes in disparate regions of the SW2 genome can confer an Fe(II) oxidation phenotype, and that these different regions are represented among the inserts of the four cosmids. It is also possible that during the ligation of pLAFR5 to SW2 genomic DNA, fragments of DNA from different regions of the SW2 genome were ligated together to form the insert of p12E4,

despite measures taken to prevent such insert-insert ligations. If, however, genes in different regions of the SW2 genome can confer Fe(II) oxidation activity and are represented among the inserts of these four cosmids, it seems that these regions may share similarities in sequence structure.

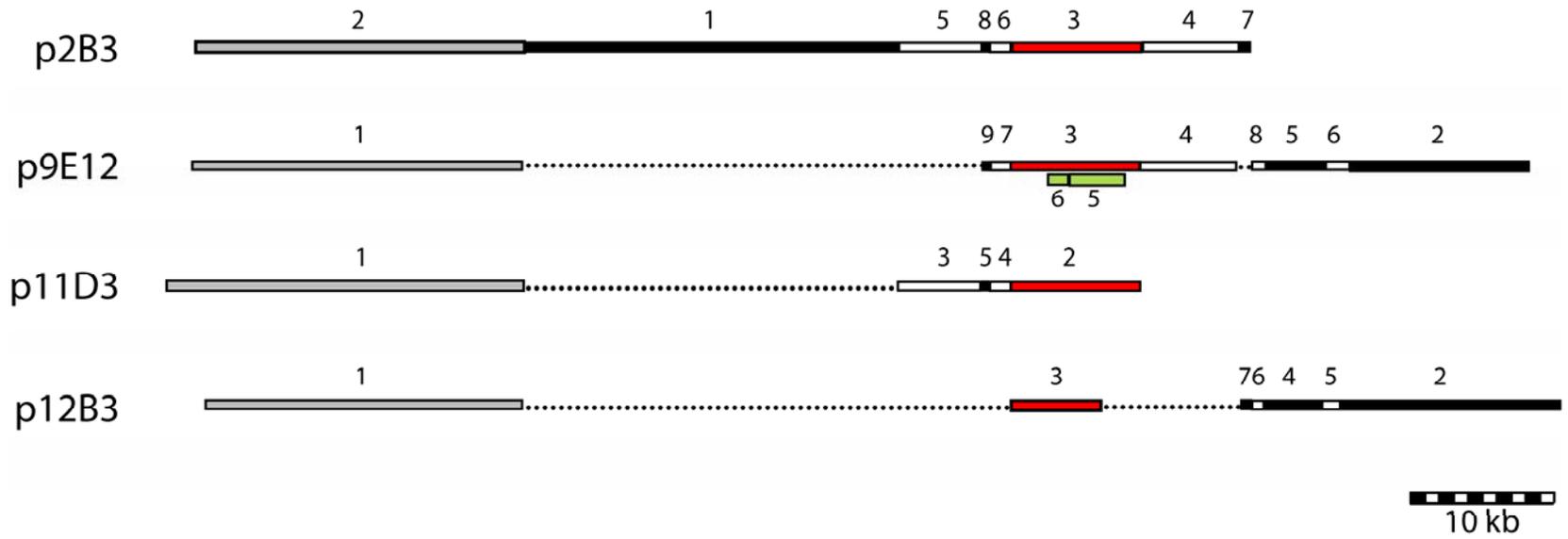


Figure 5-8.

Figure 5-8: Restriction map of p2B3, p9E12, p11B3 and p12D4 digested with *PstI*. The red bars represent fragments in the *PstI* digest common among p2B3, p9E12 and p11D3. The position of fragment 3 in the p12D4 *PstI* digest is inferred from the sizes of the other fragments in this digest relative to those in the digests of p2B3 and p9E12. This fragment likely represents a smaller piece of the common fragment in red due to partial digest conditions. The green bars represent the positions of the pH5 and pH6 inserts on fragment 3 of the p9E12 *PstI* digest inferred from sequence data. Fragment 6 in the *HindIII* digest of p9E12 was common among all the cosmids. From sequence data, fragments 7 and 3 in the p9E12 *PstI* digest are contiguous as are fragments 5, 6, and 2, and 1, although, the position of *PstI* fragment 7 with respect to fragment 3 is arbitrary (*i.e.*, fragment 7 could be on the other side of fragment 3). The positioning, however, is likely accurate given the nature of the sequence at the end of the P3 T7 and the sequence on fragment 4 (P4, 100% identical at the DNA and protein level to *E. coli*) in comparison to the sequence at the beginning of P3 T7 (97% identical at the protein level to *Rhodobacter sphaeroides* 2.4.1). The positions of p9E12 *PstI* fragments 4, 8 and 9 are inferred from comparison to the digests of the other cosmids

Nonetheless, digestion of the four cosmids with *PstI* and *HindIII* did reveal four fragments that were shared among three or more of the cosmids. In the *PstI* digest, fragments of ~9.4, 1.7 and 0.2 kb were shared among p2B3, p9E12 and p11B3 (Figure 5-7 and 5-8) and in the *HindIII* digest an approximately 1.5 kb

piece was shared among all four cosmids (Figure 5-7). That a *PstI* fragment is not shared among the four cosmids likely results from the fact that the initial digestion of the SW2 genome was a partial digest with *Sau3AI*. Thus, some fragment in the *PstI* digest of p12D4 may represent smaller piece of one of the shared fragments among p2B3, 9E12 and 11B3 (likely the ~6.5 kb fragment, Figure 5-8). To test if one of these common fragments contained the genes sufficient to confer an Fe(II) oxidation phenotype, the 9.4 and 1.7 kb fragments of the *PstI* digest were cloned into the broad-host-range vector, pBBR1MCS3 (Tc^R) and the 1.5 kb piece of the *HindIII* digest was cloned into a kanamycin resistant derivative of the same plasmid, pBBR1MCS2 (Km^R) to construct pP3, pP7 and pP6, respectively. None of these clones, conferred Fe(II) oxidation activity upon 1003 (data not shown), however, we later discovered the tetracycline resistance gene located on the vector of pBBR1MCS3 (Tc^R) is not well expressed in *Rhodobacter* strains [168]. Thus, the fact that we did not observe Fe(II) oxidation activity with any of these sub-clones may result from the fact that the construct was not well maintained.

Our next approach was to sequence the insert of p9E12 in hopes that the genes conferring Fe(II) oxidation activity could be identified through sequence data analysis. Here our guiding expectations were that these genes may not be found in the sequenced phototrophs that are unable to oxidize Fe(II), that they may contain cofactor binding motifs that would be indicative of redox activity (e.g., heme binding or multi-Cu oxidase motifs or Fe-S clusters), that they may have no homologs in the database, as no Fe(II)-oxidizing proteins involved in

photoautotrophic growth have been reported thus far, or that they would bear similarity to the genes involved in Fe(II) oxidation identified in TIE-1. In addition, given the observed Fe(II) oxidation activity of 1003 in our cell suspension assays, we also kept in mind the possibilities that the genes responsible for this phenotype may include transcriptional regulators or that gene dosage effects may play a role and introduction of additional copies of these genes could increase Fe(II) oxidation activity.

To reduce the ~ 35 kb insert of p9E12 to a more manageable size for sequencing, the *Pst*I restriction fragments 1, 2, 4, 5, and 6 were cloned into pBBR1MCS3 (Tc^R) and an additional *Hind*III fragment, 5, was cloned into pBBR1MCS2 (Km^R). These clones were designated pP1, pP2, pP4, pP5, pP6, and pH5, respectively. The inserts of pP2, pP3, and pP4, were shotgun cloned and sequenced and the inserts of pP5, pP6, pP7, pH5, and pH6 were sequenced by primer walking. Initial sequencing of pP1 using the T7 and T3 primers (primer sites are located on the pBBR1MCS backbone) showed that the insert consisted primarily of pLAFR5 (Table 5-2), thus the insert of this construct was not sequenced further. The *Pst*I fragments 8 and 9 were not cloned and sequenced because their restriction fragment size suggested that they would not contain more than one gene. In the end, approximately 78% of the p9E12 insert was sequenced. Only the inserts of pP4, pP5, pP6, and pP7 were sequenced enough to be assembled into contigs. The sequences obtained are presented in Appendix 1 and the results of the sequence analysis using ORF finder, BlastP and the CDD are summarized in Table 5-2.

ORFs on the inserts of pP1, pP2, pP5 and pP6 formed a contiguous stretch of DNA that was largely homologous in gene identity and organization to a region of the *Rhodobacter sphaeroides* 2.4.1 genome (locus tags Rsph03003804 to Rsph03003821). Of 17 predicted ORFs, only one in this sequence did not have a homolog in 2.4.1 (Table 5-2). The translated sequence of this ORF encodes a protein with 53% identity to a putative lipoprotein in *Silicibacter pomeroyi* DSS-3 and 40% identity to the putative RND (resistance-nodulation-cell division) multi-drug efflux membrane fusion protein, MexC, in *Rhodopseudomonas palustris* CGA009. In place of this ORF on the genome of *R. sphaeroides* 2.4.1, there is a 903 kb stretch of DNA with no predicted genes. On the pP2 insert portion of this sequence there is a predicted MoxR-like ATPase (COG0714). Although this predicted protein is in the same COG (Cluster of Orthologous Group) as the TIE-1 mutant A2 (putative CobS with similarity to a MoxR-like ATPase from *Rhodobacter sphaeroides* 2.4.1), this predicted protein is not a homolog of the putative CobS protein, nor is the predicted histone acetyltransferase HPA2 downstream of it a homolog of the predicted acetyltransferase downstream of CobS in TIE-1. In fact, direct homologs of the genes involved in Fe(II) oxidation in TIE-1 are not present in the sequence of the p9E12 insert that we have obtained thus far.

ORFs on the pP4 insert were identical in nucleotide and amino acid sequence to genes involved in fimbrial biosynthesis in *E. coli* (Table 5-2). This perfect identity was surprising given the large difference in GC content between *E. coli* and *Rhodobacter* species (50% vs. 68% and 69% for *Rhodobacter*

capsulatus and *sphaeroides*, respectively). Consistent with this, these ORFs have a GC content of 49% versus 63% on average for the rest of the sequenced insert. The possibility exists that these ORFs may not actually be present in SW2, and may represent contamination of *E. coli* DNA during the cosmid library construction. If true, this may also help to explain some of the ambiguities in the *Pst*I restriction map surrounding this fragment (Figure 5-8). Further, genes involved in fimbrial biosynthesis in *E. coli* have no obvious function related to Fe(II)-oxidation; therefore, these predicted ORFs were not considered further.

The inserts of pP3 and pP7 could be linked together based on sequence analysis and contained 10 predicted ORFs. Four of these ORFs had no homolog in *Rhodobacter*. These include: 1) a hypothetical protein with 100% amino acid (aa) identity to a hypothetical protein in *E. coli*, 2) AraJ, a transcriptional regulator (100% aa identity to AraJ of *E. coli*), 3) an ATP binding protein of an ABC transporter (56% aa identity to an ATP binding protein of an ABC transporter in *Erwinia carotovora*), and 4) a predicted permease (80% aa identity to a predicted permease in *Silicibacter* TM1040) (Table 5-2). Given the proximity of the first two ORFs to pP4, for the reasons described above regarding the uncertain origin of the pP4 insert, we are skeptical that these ORFs are endogenous to SW2, and thus, involved in photoautotrophic Fe(II) oxidation. Further, sequence analysis of the pP5 insert reveals that the ORF predicted to encode the ATP binding protein of an ABC transporter lies within the predicted permease ORF, and the expectation value for the permease is lower than that for the ATP binding protein. Finally, all these ORFs, but the permease lie entirely in the region of P3 that is

not predicted to be shared among the cosmids. It seems, therefore, that among these four candidates, the only one with potential to be involved in Fe(II) oxidation in SW2 is the predicted permease. This conclusion is consistent with our results with TIE-1, where we found a putative, cytoplasmic membrane permease that appears to be part of an ABC transport system to be involved in Fe(II) oxidation

Another ORF of interest on the pP3 insert sequence shares 45% aa identity to a putative serine/threonine protein kinase related protein of *Methanothermobacter thermautotrophicus* Delta H and 39% aa identity to a WD40-like repeat containing protein in *Rhodobacter sphaeroides* 2.4.1 (Table 5-2). This translated ORF contains a domain conserved in bacterial dehydrogenases [5] and serine/threonine kinases [131] containing the redox cofactor pyrroloquinoline quinone (PQQ), suggesting potential redox capabilities of this putative protein. In the PQQ containing methanol dehydrogenase from *Methylobacterium extorquens* and other related dehydrogenases, the PQQ cofactor is located in a large subunit of this protein that forms an eight bladed β -propeller structure [5]. Each of the eight blades in this structure is made up of four anti-parallel β -sheets and the interactions between these propeller blades are stabilized by series of 11 residue consensus sequences present in the sequence of all but one of the propeller motifs. Similarly, in proteins with WD-40 repeat domains, these repeats form a β -propeller structure [156]. Members of the WD-40 protein family include proteins with regulatory function and the WD repeats are assumed to be involved in protein-protein interactions [124].

Originally thought to be unique to eukaryotes, proteins with WD-40 repeats have recently been found in a number of prokaryotes including serine/threonine kinases found in *Streptomyces coelicolor* and *Thermomonospora curvata* [88, 131]. No proteins with WD-40 domains, however, have been found to contain PQQ. Thus, while it seems likely that the predicted protein from SW2 contains a β -propeller structure, whether it also contains a PQQ cofactor and possesses redox capability, remains to be determined. Encouragingly, however, recent evidence supports a role for a quinoprotein in Mn(II)-oxidation: the Mn(II) oxidizing activity of cell free extracts of *Erythrobacter* sp. SD21 is enhanced by the addition of PQQ and mutants of *Pseudomonas putida* MnB1 with disruptions in *trpE* that are defective in Mn(II) oxidation can be rescued *in vitro* when supplemented with PQQ [87].

The remaining ORFs encode predicted proteins with homologs in *Rhodobacter sphaeroides* 2.4.1 involved in amino acid biosynthesis (IlvC), carbohydrate transport and metabolism (ManB), folic acid biosynthesis (FolP and B), poly-isoprenoid biosynthesis (IspA), a protein-export membrane protein (SecF) and a permease of the drug/metabolite transporter superfamily (Table 5-2). These predicted proteins have no obvious function related to Fe(II) oxidation, although a role for a permease and a protein-export membrane protein can be envisioned given our results with TIE-1, where we found a putative, cytoplasmic membrane permease that appears to be part of an ABC transport system to be involved in Fe(II) oxidation. The putative protein-export membrane protein,

however, is located on a *Pst*I restriction fragment that is not thought to be common among the four cosmids.

Table 5-2: Summary of ORF finder, BlastP and Conserved Domain Database search results from the p9E12 insert sequence. The top BlastP matches to predicted ORFs in the pP1, pP2, pP3, pP4, pP5, pP6, pP7 pH5 and pH6 insert sequences are listed, as are conserved domains within the predicted ORF when they are present. When the top match is not a *Rhodobacter* species, the highest *Rhodobacter* or related purple non-sulfur bacterium match, when it exists, is also listed. ¹The number of amino acid residues that encode the predicted ORF. ^{2,3}Amino acid identity and similarity between the predicted ORF translation and the proteins in the database that showed the best BlastP matches. ⁴The expectation value; the lower the E value, the more significant the score. ⁵Translated fragments of the same ORF detected in different reading frames, likely due to mistakes in the sequence that result in a frame shift mutation. ⁶The ORF of this predicted ABC transporter ATP-binding protein lies within that of this predicted permease (COG0730). bp = base-pairs.

Table 5-2:

Conserved domains (function)	Protein name (Cluster of Orthologous Groups)	Matching species	Amino acids ¹	Identities ²	Positives ³	E value ⁴
pP1 – 397 bp contig from T3 end						
	non-functional <i>lacZ</i> alpha peptide	unidentified cloning vector	151	25/29 (86%)	26/29 (89%)	5e-05
pfam04956, Conjugal transfer protein TrbC	TelB	Cloning vector pLAFR	125	103/108 (95%)	104/108 (96%)	6e-51
pP1 – 709 bp contig from T7 end						
pfam06271, RDD family (function unknown).	Predicted membrane protein/domain (COG1714)	<i>Rhodobacter sphaeroides</i> 2.4.1	151	102/145 (70%)	119/145 (82%)	3e-51
pP2 – 7264 bp contig from the T3 end						
	ATP-dependent transcriptional regulator (COG2909) ⁵	<i>Rhodobacter sphaeroides</i> 2.4.1	344	195/330 (59%)	236/330 (71%)	e-105
	ATP-dependent transcriptional regulator (COG2909) ⁵	<i>Rhodobacter sphaeroides</i> 2.4.1	54	31/51 (60%)	39/51 (76%)	3e-10
COG3825, Uncharacterized protein conserved in bacteria (function unknown) COG3552, CoxE, Protein containing von Willebrand factor type A (vWA) domain (general function prediction only)	Uncharacterized protein conserved in bacteria (COG3825)	<i>Rhodobacter sphaeroides</i> 2.4.1	394	313/394 (79%)	345/394 (87%)	0.0
COG0586, DedA, Uncharacterized membrane-associated protein (function unknown)	Rhodanese-related sulfurtransferase (OG0607) Uncharacterized membrane-associated protein (COG0586)	<i>Rubrivivax gelatinosus</i> PM1 <i>Rhodobacter sphaeroides</i> 2.4.1	198	64/200 (32%) 44/139 (31%)	99/200 (49%) 66/139 (47%)	2e-15 1e-08
pfam01435, Peptidase_M48, Peptidase family M48 COG4783, Putative Zn-dependent protease,	Putative Zn-dependent protease, contains TPR repeats (COG4783)	<i>Rhodobacter sphaeroides</i> 2.4.1	145	102/134 (76%)	110/134 (82%)	2e-51

contains TPR repeats (general function prediction only)						
	hypothetical protein Rsph03003811	<i>Rhodobacter sphaeroides</i> 2.4.1	112	75/109 (68%)	87/109 (79%)	7e-37
COG1092, Predicted SAM- dependent methyltransferase (general function prediction only)	Predicted SAM-dependent methyltransferase (COG1092)	<i>Rhodobacter sphaeroides</i> 2.4.1	341	158/206 (76%)	178/206 (86%)	2e-87
	Putative lipoprotein	<i>Silicibacter pomeroyi</i> DSS-3	249	91/246 (36%)	131/246 (53%)	9e-35
	putative RND multi-drug efflux membrane fusion protein MexC	<i>Rhodopseudomonas palustris</i> CGA009		36/124 (29%)	50/124 (40%)	7.1
	Dihydroxyacid dehydratase/ phosphogluconate dehydratase	<i>Silicibacter</i> sp. TM1040	83	41/71 (57%)	49/71 (69%)	3e-14
pP2 – 1822 bp contig from the T7 end						
	DnaK suppressor protein (COG1734)	<i>Silicibacter</i> sp. TM1040	96	39/56 (69%)	47/56 (83%)	8e-17
	DnaK suppressor protein (COG1734)	<i>Rhodobacter sphaeroides</i> 2.4.1		26/30 (86%)	28/30 (93%)	2e-07
d00009, AAA, AAA- superfamily of ATPases (membrane fusion, proteolysis, and DNA replication)	MoxR-like ATPases	<i>Rhodobacter sphaeroides</i> 2.4.1	279	243/279 (87%)	263/279 (94%)	e-138
COG0714, MoxR-like ATPases (general function prediction only)						
pfam00583, Acetyltransf_1, Acetyltransferase (GNAT) family (N-acetyltransferase functions)	Histone acetyltransferase HPA2 and related acetyltransferases (COG0454)	<i>Rhodobacter sphaeroides</i> 2.4.1	205	77/134 (57%)	91/134 (67%)	1e-37
pP2 – 970 bp contig of an internal fragment						
	phosphogluconate	<i>Silicibacter pomeroyi</i> DSS-3	126	81/110	92/110	3e-41

	dehydratase			(73%)	(83%)	
pP3 – 1095 bp contig from T3 end						
	hypothetical protein b0395	<i>Escherichia coli</i> K-12	72	71/72 (98%)	72/72 (100%)	6e-36
COG1940, NagC, Transcriptional regulator/sugar kinase (transcription/ carbohydrate transport and metabolism)	AraJ	<i>Escherichia coli</i> K-12	51	45/45 (100%)	45/45 (100%)	2e-20
	Transcriptional regulator/sugar kinase (COG1940)	<i>Rhodospirillum rubrum</i> ATCC 11170		22/41 (53%)	28/41 (68%)	2e-05
COG4987, CydC, ABC-type transport system (energy production and conversion/ posttranslational modification, protein turnover, chaperones) COG1132, MdlB, ABC-type multi-drug transport system (defense mechanisms)	ABC transporter ATP- binding protein ⁶	<i>Erwinia carotovora</i> subsp. <i>atroseptica</i> SCRI1043	130	32/93 (34%)	53/93 (56%)	6e-06
pP3 – 7231 bp contig from T7 end						
pfam01450, IlvC, Acetohydroxy acid isomeroreductase, catalytic domain (catalyses the conversion of acetohydroxy acids into dihydroxy valerates)	Ketol-acid reductoisomerase (COG0059)	<i>Rhodobacter sphaeroides</i> 2.4.1	214	208/214 (97%)	210/214 (98%)	e-118
cd00216, PQQ_DH, Dehydrogenases with pyrroloquinoline quinone (PQQ) as a cofactor COG1520, WD40-like repeat (function unknown)	serine/threonine protein kinase related protein WD40-like repeat (COG1520)	<i>Methanothermobacter thermautotrophicus</i> Delta H	362	72/248 (29%)	113/248 (45%)	2e-14
		<i>Rhodobacter sphaeroides</i> 2.4.1		69/287 (24%)	112/287 (39%)	2e-05
pfam00892, DUF6, Integral membrane protein DUF6 (function unknown)	Permeases of the drug/metabolite transporter (DMT) superfamily (COG0697)	<i>Silicibacter</i> sp. TM1040 <i>Rhodobacter sphaeroides</i>	301	131/282 (46%) 106/278	172/282 (60%) 157/278	5e-59 1e-41

	Permeases of the drug/metabolite transporter (DMT) superfamily (COG0697)	2.4.1		(38%)	(56%)	
COG1109, ManB, Phosphomannomutase (carbohydrate transport and metabolism)	Phosphomannomutase (COG1109)	<i>Rhodobacter sphaeroides</i> 2.4.1	447	377/446 (84%)	400/446 (89%)	0.0
COG0294, FolP, Dihydropteroate synthase and related enzymes (coenzyme metabolism) cd00423, Pterin_binding, Pterin binding enzymes. This family includes dihydropteroate synthase (DHPS) and cobalamin-dependent methyltransferases	Dihydropteroate synthase and related enzymes (COG0294)	<i>Rhodobacter sphaeroides</i> 2.4.1	337	226/331 (68%)	259/331 (78%)	e-121
	Dihydroneopterin aldolase (COG1539)	<i>Rhodobacter sphaeroides</i> 2.4.1	88	29/34 (85%)	32/34 (94%)	4e-09
pfam00583, Acetyltransferase (GNAT) family (N-acetyltransferase functions)	Acetyltransferase, GNAT family Sortase and related acyltransferases Acetyltransferases (COG0456)	<i>Silicibacter pomeroyi</i> DSS-3 <i>Rhodobacter sphaeroides</i> 2.4.1 <i>Rhodobacter sphaeroides</i> 2.4.1	274	85/158 (53%) 42/151 (27%) 27/75 (36%)	103/158 (65%) 63/151 (41%) 35/75 (46%)	2e-37 0.006 0.040
	Predicted permeases (COG0730) ^o	<i>Silicibacter</i> sp. TM1040	47	27/41 (65%)	33/41 (80%)	5e-08
pP4 – 6281 bp consensus sequence						
pfam00419, Fimbrial protein	FimA	<i>Escherichia coli</i> K-12	80	72/72 (100%)	72/72 (100%)	3e-33
pfam00419, Fimbrial protein	FimI	<i>Escherichia coli</i> K-12	212	212/212 (100%)	212/212 (100%)	e-120
pfam00345,	FimC	<i>Escherichia coli</i> K-12	218	218/218	218/218	e-120

<p>Pili_assembly_N, Gram-negative pili assembly chaperone, N-terminal domain. C2 domain-like beta-sandwich fold</p> <p>pfam02753, Pili_assembly_C, Gram-negative pili assembly chaperone, C-terminal domain. Ig-like beta-sandwich fold</p> <p>COG3121, FimC, P pilus assembly protein, chaperone PapD (cell motility and secretion / intracellular trafficking and secretion)</p>				(100%)	(100%)	
<p>pfam00577, Fimbrial Usher protein (biogenesis of gram negative bacterial pili)</p> <p>COG3188, FimD, P pilus assembly protein, porin PapC (cell motility and secretion/intracellular trafficking and secretion)</p>	FimD ⁵	<i>Escherichia coli</i> K-12	444	403/439 (91%)	408/439 (92%)	0.0
<p>pfam00577, Fimbrial Usher protein (biogenesis of gram negative bacterial pili)</p> <p>COG3188, FimD, P pilus assembly protein, porin PapC (cell motility and secretion/intracellular trafficking and secretion)</p>	FimD ⁵	<i>Escherichia coli</i> K-12	468	454/454 (100%)	454/454 (100%)	0.0
<p>pfam00419, Fimbrial protein</p> <p>COG3539, FimA, P pilus assembly protein, pilin FimA</p>	FimF	<i>Escherichia coli</i> K-12	176	176/176 (100%)	176/176 (100%)	4e-96

(cell motility and secretion/intracellular trafficking and secretion)						
pfam00419, Fimbrial protein COG3539, FimA, P pilus assembly protein, pilin FimA (cell motility and secretion/intracellular trafficking and secretion)	FimG	<i>Escherichia coli</i> K-12	169	167/167 (100%)	167/167 (100%)	3e-89
pfam00419, Fimbrial protein	FimH	<i>Escherichia coli</i> K-12	307	287/287 (100%)	287/287 (100%)	e-164
pP5 – 3955 bp consensus sequence						
	hypothetical protein Rsph03003817	<i>Rhodobacter sphaeroides</i> 2.4.1	43	34/38 (89%)	36/38 (94%)	3e-12
COG1391, GlnE, Glutamine synthetase adenylyltransferase (posttranslational modification, protein turnover, chaperones/signal transduction mechanisms)	Glutamine synthetase adenylyltransferase (COG1391) ⁵	<i>Rhodobacter sphaeroides</i> 2.4.1	340	174/245 (71%)	196/245 (80%)	2e-92
COG1391, GlnE, Glutamine synthetase adenylyltransferase (posttranslational modification, protein turnover, chaperones/signal transduction mechanisms)	Glutamine synthetase adenylyltransferase (COG1391) ⁵	<i>Rhodobacter sphaeroides</i> 2.4.1	373	276/351 (78%)	298/351 (84%)	e-153
COG1391, GlnE, Glutamine synthetase adenylyltransferase (posttranslational modification, protein turnover, chaperones/signal transduction mechanisms)	Glutamine synthetase adenylyltransferase (COG1391) ⁵	<i>Rhodobacter sphaeroides</i> 2.4.1	172	110/165 (66%)	124/165 (75%)	4e-53
cd00002, YbaK, protein family (function unknown) COG2606, EbsC,	Uncharacterized conserved protein (COG2606)	<i>Rhodobacter sphaeroides</i> 2.4.1	156	110/155 (70%)	122/155 (78%)	9e-56

Uncharacterized conserved protein (function unknown)						
	hypothetical protein STM1w01000948	<i>Silicibacter</i> sp. TM1040	117	57/91 (62%)	66/91 (72%)	2e-25
	hypothetical protein Rsph03003820	<i>Rhodobacter sphaeroides</i> 2.4.1		56/95 (58%)	66/95 (69%)	4e-24
pP6 – 1770 bp consensus sequence						
COG0800, Eda, 2-keto-3-deoxy-6-phosphogluconate aldolase (carbohydrate transport and metabolism)	2-keto-3-deoxy-6-phosphogluconate aldolase (COG0800)	<i>Rhodobacter sphaeroides</i> 2.4.1	258	72/96 (75%)	78/96 (81%)	4e-32
pfam00920, ILVD_EDD, Dehydratase family.	Dihydroxyacid dehydratase/ phosphogluconate dehydratase (COG0129)	<i>Rhodobacter sphaeroides</i> 2.4.1	90	62/90 (68%)	69/90 (76%)	2e-25
fam06983, 3-dmu-9_3-mt, 3-demethylubiquinone-9 3-methyltransferase COG3865, Uncharacterized protein conserved in bacteria (function unknown)	Uncharacterized protein conserved in bacteria (COG3865)	<i>Rhodobacter sphaeroides</i> 2.4.1	153	87/157 (55%)	101/157 (64%)	1e-37
pP7 – 1299 bp consensus sequence						
COG0059, IlvC, Ketol-acid reductoisomerase (amino acid transport and metabolism/ coenzyme metabolism)	Ketol-acid reductoisomerase (COG0059)	<i>Rhodobacter sphaeroides</i> 2.4.1	185	60/65 (92%)	62/65 (95%)	5e-27
COG0142, IspA, Geranylgeranyl pyrophosphate synthase (coenzyme metabolism)	Geranylgeranyl pyrophosphate synthase (COG0142)	<i>Rhodobacter sphaeroides</i> 2.4.1	102	33/40 (82%)	35/40 (87%)	4e-10
pfam02355, SecD_SecF (protein export membrane proteins)	SecF - Protein-export membrane protein	<i>Rhodobacter capsulatus</i> SB1003	70	43/60 (71%)	52/60 (86%)	2e-17
COG0341, SecF, Preprotein translocase subunit SecF (intracellular trafficking and	Preprotein translocase subunit SecF (COG0341)	<i>Rhodobacter sphaeroides</i> 2.4.1		41/60 (68%)	52/60 (86%)	3e-17

secretion)						
pH5 – 949 bp from T3 end						
	Phosphomannomutase (COG1109)	<i>Rhodobacter sphaeroides</i> 2.4.1		201/241 (83%)	209/241 (86%)	e-109
pH5 – 912 bp from T7 end						
	Predicted permeases (COG0730) ⁶	<i>Silicibacter</i> sp. TM1040	219	42/57 (73%)	44/57 (77%)	1e-15
COG4987, CydC, ABC-type transport system (energy production and conversion/ posttranslational modification, protein turnover, chaperones)	ABC transporter ATP- binding protein ⁶	<i>Erwinia carotovora</i> subsp. <i>atroseptica</i> SCRI1043		32/93 (34%)	53/93 (56%)	6e-06
COG1132, MdlB, ABC-type multi-drug transport system (defense mechanisms)						
pH6 – 269 bp from T3 end						
	No significant hits to the three predicted ORFs		41, 76, & 85			
pH6 – 322 bp from T7 end						
COG1109, ManB, Phosphomannomutase (carbohydrate transport and metabolism)	Phosphomannomutase (COG1109)	<i>Rhodobacter sphaeroides</i> 2.4.1	74	51/73 (69%)	57/73 (78%)	4e-20

Because the sequence of the pP3 insert is only approximately 89% complete, more ORFs may exist in the unsequenced region. The inserts of pH5 and pH6 lie within the pP3 sequence (Figure 5-8), but do not provide additional sequence information beyond that which was obtained from the pP3 insert sequence.

The number of predicted ORFs not present in other *Rhodobacter* species on the sequence of the pP3 insert, the lack of conserved organization to a region of the 2.4.1 genome (as was observed with the inserts of pP1, pP2, pP5 and pP6), the fact that this P3 fragment is common to three of the four identified cosmids (with a region of this fragment likely being common to the to the fourth cosmid as well) and that a *HindIII* fragment within it is common to all four cosmids (Figure 5-7 and 5-8) seemed to suggest that pP3 may contain the genes required for the observed Fe(II) oxidation activity. Revisiting our sub-cloning results where we found that the pP3 clone did not confer Fe(II) oxidation activity to 1003, we realized that the particular tetracycline resistance gene located on the vector of this construct (from pBR322) is not well expressed in *Rhodobacter* strains [168]. Thus, as stated earlier, we could not tell if this sub-clone did not confer Fe(II) oxidation activity to 1003 because it did not contain the required genes or because the construct was not well maintained. Therefore, this restriction fragment was re-cloned in both orientations into pBBR1MCS5 (Gm^R) to generate pP3-gm1 and pP3-gm2. When tested, both of these constructs conferred Fe(II) oxidation activity upon 1003 (Figure 5-9).

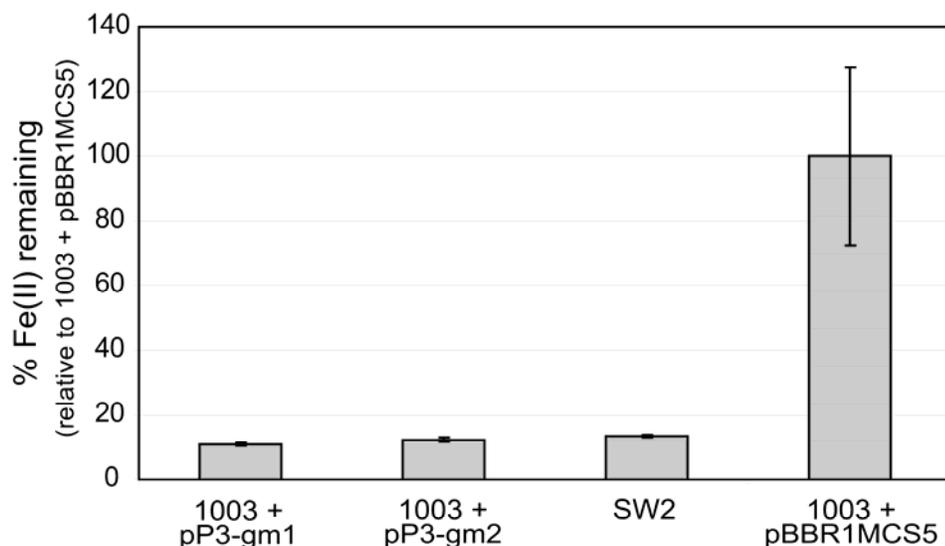


Figure 5-9: The 9.4 kb *Pst*I fragment of p9E12 confers Fe(II) oxidation activity to *Rhodobacter capsulatus* SB1003 when cloned into pBBR1MCS5(Gm^R). This activity is independent of the orientation in which the fragment is cloned suggesting that the gene(s) responsible for the observed phenotype is cloned with its endogenous promoter.

That this activity is independent of the orientation in which the fragment is cloned suggests that the gene(s) responsible for the observed phenotype are cloned with their endogenous promoter. Promoter prediction analysis indicates that σ^{70} promoter consensus sequences lie 182 bp upstream of the putative PQQ containing protein, 655 bp upstream from the predicted permease with no homolog in *Rhodobacter*, and 198 bp upstream of the predicted permease of the drug/metabolite transporter superfamily.

GAGCATTTCACCAAGGAAATCCTGCCGCTGTTCCGCCACCGACAACACCTTTGGCCCGACACGCCGATTGCACCACCTGCCACTTCTCGAACCCAGGAACCG
 CCCAGCTTCCACGAG**CTGAAC**CTGACCACCTATGA**GGGCATCA**GCTGGG**GGCGG**ATTCCGGTGGCCAAAGGTGTCGACAATGCCACCAAGGTGATCATTTCCG
 GGTGACCCCGGAGGCTCGAAGGTGTTCCAGCACCTGACCGAAGACCGCATGCCGCCCGCATCGACCCCTCGGAAGACCGCGACCATCCGAACACCCAGA
 TCCTGTTCCGCTGGATCAAGCAAGGGGCCAA**ATCGAGTAACGCCAACCGCGCCTTGTGCTGCAATCCCTGATCCCGCGCCTGGTTTTAGGCCGGCG**
GTCTTGGCTGCGCAAGACCTGCCGAACTTCCGCGCTGCCGCGCTGGAAACCGGCAATGCCGTTCTGAACCCGCTCGCTATGGTAGGATCGCCTGC
TGTTCTGCCGGCACAAGACCATCGCGCGATTGCCCGGATGCTGCCAGCCGCTGTGGCAGCTGCCGCACGGCTTTGACAAAGCCCGCGCAATTCGGC
CGCGCCCGCGCGCGCAAGGTGATCTCGCGCGCGCGTCTGGCTGGCCGGCTATGATGCCGCCACCGGACCGAACTTGGCGGCATCGCGCAAAAG
ATTAGATCGGTGTGCCCTTCTGACCCCGACCCACACGCTGTTGGCGATGGTACTGGCTGATCGCGCTGGATACCGCGACCGCGCGCGAATTGTGGC
GCTTTGCCGCAATCGAGGATACCATCATCGCCTATGCCCGGTGGCCGATGCCGATACCCGCTATGTCCGCCCGCGGACCGCGCGCTTATAGCCTGTG
GCTGGCGATGGCAGCTGAAATGGCGGTGATGGCCGGCGCAATGGCAATACCTTGGCCGATAGCATCGAGGACGGGATTCTGGTGGCGCGCAC
CTATCAGAGAACCTCAAGGGCTGTGCTGGCCGATGGCCGGAGTTGTGGAGCTTTTACGCCGGCAATTTATCAACTCGCAACTGGTGGCGGACGGT
TGGCGTATCTGGTCAACCACCGGCTATGCTATCG**TCGACACCACAG**GGTGTGTA**CCGCTG**TC**GCTATCAGACCACCGATTACGACAACACCGG**
CAGCAACTGGCCCTCGTGGTGGCCGAACTGCAAGCCTTTGACGCCAAGCTTTATGGCTGTGCTGACCATGTGCTGACCCGGCTGGATAGCGCCACC
GGCACCAGCATCGGGAATGGCGGATGCCGACAAGATCCGCTATGGCGCTTCCCGTAGCCGGTATGGCATCGCCTCCCGACCGGAAAGCGCGGAG
GTGCTGTGACCGCTTCCCTGACCGGCCCTTCCCGCCGAGCAACGCCACCTATAACCCCGGCC**ATGACCGACACCCCGCTCAGAAAGCCCGGCTT**
GCCAACTGGTGGGCTATCGCCTGTGCGGTGATCTGGGGTGGCTGCTTATGGGGGTGAAGTGGCGCTGGCGGGTTGGCGGCTACGATCGCGG
CCATCCGATCGCGCTGGCCGCATCGCGCTTTCTGATCGCGCGACATGGGTCTGGCCTTCCCGCGACCGCGGGTCTGGCGCATGCCATCGG
CATGGGGTCTCTCAACGCCCTGCCGTTTGCATGCTCAGTTGGGGGAGTTGCATGTGGCCCTCGGGCTTTGCCGCAATCACCATGGCGCGGTGGCG
CTGTTACCTGCTCTTGGCGCACCGGCTGATCCGGGGGAGCAGATGACGGGTGGAAACTGGCGCTGGGCTTGGCTCGCCGCGGTTGCTGGTGG
TGATCGGCCAAGGGCGCTTGCCTCGTGGGGGGGATGTCGAAAATCTGGCGCGCTGGCCTGCTCGCCTCGACCTGTGCTATGCCATCGCGCCAT
CATCACCCCGCGCTGCCCGCAGTGGCGCTGGTGGCGTTCTGACCCGCCCTGATCGCGCGAGCGTGATGATGATGATGATGATGATGATGATGATGATG
AGGACCCCGGAGCTTGGCGCGCCACCCAAAGCGCGCTGCTGGCGCTGCTTACCTGGGCTTGGCCCGACCGCGCTGACCCGCTGCTGCTGTGTA
GGTATCACCCCGCTGGCCACCTTCTGACTGTGCAACTACCCAGTGGCGCTGTGGTGTGCTGTTGGCAGCATCTTGGCAGGAAATTCG
CGCCAGCTTCTTGGCGCGCTGGCCTGATCTGGCAGGCTCGCCCTCAGCAACATCCGCCGCAAGGCTTGACCCTGATTTACCTTTGCTCAAATATCC
 TGGGGGGTAGGGCCTATCGCGCGAGGGGGGCAACGCCCCCTTCTGCTCAAATCAAACCGCGCTTCCACCGCGCGCCAAATGCTGCCACACCATC
 GGCCAGCAGCAGCTCTGCTCACATTCCGCCATCACCCGACCAAAAGGCTCGGTGCCGCTGCGGATCAGCAGCCCGCCGAGCCCTGCAACCCGCGCT
 CGGCATCAGCAATCACCGCCCTGCACCGCCCGCGCGCCAAAGGGCTGCAACCCCGCACCGTAAACGCACTTCTCAGCATCTGCGGACCGCTGAAAGGTT
 TGGCCAGCGCGGATGCCGGCTGTTCCGGTGCGCCCATCTCGGCCAGAAACTGCAACCCGGCAATCAGCCCGTCCGCCGTTGGTGGCCTAAATCGGTATCAC
 GATATGGCCGAGCTGCTGCCGCCAGTTGAAACGCCCGCGCATCGCCTGACCCACATAGCGGTGCGCGACATTTGGTGGCTCCAGCCGACGCCCG
 CGCGCTCCAGAAACCGCTCCAGCCCGAGGTTCCGACATCACCGTGGCCACCGCTGCCCGCTGCAGCCGCTGCAGCCGCCCTCTCCGCGCGCGCGCAG
 GCGCATGATCTGGTCCGATCCGCCACCTTCCCGGTCTGATCCAGAATCATCACCCGGTCCGCGTCCGATCCAGACAGATGCCGACATCGCGCCATGC
 GCCACCACCGCCTCGCGCGCAGTCTGGGTATAGGTCCAGCCACACCGTCTGTTGATATTGGTGCATTCCGCGCCACCCCGGATCACCTCGCGGCC
 CAATTTCCACAGCACTCGGGGGCGCACGATAGGCCCGCGCATTGGCGCAATCTATCACACCTTACAGCCATCAAGCCGACGCCGACGGGAAAGGTGG
 TCTTGGCATATTCTGATAGCGCCCGCGGCCATCGTGCATCGCTTGGCCCGCGCGATGTTCTGCGGCTGCGCCAGCGCAATCTCGCCCGCAGAAATCGCC
 TCGATCTCTTTCCGCATCATCCGACAGTTGAAGCCGTCCGGGGCGAAGAACTTATGATCCATTGTCTGATCGGGTGTGGAGGCGCTGATCATCAC
 CCGAGATCGCGCGCATGCTGCGGTGAGAAACCCACCGGAGGCTGGCACCCCGCAGCAGCAGCAGTTCATCCCGTGGAGGTCAACCCGGCGG
 TCAGCGCTTTTCCAGCATGATCCCGACAGCCCGTCTCTTCCGATCACACCCCGGTCCGCGCAGTCCGCTGCGCGGAGGCGCGGCAAGAACCCCGCCG
 CGCCCCAGCGCAAGGCCATCTCGGGGTATCGGGTAGGTATTGGCCCGCCCGCACCCCATCGGTGCCGAAAGGTGCTCGGTGATGCGCCCTCACT
 CTCTAACACAGCCGCCACAGCGCCAGCGCTGCCGGTTCACACACATCATGACCCCGCAGGATCTGGCCCGCTTGGCCACCCCGGCCAGCGCCACC
 GCCAGCGACCCCGCGCGCGCGGTGCCCTGCGCCTCGCCGCAATGTTGGCGATGAAATTTTTCGCGGACACCCCGCAGCAGAAATGGCGAGCCAGAC
 CATGGAACA**GGCACA**CGCCCGCGATCAGCGTCA**GGTTATGCT**GCAAGG**CT**CTTCCAAAACCGATGCCGGGATCAACCCAGATCGGGGCGCGGGGATCCCC
 GCGCGCAGGCCAGCGCCACCGCTTCCGCCAGATAGTCTGAAACATCCAGAGCCACGTCATCATATTGGGGATGCTCTGATGGTGTACCTG
 CGCGTGCATCAGGACAGCCGGCGCATCCGCTGCGCCACACCTCCGCCATGCCCGGCTATAGCCAAAGCGGACACGTCATTGACCATGCTGCCCGG
 CCGCAGTCCCGCTGCGCACCCGCCCTTGGGGTGTGATGAAATGGCGGTGGCAATCCCGCCCGCGCAGCGCGCAATCACCGGGGCGGTG
 GCGAAATCTCTCGGCTCCGCCACCTCGCGCGCCCTGGCCGAGTCTTTCGCCCGCGATGTGATGATCTCGGCCCGCCCTGCCATCATCCGGCC
 TGGCGGTGGCGGCATCCGGGGCAGAAACCGCCCGCGTCCGAAAGCTGTGGGTGTGACGTTTCCAGGATACCCATAGCCCGCGCGCGACAGGTC
 GCCGGCCAA**ATCGCGGACCGCGCGGTGACGCGCGCGCTGCCGCGAGATCGCGCACCGGAATCAGCTCGCGCACCGGCTCGCGGACAG**
CACCTGACCCCGTCAACACAGCACAGCCCGCGCGCGTCCGCGCATCGGTCATCAATCGGTGCGAAGTATCCAT
GCCCCCTGTTAGGGCTTTGCCCGCGCTTGGCAAGCCCTTCCGCCCTAAGCCCGCAGCACCGGACCGCGCTTCCCGCGGAAAGTGTACAGCGCGT
GAACGATATCCGCGCGCGCCAGCATTTCGCCAGCCACAGCGCCAGCGCACCGGATCGTGGCGCGCGGCTGCGCGATGCCCGCGCATG
CGCGCTGTGAAACCGCATCATGCATCGGTGGCACCCAGCGCATGAAATCCCGTTGACCGAGGCGCAGGTGGCGCAGCATTACATCAGCGCGCGG
CGGTGATCTGCTGCCATGTGGCGGAAGGCGCTGGGCGAGGTGATCGGCTTCAATCGCTGGACCGCACCCCGGCTTGGCCGACCGGCTGGGGGATTCG
GCACCTTCGATCCCGCATCTGAAACGACAGCGCGCTGGGGCGCGCTTTTCCCGCGACCTGTGCGCGGCAAGGCGCGGCTGGGGTGGCGGTGATCA
ACGCCACATCCGCGCGCAATGCGCGCGGGCTGGGCTATTACCGCGCGCGGTTTACCGGATTACGCCACTGACCCGGCTTGGCTGCGCGAGC
GCACCGAGTCCGTCGATCAGCCCGCGCTTGCATCTGGCCTGACGCGGGTTTACTTCCCGCGCGCAGCCCGATGAAGGC**TCGCCAATCCCGGTC**
CGCATGTTTCAAGTACGCTGCCCTTGGCACCTGCTACCTTGGCGCGCTTGGCGGGGTTGTTGACGCCATTGCGCGCGGTGGCGGCTCATC
ACTCTGCCGCGCTGTTGCTGGCGGGG

Figure 5-10: Sequence region of the pP3 insert that contains genes predicted to encode a putative PQQ containing protein (bold red), a permease of the drug/metabolite transporter superfamily (bold blue) an acetyltransferase of the GNAT family (bold pink), and a permease with no homolog in *Rhodobacter* (bold green). Candidate promoters for these genes are highlighted in the color representative of the gene sequence.

DISCUSSION AND FUTURE WORK

Fe(II) oxidation activity of Rhodobacter capsulatus SB1003 and Rhodopseudomonas palustris CGA009.

It is intriguing that *Rhodobacter capsulatus* SB1003 pre-grown on H₂ has Fe(II) oxidation activity in our cell suspension assays (albeit less than that of the cosmid containing strains) (Figure 5-4). In addition, it has been observed that photoheterotrophically-grown cells of *Rhodopseudomonas palustris* CGA009 also have Fe(II) oxidation activity [83]. In the case of CGA009, the observed activity is equivalent to that of *Rhodopseudomonas palustris* TIE-1 under these conditions. This indicates that Fe(II) oxidation can be decoupled from growth, as neither 1003 or CGA009 can grow photoautotrophically on Fe(II).

The cell suspension assay, however, did not decouple Fe(II) oxidation from the photosynthetic apparatus, as no Fe(II) oxidation occurred in the dark (data not shown). This suggests that either we have not yet identified the specific conditions which allow 1003 and CGA009 to grow photoautotrophically on Fe(II) or TIE-1 and SW2 contain components not present in CGA009 or 1003 that allow them to conserve energy for growth from Fe(II) oxidation. Given that both of the genes identified in our transposon mutagenesis screen are also present in *R. palustris* strain CGA009, if the latter is the case, it is possible that essential genes for this process are missing from 1003 and CGA009, mutated, or not expressed. To resolve this, a screen to identify TIE-1 mutants that are incapable of phototrophic growth on Fe(II), rather than oxidation activity, could be

performed. Alternatively, CGA009 or 1003 could be complemented for growth on Fe(II) through provision of genes from TIE-1 or SW2.

Identification of genes involved in photoautotrophic Fe(II)-oxidation by Rhodopseudomonas palustris strain TIE-1

Of 12,000 mutants screened for loss of Fe(II) oxidation activity, six were identified as being specifically defective in Fe(II) oxidation, with only two different genes being represented among these mutants. Theoretically, our screen is only ~88% saturated. Thus, that five of six mutants identified contained disruptions in the same gene suggests that our screening strategy is not ideal to identify mutants defective in Fe(II) oxidation.

Nonetheless, the two mutants identified in this study provide new insight into the mechanism of Fe(II) oxidation in *Rhodopseudomonas palustris* TIE-1. Mutant strain A2 contains a disruption in a homolog of a cobalt chelatase (CobS). Because the structures of cobaltochelataes and ferrochelataes (which insert Fe(II) into porphyrin rings) are alike, it has been suggested that they have similar enzymatic activities [42, 144]. While it is possible that the phenotype of A2 might be due to the disruption of an enzyme that inserts Fe(II) into a protein or a cofactor that is involved in Fe(II)-oxidation, this seems unlikely because cobaltochelataes and ferrochelataes are typically different at the amino acid level [42]. Instead, a protein involved in Fe(II) oxidation may require cobalamin as cofactor; if true, this would represent a novel use for cobalamin [144]. Mutant strain 76H3 is disrupted in a gene that appears to encode a component of an

ABC transport system that is located in the cytoplasmic membrane. While a variety of molecules could be transported by this system, whatever is being transported (e.g., the Fe(II) oxidase or a protein required for its assembly) likely resides at least momentarily in the periplasm. This raises the question of where Fe(II) is oxidized in the cell? Because Fe(II) is known to enter the periplasmic space of gram negative bacteria through porins in the outer membrane [179], it is conceivable that Fe(II) could be oxidized in this compartment. Alternatively, the Fe(II) oxidase could reside in the outer-membrane and face the external environment, as has been inferred for Fe(II) oxidizing acidophilic bacteria [7, 190]. Determining what catalyzes Fe(II) oxidation and where it is localized is an important next step in our investigation of the molecular basis of phototrophic Fe(II) oxidation.

Identification of candidate genes involved in photoautotrophic Fe(II)-oxidation by Rhodobacter sp. SW2

Identification of the gene(s) on the insert of p9E12 that confer the observed Fe(II) oxidation phenotype proved challenging. Our attempt to identify these genes through *in vitro* mutagenesis was unsuccessful, as all cosmids that lost the ability to confer Fe(II) oxidation contained transposons in different sites in the pLAFR5 backbone rather than the SW2 DNA insert. Given the nature of the genes that these insertions disrupt (*trfA*, *tetA*, *kfrA*, *trbD/E*, and *korF/G*), it is likely that the loss of the Fe(II) oxidation phenotype results from defects in cosmid stability, replication, and/or maintenance. Why the transposon preferentially

inserts into the pLAFR5 backbone was not investigated, however, it is known that the *tetA* gene on pBR322 contains numerous “hot-spots” for Tn5 insertion [15]. Although the *tetA* gene located on pBR322 is different from that on pLAFR5 [168], these genes share 75% sequence identity (determined using Blast 2 sequences; <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>). Thus, such hot-spots likely exist within the *tetA* gene of pLAFR5 and perhaps other genes on the backbone of this cosmid.

At this point, it is still unclear which gene(s) from *Rhodobacter* SW2 on the insert of pE12 (or pB3, p11D3, and p12D4) confer the observed Fe(II) oxidation activity to *Rhodobacter capsulatus* SB1003; however, sub-cloning of the *Pst*I restriction fragment 3 in pBBR1MCS5 (Gm^R) has allowed us to confine our search to the genes present on this fragment. Sequence analysis of this 9.4 kb fragment suggests that a potential gene responsible for the Fe(II) oxidation phenotype may encode a predicted permease based on the fact it is the only ORFs in sequence common to the four cosmids for which a homolog in other *Rhodobacter* species was not found. In addition, one of the genes found to be involved in Fe(II) oxidation in TIE-1 encodes a predicted permease. Directly upstream of this permease in SW2 is a putative acetyltransferase. In TIE-1, an predicted acetyltransferase is found downstream of the permease involved in Fe(II) oxidation. The similarities here are encouraging and warrant further investigation.

In accordance with our findings that genes required for Fe(II) oxidation in TIE-1 are also present in CGA009, a strain unable to grow photoautotrophically

on Fe(II), a protein containing a β -propeller structure and possibly the redox cofactor pyrroloquinoline quinone (PQQ) with similarity to a WD-like protein in *Rhodobacter sphaeroides* 2.4.1 also represents a potential Fe(II) oxidation gene candidate. A quinoprotein has recently been implicated in the Mn(II)-oxidizing activity of *Erythrobacter* sp. SD21 and *Pseudomonas putida* MnB1 [87], providing precedent for the involvement of PQQ containing enzymes in metal oxidation reactions. In addition, a permease of the drug/metabolite transporter superfamily with a homolog in *Rhodobacter sphaeroides* 2.4.1 represents a candidate based on similarity to TIE-1 mutant 76H3 and the fact that it lies in a region common to the four cosmids.

We find that this 9.4 kb fragment confers Fe(II) oxidation activity upon 1003 independent of the orientation in which it is cloned. This suggests that the gene(s) responsible for this phenotype is cloned with its endogenous promoter. Analysis of the sequence upstream of these predicted candidate ORFs reveals putative σ^{70} promoter consensus sequences 182 bp upstream of the predicted start of translation (T_L start) for the putative PQQ containing protein, 1427 bp upstream from the T_L start for the predicted permease with no homolog in *Rhodobacter*, and 198 bp upstream of the T_L start for the predicted permease of the drug/metabolite transporter superfamily. For comparison, in genes of *E.coli* K12, the distance between the transcription start site associated with the promoter and the translation start site is between 0–920 bp with 95% of 771 promoters analyzed being at a distance <325 bp upstream of the T_L start site [28].

The promoter 1427 bp upstream from the T_L start for the predicted permease with no homolog in *Rhodobacter* is well out of this range, however, it is possible that this permease is co-transcribed with the upstream acetyltransferase. This predicted promoter lies 655 bp upstream of the acetyltransferase and while this distance is above average by comparison to *E. coli*, it is within range.

To identify the specific genes responsible for the observed Fe(II) oxidation phenotype, we may take two approaches which include: 1) cloning these candidates to test them specifically for their ability to confer Fe(II) oxidation activity to 1003; 2) making directed knock-outs of these candidates using the method of Datsenko and Wanner [44]. In this method, particular genes on a chromosome or a construct can be replaced with an antibiotic resistance gene (generated by PCR and designed to have a 36 nucleotide extension with homology to the gene(s) of interest) using the phage λ Red recombinase to promote recombination. In anticipation of the possibility that the genes involved are not present in our current sequence, we are also closing the sequence gaps of the pP3 insert. Lastly, given that the pP3-gm1 and pP3-gm2 clones contain a smaller fragment of the p9E12 insert and are cloned in the BBR1MCS5 vector (which confers gentamicin resistance) rather than pLAFR5, if our direct approaches do not work, we will again attempt to use an *in vitro* mutagenesis approach to identify these genes.

Once these genes are identified we will be able to address whether the Fe(II) oxidation system in these phototrophs is similar to that in other Fe(II)-oxidizing bacteria such as *A. ferrooxidans*. Uncovering the degree to which

electron transfer from Fe(II) is conserved amongst phylogenetically divergent species may in turn provide information on the origins of this ancient metabolism.