Winter 2007

The identification, characterization, and expression of truncated hemoglobin genes in the nitrogen-fixing actinorhizal symbiont Frankia

James Michael Niemann
University of New Hampshire, Durham

THE IDENTIFICATION, CHARACTERIZATION, AND EXPRESSION OF TRUNCATED HEMOGLOBIN GENES IN THE NITROGEN-FIXING ACTINORHIZAL SYMBIONT *FRANKIA*

BY

JAMES MICHAEL NIEMANN

Bachelor of Science, University of New Hampshire, 2001

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

in

Microbiology

December, 2007

This thesis has been examined and approved.

Thesis Director, Dr. Louis Tisa, Associate Professor of Microbiology

Dr. Robert Zsigray, Professor Emeritus of Microbiology

Dr, Steven Torosian, Affiliate Assistant Professor of Microbiology

12/13/07

Date

DEDICATION

This thesis and all of the work accomplished during this research could not have been possible without the continued love and support of my parents and my wife. I am forever indebted to their never-ending belief in me.

ACKNOWLEDGEMENTS

I would like to thank Bob Mooney for all of his technical support during this research. I could not have asked for better thesis committee members than Dr. Robert Zsigray and Dr. Steven Torosian. All of their input was greatly valued. Thanks to Dr. John Tjepkema whose feedback as our collaborator on some of this research was of great help. I would also like to acknowledge all of the members of the Tisa Lab, especially Erik Janicki and Dr. Brandye Michaels for all of there constructive input during my graduate research.

Last, but certainly not least, I want to thank my graduate advisor Dr. Louis Tisa for his knowledge and guidance that helped shape me to think and act as a scientist, but most of all for his patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

THE IDENTIFICATION, CHARACTERIZATION, AND EXPRESSION OF TRUNCATED HEMOGLOBIN GENES IN THE NITROGEN-FIXING ACTINORHIZAL SYMBIONT *FRANKIA*

by

James Michael Niemann

University of New Hampshire, December, 2007

Using a molecular approach (PCR), a group II truncated hemoglobin (trHb) gene was identified in several diverse isolates of *Frankia*. An analysis of three draft genome sequences for *Frankia* isolates EAN1pec, Ccl3, and ACN14a also revealed the presence of second trHb, homologous to group I trHbs. Phylogenetic analysis suggested that the *Frankia* trHb genes were grouped based on their respective genotype and clustered closest to *Mycobacterium* trHb genes.

Frankia strain Ccl3 was grown under a variety of environmental stimuli to evaluate the expression of trHbN and trHbO genes. Nitrogen status did not affect expression of either gene, while oxidative stress caused a decrease in expression levels for both genes. The expression of trHbO increased under low oxygen environments, suggesting a role in increasing respiration rates. The expression of trHbN increased in response to spontaneously generated nitric oxide, suggesting a role in the protection from reactive nitrogen species.

CHAPTER I

INTRODUCTION

General Overview of *Frankia*

Frankia, a member of the order Actinomycetales, is the nitrogen-fixing bacterial partner of actinorhizal symbiosis. This gram positive, filamentous microbe is capable of fixing nitrogen in association with over 200 species of woody, dicotyledonous shrubs and trees comprising 25 genera of 8 plant families. There are several reviews that describe this plant-microbe interaction in much greater detail (Benson & Sylvester, 1993; Huss-Danell, 1997; Wall, 2000; Schwencke & Caru, 2001; Vessey et al., 2005).

As a contributor to the global biological nitrogen budget, the *Frankia*-actinorhizal symbiosis rivals the *Rhizobium*-legume symbiosis and is responsible for ~50% of the terrestrially fixed nitrogen. Unlike *Rhizobium*, which are dependent on its host to fix nitrogen or require reduced partial pressures of oxygen, *Frankia* species are capable of fixing nitrogen in a free-living state under atmosphic partial pressures of oxygen. Another point of contrast is that most *Frankia* sp. have broad host-recognition specificity whereas *Rhizobium* sp. have a more narrow host-recognition specificity. In fact, some *Frankia* strains are capable of infecting a variety of plants from different families (Benson & Sylvester, 1993; Huss-Danell, 1997; Wall, 2000; Schwencke & Caru, 2001;

Vessey et al., 2005). In either case, the nature of the relationship is the same: the microbe provides the plant with a source of fixed nitrogen and the plant provides a source of a carbon and energy.

The ecological range of actinorhizal plants is quite vast and they can be found on every continent with exception to Antarctica. These perennial, dicotyledonous angiosperms can be found inhabiting arctic tundra, glacial tills, alpines, temperate forests, wetlands, riparian zones, dry chaparral and xeric shrub lands, coastal dunes, and the tropics. The ecological and economical benefits of these important trees and shrubs are: soil restoration and land reclamation, biomass for pulp and timber, windbreaks, nurse cropping, fuelwood, use as ornamental and horticultural plants, and in some cases as a food source (Benson & Sylvester, 1993; Huss-Danell, 1997; Wall, 2000; Schwencke & Caru, 2001; Richards et al., 2002; Vessey et al., 2005). Recently, the genomes of three phylogenetically distinct *Frankia* strains (Ccl3, ACN14a, and EAN1pec) were sequenced (Normand et al., 2007ab). These genome sequence databases have generated a pipeline of information about this symbiosis, and raised many more questions about *Frankia* sp. and their relationship with actinorhizal plants.

Frankia is capable of developing two unique morphological characteristics: spores and vesicles (Benson & Sylvester, 1993; Huss-Danell, 1997; Wall, 2000; Schwencke & Caru, 2001; Vessey et al., 2005). Spores are produced in mulitlocular sporangia that stem from the hyphae terminally or intercalary. Each sporangium may contain hundreds of spores. Upon their release from the sporangia the spores may germinate when environmental conditions are

favorable. Vesicles are expressed during nitrogen limiting conditions and are functionally analogous to heterocysts formed by nitrogen-fixing cyanobacteria. These specialized structures function to compartmentalize nitrogenase, the enzyme responsible for nitrogen fixation, and serve as the localized site of nitrogen fixation. Several lines of evidence indicate that the laminate layer of hopanoid lipids surrounding the vesicle structure prevents the diffusion of oxygen, protecting the oxygen-labile enzyme. As oxygen levels increase, the layers of lipid have been shown to proportionally thicken as to restrict further diffusion of oxygen (Berry et al., 2003).

General Overview of Truncated Hemoglobins

Truncated hemoglobins (trHbs) are small heme proteins within the hemoglobin (Hb) superfamily that includes bacterial flavohemoglobin, *Vitreoscilla* hemoglobin, (non-) vertebrate hemoglobin and myoglobin (Mb), and plant (non-) symbiotic hemoglobin (Pesce et al., 2000; Milani et al., 2001a; Wittenberg et al., 2002; Sarma et al., 2005). TrHbs are distributed among eubacteria, cyanobacteria, protozoans, and higher plants. As of yet, there have been no trHbs discovered in archaea or metazoa. A number of trHbs have been identified and studied in various organisms including the following: the actinobacteria *Mycobacterium* sp. (Couture et al., 1999a; Milani et al., 2001ab; Ouellet, et al., 2002; Pathania et al., 2002ab; Milani et al., 2003; Liu et al., 2004; Milani et al., 2004), and *Thermobifida fusca* (Bonamore et al., 2005); the eubacteria *Bacillus subtilis* (Choudhary et al., 2005), *Campylobacter jejuni* (Wainwright et al., 2005),

Nostoc sp. (Hill et al., 1996), and *Synechocystis* sp. (Couture et al., 2000; Scott & Lecomte, 2000); the unicellular eukaryotes *Paramecium caudatum* (Das et al., 2001), *Tetrahymena pyriformis* (Korenaga et al., 2000) and *Chlamydomonas eugametos* (Couture et al., 1999b), and in the plants *Arabidopsis thaliana* (Watts et al., 2001), and *Medicago truncatula* (Vieweg et al., 2005). An analysis of sequenced genomes has also revealed the presence of these proteins in several other organisms (Wittenberg et al., 2002; Vuletich & Lecomte, 2006).

These trHbs have very little homology with other Hbs in the superfamily. The genes have undergone severe deletions resulting in a protein that is typically 20-40 amino acids shorter than other globins (Pesce et al., 2000; Milani et al., 2001a; Wittenberg et al., 2002; Sarma et al., 2005; Vuletich & Lecomte, 2006). Unlike the traditional three-on-three alpha helical configuration, trHbs display a novel two-on-two alpha helical sandwich that still conserves the classical globin fold (Figure 1). This two-on-two arrangement is composed of four main helices (B, E, G, and H) that form two anti-parallel helix pairs (B/E and G/H). The terminal end of the A helix is almost completely deleted, the CD-D region of the protein has been reduced to three residues, and the C-terminal H helix has lost about 10 amino acids. The F helix has been replaced with a "Pre-F extended loop" containing a single turn that houses the heme-coordinating F8 histidine that is conserved in all known hemoglobin.

In order to maintain the traditional globin fold in trHbs, three conserved glycine motifs are present to overcome the steric hindrance introduced from these extreme amino acid deletions. These glycine residues are located at the

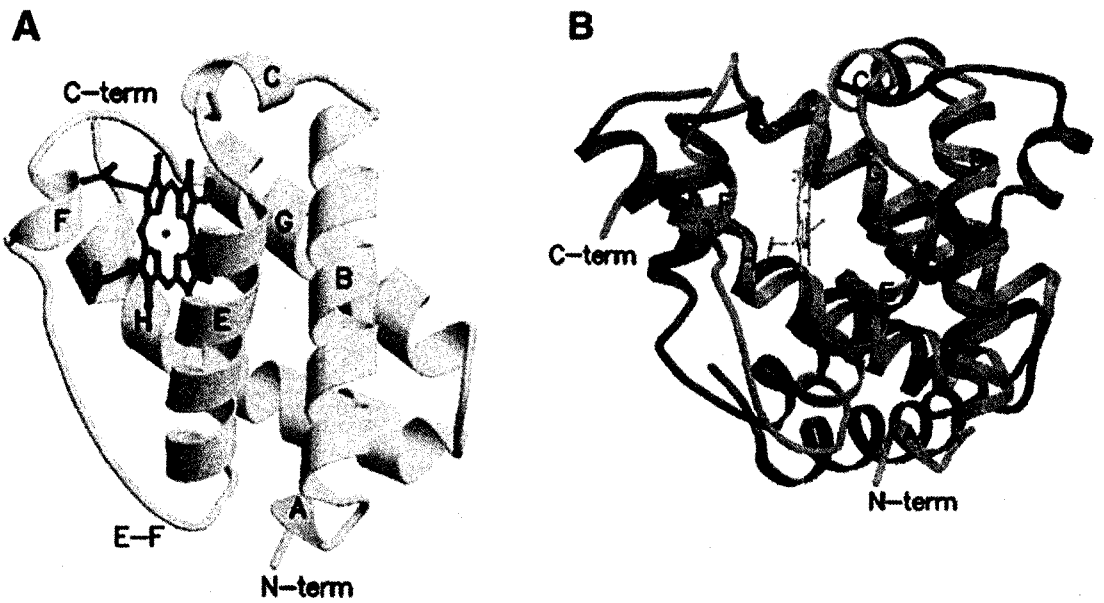


Figure 1. (A) A ribbon stereo view of *Paramecium caudatum* trHb tertiary structure, including the heme group. (B) A stereo view of the structural overlay of *Chlamydomonas eugametos* trHb (green) and sperm whale Mb (red). Images taken from Pesce et al., 2000.

interhelical regions of the AB and EF helices and before the single turn in the Pre-F loop. Another residue almost completely conserved among all trHbs is a B10 tyrosine that aids in the stabilization of heme-bound ligand (Pesce et al., 2000; Milani et al., 2001a; Wittenberg et al., 2002; Sarma et al., 2005; Vuletich & Lecomte, 2006). Together with B10 residue, the distal residues CD1, E7, E11, E15, and G8 are instrumental in the stabilization of heme interactions with ligand. Although the residues in these positions are not as conserved as the nearly universal B10Y, their distinct arrangements have evolved in certain bacterial groups, suggesting functional adaptation (Wittenberg et al., 2002; Milani et al., 2005).

There are three distinct groups of truncated hemoglobins that have been designated group I (trHbN), group II (trHbO), and group III (trHbP). Within the group I trHbs are two distinct subgroups that branch phylogentically. Amino acid homology among trHb paralogs within the same species tends to be low, whereas the homology among trHb orthologs in different species is rather high. Many of the organisms studied thus far have been found to contain different numbers of trHb paralogs. For example, the actinobacteria of the *Mycobacterium* genus display several layers of complexity. *Mycobacterium avium*, an opportunistic pathogen, possesses all three trHb types (N, O, & P); *Mycobacterium tuberculosis*, a facultative pathogen, has only two trHb types (N & O); and the obligate intracellular pathogen, *Mycobacterium leprae*, has only one trHb group (O) present in its genome (Wittenberg et al., 2002; Milani et al., 2005; Sarma et al., 2005). This trimming of paralog numbers and types within this

genus correlates with their respective genome sizes, and is evidence for an effect of reductive evolution. The presence of gene paralogs suggests potential separate function for each trHb. Phylogenetic analysis of 111 trHb genes suggests that trHbO was the original gene present in the last common ancestor of actinobacteria and proteobacteria and that trHbN and trHbP genes arose through gene duplications and horizontal transfer events (Vuletich & Lecomte, 2006). This hypothesis would also explain the presence of trHbO in every *Mycobacterium* genome and other organisms possessing trHb genes. Interestingly, Vuletich & Lecomte (2006) proposed that the trHb globin fold existed prior to the vertebrate globin fold.

Group I trHbs (trHbN) share the least amount of homology within their respective subgroup compared to trHbO and trHbP. Unlike the other trHbs groups, group I trHbs share a unique structural feature. The four main helices of these hemoglobin proteins create a protein tunnel lined with hydrophobic residues. This tunnel connects the molecular surface of the protein with the heme distal site to essentially create a path for ligand diffusion, storage, and/or multi-ligand reactions (Pesce et al., 2000; Milani et al., 2001ab; Wittenberg et al., 2002; Milani et al., 2005; Vuletich & Lecomte, 2006). This structure has not been observed before in (non-) vertebrate Hbs.

Most research on trHbN has focused on the *Mycobacterium* sp. (Couture et al., 1999; Milani et al., 2001; Ouellett et al., 2002; Pathania et al., 2002b; and Milani et al., 2004). The *Mycobacterium* sp. *glbN* gene product (HbN) has been proposed to have catalytic function role as a oxygen-dependent nitric oxide (NO)

dioxygenase to protect the microbe from reactive NO. During the intracellular infection of macrophages by *Mycobacterium tuberculosis* in lung tissue, the infected cells are surrounded by other macrophages and bombarded with reactive nitric oxide to kill the microbe. The heme of HbN stores an oxygen molecule, which reacts with nitric oxide (NO) to form harmless nitrate (Ouellett et al., 2002; and Pathania et al., 2002b).

In *Nostoc* sp., trHbN is postulated to function differently and may be involved in the process of nitrogen fixation (Hill et al., 1996). This hypothesis is supported by two lines of evidence: up-regulation of the gene when *Nostoc* is grown without a combined nitrogen source, and the location of *glbN* within the intergenic region of two nitrogen fixing genes, *nifU* and *nifH.* Other trHbN hemoglobins from organisms such as *Tetrahymena* (Korenaga et al., 2000), and *Synechocystis* (Couture et al., 2000; Scott & Lecomte, 2000) have been evaluated, but only for structural studies. Their functions remain to be determined or predicted.

Group II trHbs (trHbO) share more homology within their subgroup compared to their group I orthologs. These trHb proteins lack the hydrophobic ligand tunnel present in trHbN, but possess a shallow depression on the proximal side of the heme that is believed to act as a docking site for a ligand or the reaction of multi-ligands (Milani et al., 2005; Vuletich & Lecomte, 2006).

Similar to the trHbN paralogs, group II trHbs have also been extensively studied in *Mycobacterium* sp. (Pathania et al., 2002a; Milani et al., 2003; Liu et al., 2004, Visca et al., 2002ab; and Ascenzi et al., 2006). The trHbO protein is

postulated to help increase *M. tuberculosis* respiration rates under hypoxic intracellular environments aiding its persistence in the lung where the microbe competes with lung tissue cells for oxygen. TrHbO is hypothesized to increase respiration by delivering oxygen to terminal cytochrome oxidases. This hypothesis is supported by the observation that these hemoglobins are localized at the cell membrane where oxygen delivery to these respiratory enzymes would be expected (Pathania et al., 2002a). The proposed function of *M. leprae* trHbO is analogous to that of *M. tuberculosis* trHbN to help protect the microbe from reactive nitrogen species produced from the host. In contrast with trHbN that acts as a catalyst to react oxygen with NO, the oxygenated form of trHbO is speculated to react directly with peroxinitrite ($ONOO^-$), the toxic intermediate of NO reactions with superoxide radicals (Visca et al., 2002ab; and Ascenzi et al., 2006). Other trHbO hemoglobins from organisms including plants (Watts et al., 2001), *Thermobifida fusca* (Bonamore et al., 2005), *Bacillus subtilis* (Choudhary et al., 2005), and *Campylobacter jejuni* (Wainwright et al., 2005) have been evaluated primarily for structural characterization studies.

Among the subgroups, group III trHbs (trHbP) are the most highly conserved. Phylogenetically, these trHbs resemble the group II trHbs more closely than group I (Vuletich & Lecomte, 2006). Unlike their paralogs, trHbP lack the conserved glycine motifs important for maintaining the globin fold. Instead it is believed that they maintain their structural integrity through other interhelical interactions. Since no members of the group III trHbs have been characterized to date, little is known about their function.

Hemoglobin in *Frankia*

Hemoglobin has been identified and/or isolated previously from the root nodules of various actinorhizal plants (Tjepkema, 1982; Tjepkema & Asa, 1987; Kortt et al., 1988; Jacobsen-Lyon et al., 1995; Pathirana & Tjepkema, 1995; Suharjo & Tjepkema, 1995; Sasakura et al., 2006). However, there had been speculation as to whether or not the hemoglobin isolated was being produced by the microbial symbiont or the host plant. Hemoglobins of similar molecular weight to trHbs were found associated with nodule fractions of *Alnus glutinosa* (Suharjo et al., 1995) and *Myrica gale* (Pathirana & Tjepkema, 1995) that were infected with *Frankia* hyphae. Hemoglobin eluted from *Myrica gale* nodule fractions corresponded to a molecular mass of 38.5 kDa and a smaller amount of 16.7 kDa (Pathirana & Tjepkema, 1995). It was thought that the larger elution was a Hb dimer produced by the host plant and the smaller of the two came from *Frankia*. This led to subsequent studies aimed at isolating Hb from growing cultures of *Frankia*.

The isolation of hemoglobin from *in vitro* cultures of several genetically diverse *Frankia* isolates indicated that the microbes were capable of producing hemoglobin (Tjepkema et al., 2002; Beckwith et al., 2002). Several lines of biochemical evidence support the hypothesis that *Frankia* hemoglobins are trHbs. First, the molecular mass of the hemoglobins isolated from strain Ccl3 and EAN1pec are 14.1 and 13.4 kDa, respectively, which is consistent with known molecular mass of trHbs. The oxygen dissociation rates for the Ccl3 and EAN1pec hemoglobins are 56 s^{-1} and 131 s^{-1}, respectively. These values are

high when compared to other hemoglobins, suggesting a role well suited for diffusion of oxygen over short distances.

Initial studies to determine the function of *Frankia* Hb showed no significant differences in hemoglobin production for cells grown in media with or without a combined nitrogen source (Beckwith et al., 2002). This result suggests that *Frankia* hemoglobin does not function in nitrogen fixation. Schwintzer and Tjepkema (2005) evaluated the production of *Frankia* strain Arl3 hemoglobin with respect to hypoxic (1%), ambient (20%), and hyperoxic (40%) oxygen concentrations in growth media with (N+) or without (N-) a combined nitrogen source. For N+ cultures, hemoglobin production was the greatest under hypoxic conditions, suggesting a possible role as an oxygen shuttle to terminal oxidases during respiration. Hemoglobin production under each of the oxygen conditions is lower for N- grown cultures compared to N+ cultures. These data support the idea that *Frankia* hemoglobin does not act to protect oxygen-labile nitrogenase.

Research Goals

The overall purpose of this study was to identify and characterize truncated hemoglobin genes in *Frankia* and to evaluate their expression under various environmental stimuli. This thesis began before the completion of the *Frankia* genome sequence projects and was initially focused on the hypothesis of a single hemoglobin gene. The original goal was to clone and sequence this gene. During the course of this work, the completed *Frankia* genomes revealed the existence of two hemoglobin paralogs. The first section of this thesis covers

the molecular approaches used to sequence trHbO genes in several genetically diverse *Frankia* strains. These data along with gene sequences from other genomes were used to access the phylogeny of *Frankia* trHbN and trHbO genes with known homologues of other actinomycetes. The analysis of promoter regions and organization of these genes in their respective genomes was also evaluated to further characterize trHbN and trHbO in *Frankia*. The last section of this thesis focused on the use of RT-PCR and Real-time RT-PCR to evaluate the relative expression of trHbN and trHbO under different growth conditions to assess functionality. I hypothesized that the *Frankia* trHbO gene (*glbO*) is expressed under hypoxic conditions and may act as an oxygen shuttle for terminal respiratory oxidases. This role would allow *Frankia* to persist in nodules or other environments where oxygen levels are lower than ambient oxygen. I also postulated that the *Frankia* trHbN gene (*glb*N) plays a role in nitric oxide detoxification under nitrosative stress conditions. This function would be useful in protecting *Frankia* from nitric oxide produced during oxidative and nitrosative bursts by the host plant during infection. Based on previous data, I did not expect a role for either of these hemoglobins in nitrogen fixation.

CHAPTER II

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

All of the *Frankia* strains used in this study and their respective growth incubation conditions are listed in Table 1. Cultures were grown and maintained in basal medium under nitrogen-fixing repressed conditions with the addition of 5 mM NH_4Cl as a nitrogen source, as described previously (Tisa et al., 1999). The basal medium consisted of MOPS-phosphate buffer (50 mM MOPS, 10 mM K_2HPO_4, pH 6.8) containing Metals Mix and a carbon source. Metals Mix contained the following components: 1 mM Na_2MoO_4, 2 mM $MgSO_4$, 20 μM $FeCl_3$ with 100 μM nitrilotriacetic acid (NTA), and modified trace salts solution (Tisa et al. 1983). *Streptomyces coelicolor* NRRL B-16638 and *Escherichia coli* HB101 were grown in YEME (Hopwood et al., 1985) at 28°C and Luria-Bertani Broth (1% tryptone, 0.5% yeast extract, and 1% NaCl) at 37°C, respectively.

Gene Sequences Used

The following trHb sequences (with their respective accession numbers) were used in this study: *S. coelicolor* trHbO (CAB71209); *Streptomyces avertimilis* trHbO (BAC73082); *Mycobacterium bovis* trHbN (CAD96236), trHbO (CAD97358); *Mycobacterium tuberculosis* trHbN (CAA98320), trHbO

Table 1. *Frankia* strains used in this study.

Frankia Strains	Host Plant	Carbon Source Used	Incubation Temp. (°C)	Phylogenetic Group	Reference
ACN1^{AG}	*Alnus viridis ssp. crispa*	Succinate	30	Group I	Lalonde *et al*, 1981
CcI1.7	*Colletia cruciata*	Propionate	28	Group III	Meesters *et al*, 1985
CcI3	*Casuarina cunninghamiana*	Propionate	30	Group I	Zhang *et al*, 1984
CN3	*Coriaria nepalensis*	Succinate	30	Atypical	Mirza *et al*, 1994
CpI1 propionate variant	*Comptonia peregrina*	Propionate	30	Group I	Callaham *et al*, 1978; Tisa *et al*, 1983
EAN1pec	*Elaeagnus augustifolia*	Fructose	25	Group III	Lalonde *et al*, 1981
EuI1c	*Elaeagnus umbellata*	Glucose	30	Group III	Baker *et al*, 1980
EUN1f	*Elaeagnus umbellata*	Propionate	30	Group III	Lalonde *et al*, 1981

(CAA16047); *Mycobacterium avium* subsp. *paratuberculosis* trHbN (AAS03570), trHbO (AAS04608), trHbP (AAS05726); *Mycobacterium leprae* trHbO (CAC31634); *Corynebacterium glutamicum* trHbO (CAF21110); *Corynebacterium efficiens* trHbO (BAC19155); *Corynebacterium diptherae* trHbO (CAE50330); Sperm whale (*Physeter catodon*) Mb (P02185). TrHbO sequences from *Arthrobacter sp.*, *Brevibacterium linens*, *Kinecoccus radiotolerans*, *Leifsonia xyli*, *Nocardia farcinica*, and *Thermobifia fusca* were all obtained using the Integrated Microbial Genomes System from the Joint Genome Institute (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). Preliminary sequence data for *Frankia* strains EAN1pec and Ccl3 trHbO and trHbN were obtained from D. Benson and L. Tisa. Preliminary sequence data for strain ACN14a trHbO and trHbN were obtained from P. Normand.

Polymerase Chain Reaction (PCR)

Genomic DNA (gDNA) from *Frankia* was isolated using the CTAB method (Wilson, 1989). All of the primer sets used in this study are listed in Table 2. A 202 bp trHb amplicon was amplified by PCR with the primer set TrHb-F (5'-GTCGGCGGGGAGGAGACCTTC-3') and TrHb-R (5'-CGTGCCGCATCCGC AGCCGCGG-3') and 250ng of template DNA. The PCR was performed in 50 µl reaction volumes with 0.5 µM of each primer using the Failsafe™ PCR System (Epicentre Tech.) according to the manufacturer's recommendations. The TrHb primers were designed from *S. coelicolor* (bp positions 61 to 81 and positions 241 to 262) within conserved regions of the trHbO gene between *S. coelicolor*,

Table 2. Primer sets used in this study for PCR, RT-PCR, and qRT-PCR.

Primer	Primer Sequence	Target	Product Size	Reference
TrHb-F	5'-GTC GGC GGG GAG GAG ACC TTC-3'	*HbO* gene	202 bp	Niemann *et al.*, 2005
TrHb-R	5'-CGT GCC GCA TCC GCA GCC GCG G-3'			
HbNcci92	5′-CAC CCC TCT TTG CCA ACC-3′	trHbN mRNA	127 bp	This study
HbNcci219	5′-CCT CAC CGA CGC CCA CTT-3′			
HbOcci299	5′-GGG ACG CCT GGC TGA AGA-3′	trHbO mRNA	76 bp	This study
HbOcci375	5′-CCA GAG CTG CCT GTC GAG ATC-3′			
DB41	5′-TTC TTC ATC CAC GAC CCG-3′	*glnA* gene	477 bp	Clawson et al., 2004
DB44	5′-GGC TTC GGC ATG AAG GT-3′			

M. tuberculosis, *M. leprae*, and *C. glutamicum* identified using ClustalX cluster analysis software (Thompson et al., 1997). Thermocycling parameters were as follows: i) initial denaturation step at 95°C for 2 min; ii) 35 cycles of denaturization at 95°C for 1 min, primer annealing at 61°C for 45 sec, and primer extension at 72°C for 1 min; and iii) a final extension step at 72°C for 5 min. The amplicons were resolved using agarose gel electrophoresis in a submarine gel electrophoresis unit (Hoefer Scientific Instruments, San Francisco, CA) as described by Sambrook et al. (1989). The agarose gel was cast with 2% agarose in a 1X Tris-Borate-EDTA (TBE) electrophoresis buffer (89mM Tris base; 89mM Boric Acid; 50mM EDTA, pH 8.0) (Sambrook et al., 1989). The electrophoresis was performed in this same buffer.

DNA Sequencing and Phylogenetic Analyses

PCR products were purified with a Qiaquick® PCR Purification Kit (Qiagen), according to the manufacturer's recommendations, and used as a template for DNA sequencing reactions with a DYEnamic ET terminator cycle sequencing Kit (Amershan Pharmacia Biotech) in an ABI PRISM 377 sequencer (Perkin Elmer). Both the TrHb-F and TrHb-R primers were used for these sequencing reactions. Sequence analyses were performed using the SeqEd program version 1.0.3 (Applied Biosystems). The sequences were compared to those available from the GenBank and EMBL database using the BLAST program (Altschul et al., 1997). Partial sequences of the *Frankia* trHb amplicons have been deposited in GenBank under the following accession numbers:

AY768545 (EUN1f), AY768546 (EAN1pec), AY768547 (Eul1c), and AY768548 (CN3). Frameplot 2.3.2 (Ishikawa and Hotta, 1999) was used to establish an open reading frame for each sequence. Hemoglobin sequences were aligned using ClustalX (Thompson et al., 1997). For phylogenetic analysis, neighbor-joining trees were constructed from 1000 bootstrap replicates using PAUP 4.0b10 (Swofford, 2003).

Sequence Analyses: Gene Neighborhoods and Promoter Analyses

The gene neighborhoods of *Frankia* trHbs and of known trHb genes in other actinomycetes were compared using the gene neighborhood program in the Integrated Microbial Genome database through the Joint Genome Institute (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). Putative promoters and ribosomal binding sites (RBS) of Frankia trHbs were identified by manually searching the upstream gene regions available from the preliminary sequence data from EAN1pec, Ccl3, and ACN14a draft genomes.

RNA Extraction Methods

Prior to any RNA work, all reagents and glassware were directly treated with or prepared from 0.1% diethylpyrocarbonate (DEPC) water. Reagents were treated for at least 1 h followed by autoclaving to inactivate the DEPC.

Two methods of total RNA extraction, RNAwiz (Ambion, Foster City, CA), a commercially available reagent and the Triton X-100 boiling method (Sung et

al., 2003), were tested and compared to evaluate which method would provide the highest quality and yield of RNA.

For these initial experiments, 400 mg of frozen EAN1pec hyphae stored at -80° C were thawed in 1 mL of Tris-EDTA (TE) buffer (10 mM Tris-HCl, pH 7.4; 1 mM EDTA, pH 8.0) and fragmented with a tissue homogenizer. The fragmented hyphae were collected by centrifugation at 12,000 x g for 15 min. After the supernatant was removed, the pellet was resuspended in 1ml of Tris-EDTA buffer. This wash step was repeated two more times. The final pellet was resuspended in 600uL of TE buffer and the cell suspension was aliquoted into two separate tubes of equal volume (300uL containing 200 mg of cell mass each) to be used for each RNA isolation method.

RNAwiz® extraction method. The manufacturer's basic protocol was followed to extract total RNA from *Frankia* EAN1pec, and an enzymatic digestion was utilized to aid in cell lysis. The cell suspension was incubated with lysozyme (Sigma) at 1 mg/mL (final concentration) at 37°C. After 15 min, 800 ul RNAwiz® reagent was added to the suspension, which was mixed with a vortex mixer. After 5 min incubation at room temperature, chloroform (0.2X starting volume) was added to the homogenate. The tube was repeatedly inverted for 20 sec and incubated at room temperature for 10 min. The tube was centrifuged at 10,000 x g for 15 min at 4°C to separate the phases. The aqueous phase was transferred to a fresh tube and 0.5 X starting volume of DEPC-treated water was added to it. After the tube was mixed well, one starting volume of isopropanol was added and mixed to precipitate the RNA. Following a 10 min incubation at room

temperature, the tube was centrifuged at 10,000 x g in a microfuge for 15 min at 4°C to recover the RNA. The RNA pellet was washed with 1ml of cold 75% ethanol. The RNA pellet was centrifuged at 10,000 x g for 15 min at 4°C. After the ethanol was decanted, the pellet was air dried for 15 min. The RNA pellet was resuspended in 200 ul of DEPC-treated water and stored at -80°C.

Triton X-100 Boiling Method. The TE-washed cells (200ug) were concentrated by centrifugation at 10,000 x g for 15 min and suspended in 1ml TE buffer (pH 7.5) containing 0.2% Triton X-100. The cell suspension was incubated for 10 min at 100°C. The tops of the microfuge tubes were taped down to prevent them from popping open. After the incubation, the tubes were transferred to an ice bath to cool. An equal volume of chloroform was added and the tubes were inverted several times. The tube was centrifuged at 13,000 rpm in a microfuge for 10 min at 4°C to separate the phases. The aqueous phase was transferred to a fresh microfuge tube and the chloroform extraction step was repeated two more times with the aqueous phase. The RNA from the aqueous phase was precipitated by adding $1/10^{th}$ volume of 3 M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol, and incubation at -20°C overnight. The RNA was recovered by centrifugation at 10,000 x g in a microfuge for 15 min at 4°C. The RNA pellet was washed twice with 1.5 mL of cold 75% ethanol by centrifugation at 10,000 x g in a microfuge for 15 min at 4°C, which was followed by a final wash with 1 mL of cold absolute ethanol. After centrifugation at 10,000 x g for 15 min at 4°C, the ethanol was removed and pellet was air dried for 15

min. The RNA was suspended in 200ul of DEPC-treated water and stored at -80°C.

For the initial RNA experiments, 50ul aliquots from each RNA sample were treated with DNaseI using a DNA-free™ kit (Ambion), RNase A (Qiagen), or both. Five microliters of 10X DNase buffer was added to each tube. Two units of DNaseI and/or 7 U of RNase A were added to the RNA samples and incubated at 37°C for 30 min. These RNA samples were used for the determination of the RNA quality and quantity (see below).

Determination of RNA Quality and Quantity

RNA was quantified by measuring the absorbance at 260 nm using a Beckman DU 640 spectrophotometer (Beckman Instruments, Inc.). The quality of RNA was tested by separating 1 ug RNA on a 1.2% (w/v) agarose gel in MOPS electrophoresis buffer (20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA, pH 7.0) containing 0.66 M formaldehyde to denature the RNA. One microgram of RNA samples suspended in 6X Loading buffer were used to load the gel. Gels were run in 1X MOPS buffer at 70V and stained in ethidium bromide to visualize strong 5S, 16S, and 23S rRNA bands.

Environmental Test Conditions

Unless otherwise mentioned, *Frankia* strain CcI3 cultures were incubated under the following conditions to observe their effect on gene expression. For these experiments 1-2 week old static cultures grown at 28°C were used. The

cells were washed three times with MOPS-phosphate buffer by centrifugation at 10,000 x g prior to inoculation. To test the effect of combined nitrogen on TrHb production, *Frankia* strain Ccl3 cultures were incubated for 6 days at 28°C in propionate growth medium lacking a combined nitrogen source or in propionate growth medium containing NH_4Cl as nitrogen source. To test the effect of oxygen, cultures were grown under aerated conditions and under static culture conditions for 7 days. For the aerated culture, a constant stream of sterile air was bubbled through the culture by the use of a sparger and an aquarium pump. For oxidative stress conditions, cultures were incubated for 4 days in propionate-NH_4Cl growth medium containing 0.1mM paraquat, 0.1 mM H_2O_2, or no additions (control cultures). For nitrosive stress conditions, cultures were incubated for 3 h in propionate-NH_4Cl growth medium containing either 400 μM S-nitroso-N-acetylpenicillamine (SNAP) (Calbiochem), and/or 400 μM carboxy PT10 (Cayman). SNAP spontaneously generates NO and carboxyl PT10 is a NO chelator.

One-Step Reverse Transcriptase PCR (RT-PCR)

Total RNA was isolated from *Frankia* strains cultures under the test conditions using the Triton X-100 method described above. The final recovered RNA pellet was resuspended in 200 uL of DEPC treated water. DNA was removed from RNA samples with DNase treatment using RNase-free DNaseI (New England Biolabs). RNA samples were mixed with 5 U of DNaseI, 1/10th volume of 10X DNase buffer, 200 U of RNaseOUT (Invitrogen) and the reaction

mixture was brought to a final volume of 100 uL with DEPC-treated water. The samples were incubated at 37°C for 30 min. To stop the reaction, 1 uL of 0.5 M EDTA was added to each reaction and the nuclease was heat inactivated by incubating at 75°C for 10 min. RNA samples were stored at -80°C until they were used for RT-PCR reactions.

RT-PCR was performed using a Titan One-Tube RT-PCR System (Roche) according to the manufacturer's recommendations. The reactions were performed in 50 µl volumes with 200 ng of RNA template and 0.4 µM of each primer. For amplification of TrHbN mRNA, the primers, HbNcci92 (5′-CACCCCTCTTTGCCAACC-3′) and HbNcci219 (5′-CCTCACCGACGCCCACTT-3′) were used, while the primers HbOcci299 (5′-GGGACGCCTGGCTGAAGA-3′) and HbOcci375 (5′-CCAGAGCTGCCTGTCGAGATC-3′) were used for amplification of TrHbO mRNA. To avoid non-specific priming from their respective paralogs these trHb primers were chosen to bind to non-homologous regions of the TrHbN and TrHbO genes. These regions were identified using ClustalX cluster analysis software. For the amplification of *glnA* mRNA, the primers DB41 (5′-TTCTTCATCCACGACCCG-3′) and DB44 (5′-GGCTTCGGCAT GAAGGT-3′) were used (Clawson et al., 2004).

Thermocycling parameters for trHbN and trHbO were as follows: i) reverse transcription at 55°C for 30 min; ii) initial denaturing at 94°C for 2 min; iii) 9 cycles of denaturing at 94°C for 20 sec, primer annealing at 55°C for 30 sec, and primer extension at 68°C for 45 sec; iv) 24 cycles of denaturing at 94°C for 20 sec, primer annealing at 55°C for 30 sec, and primer extension at 68°C for 50 sec with

an additional 5 sec added to each progressive cycle; and v) a final extension step at 68°C for 4 min. Thermocycling parameters for GlnA were the same as described for trHbN and trHbO, except that the annealing temperature was 51°C. The amplicons were resolved by gel electrophoresis in 2% agarose matrix in 1X Tris-Borate-EDTA (TBE) electrophoresis buffer (pH 8.0) according to Sambrook et al. (1989). The respective band intensities were quantified using Quantity One® software (BIORAD).

Preparation of cDNA for Real-Time RT-PCR

The Triton X100 method described above was used to isolate the total RNA from ~2 mg (dry weight) of *Frankia* strain CcI3 grown under the test conditions described in the section above. DNA was removed from RNA samples with DNase treatment using RNase-free DNaseI (New England Biolabs). RNA samples were mixed with 5U of DnaseI, 1/10th volume of 10X DNase buffer, 200U of RNaseOUT (Invitrogen) and brought to a final volume of 100 uL with DEPC-treated water. The samples were incubated at 37°C for 30min. To stop the reaction, 1 uL of 0.5 M EDTA was added to each reaction and the nuclease was heat inactivated by incubating at 75°C for 10 min. Three successive extraction steps with equal volumes of water-saturated phenol, phenol:chloroform (50:50), and chloroform were performed on the DNase-treated RNA to remove inactivated DNase and purify the RNA . The DNase-treated RNA was precipitated by the addition of a 1/10th volume of 3M sodium acetate (pH 5.2) and 2.5 volumes of ice-cold absolute ethanol followed by incubation at -80°C

overnight. The precipitated RNA was collected by centrifugation 10,000 x g in a microfuge for 15 min at 4°C and washed with cold 80% ethanol. The washed pellet was air-dried and suspended in nuclease-free water and stored at -80°C until used for Real-Time PCR.

cDNA synthesis was performed using SuperScript III Reverse Transcriptase (Invitrogen) according the manufacturer's directions. Both trHbN and trHbO mRNA were reverse transcribed separately in 20 uL reaction volumes using the HbNcci219 and HbOcci375 gene-specific primers, respectively. Each reaction mixture contained 1uL 10 mM dNTP mix (Invitrogen), 1 uL of the gene-specific reverse primer (2 pmol), 1.5 ug RNA template and nuclease-free water to a final volume of 13 uL. The reaction mixture was incubated at 65°C for 5 min., and placed on ice for at least 1 min. To each tube, 4 uL 5X First-Strand buffer, 1 uL 0.1M DTT, 1 uL of RNaseOUT (Invitrogen), and 1 uL of SuperScript III RT (200 U) were added, and the mixture was incubated at 55°C for 50 min. The reactions were inactivated by incubating at 70°C for 15 min.

To remove the RNA from cDNA samples, 1ul (2U) of Ribonuclease H (Invitrogen) and 69 ul of DEPC-treated H_2O were added to each tube, and the samples were incubated at 37°C for 30 min. The addition of nuclease-free water prior to extraction was necessary to increase the sample volume for cDNA recovery. cDNA was extracted with the addition of an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol (saturated with 10mM Tris-HCl, 1mM EDTA, pH 8.0). After the samples were mixed on a vortexer for 30 sec, they were centrifuged at 13,000 rpm for 2 min at 4°C and the aqueous phase was

transferred to a fresh microfuge tube. cDNA was precipitated with the addition of 1/10th volume of 7.5M ammonium acetate and 2.5 volumes of chilled absolute ethanol. The cDNA was recovered by centrifugation at 10,000 x g for 20 min. at 4°C. After the supernatant was decanted, the pellet was washed with chilled 80% ethanol, inverting the tube several times, and centrifugation at 10,000 x g for 5 min. at 4°C. The pellet was suspended in 35ul nuclease-free water.

Real-Time RT-PCR

Gene transcripts were quantified by amplification of cDNA with Power SYBR® Green PCR Master Mix (Applied Biosystems) according to the manufacturer's recommendations. The same primers specific to trHbN (HbNcci92 and HbNcci219) and trHbO (HbOcci299 and HbOcci375) that were used for the RT-PCR experiments were used for Real-time PCR. Prior to running experimental samples the primer concentrations were optimized to determine the minimum primer concentrations that gave the lowest threshold cycle (C_t) and maximized the magnitude of the signal while minimizing non-specific priming. One hundred nanograms of *Frankia* strain Ccl3 gDNA was used in a series of reactions that utilized various combinations of forward and reverse primer concentrations (i.e. 50, 300, and 900 mM). The exact protocol for primer optimization is outlined in the SYBR® Green PCR Master Mix product insert. For the experimental conditions, each reaction was performed in a 50 uL total volume with 10ul of cDNA (~1.5ug), 1ul each of forward and reverse specific primers (10 pmoles each), 25ul of 2X Power SYBR® Green PCR Master Mix,

and 13ul of nuclease-free water using MicroAmp® Optical reaction tubes and MicroAmp® Optical caps (Applied Biosystems). Triplicate amplication of all standards, unknowns and controls was performed in an ABI GeneAmp® 5700 Sequence Detection System adapted on a 96-well GeneAmp® 9600 PCR System (Applied Biosystems). Thermocycling parameters were as follows: activation of enzyme at 95°C for 10 min; followed by 40 cycles of a two-step cycle denaturization at 95°C for 15 sec and primer annealing/extension at 63°C for 1 min. Data analysis was performed using the GeneAmp® 5700 Sequence Detection System Software (v. 1.3) (Applied Biosystems).

A standard curve was generated from a dilution series of *Frankia* gDNA (1 ng to 1 ug gDNA) and also used to determine the efficiency of the primer sets. The standard curve and experimental reactions were also used to establish melting curves for each individual reaction. This determined the specificity of the primer sets and identified any nonspecific amplification that may have occurred. Upon completion of thermocycling process, each of the reactions were heated from 60°C to 90°C and the fluorescence was measured sequentially as the PCR amplicons melted apart. Upon plotting the derivative of the fluorescence vs. temperature, any non-specific priming could be identified as a peak observed separate from the peak of interest.

The threshold value for fluorescence of the reactions (samples and standard curves) was set manually. The number of reaction cycles in which the fluorescence of SYBR green in an individual reaction crossed the threshold was identified as the threshold cycle (C_t). The C_t of unknown cDNA samples was

compared to the C_t values of the standard curve to determine the relative amount of mRNA that was present. The PCR efficiencies for each primer set were determined by plotting ln (T) vs. Ct, where T is template concentration.

CHAPTER III

RESULTS

The Molecular Identification of Truncated Hemoglobin in *Frankia*

Identification of *Frankia* trHb amplicons.

An analysis of trHb amino acid sequences indicates that the actinomycetes form a single clade within the trHbO subgroup of truncated hemoglobins (Wittenberg et al., 2002). A PCR approach was used to identify truncated hemoglobin (trHb) genes in *Frankia*. Sequence alignment of trHbO genes from the actinomycetes *S. coelicolor*, *M. leprae*, *M. tuberculosis*, and *C. glutamicum* revealed conserved regions within the trHbO gene (*HbO*) that were potential candidates for primer design (Fig. 2). These aligned gene sequences were used to create the primers TrHb-F and TrHb-R corresponding to the base pair positions 61 to 81 and 241 to 262 of the *S. coelicolor* trHbO gene. The primer set was tested on 7 *Frankia* isolates representing phylogenetically distinct strains and yielded the expected 200bp amplicon (Fig. 3). A larger amplicon about 400bp was also observed with strain Cc1.17. As expected, *S. coelicolor* yielded a 200bp amplicon while *E. coli*, which is not known to produce a trHb (Wittenberg et al., 2002), failed to yield any PCR product.

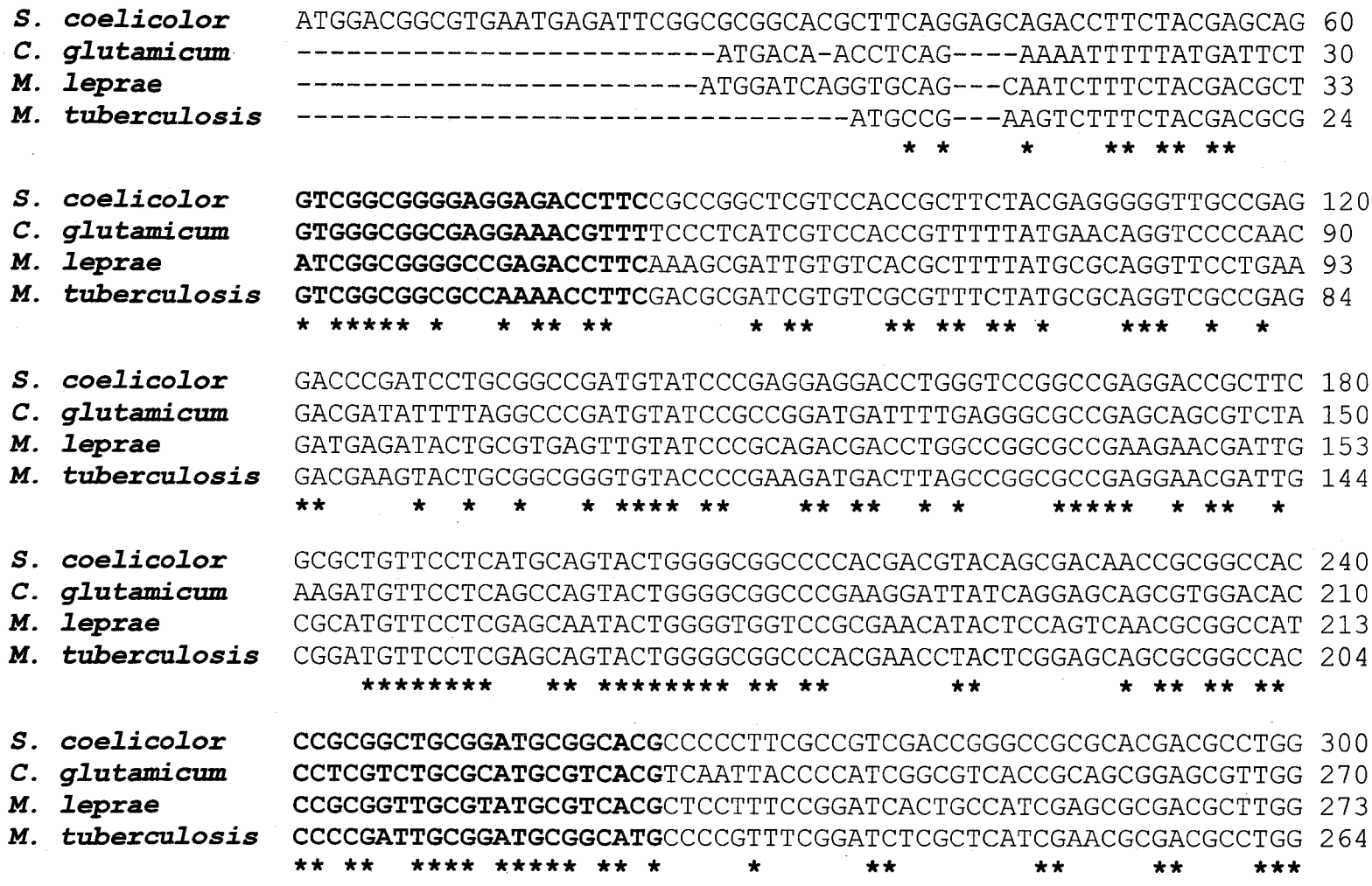


Figure 2. Cluster analysis of trHbO genes from actinomycetes used to design *Frankia* trHb primers. Regions used to design the forward (trHb-F) and reverse (trHb-R) primers are highlighted in blue and green, respectively. Conserved base pairs are denoted by asterisks (*).

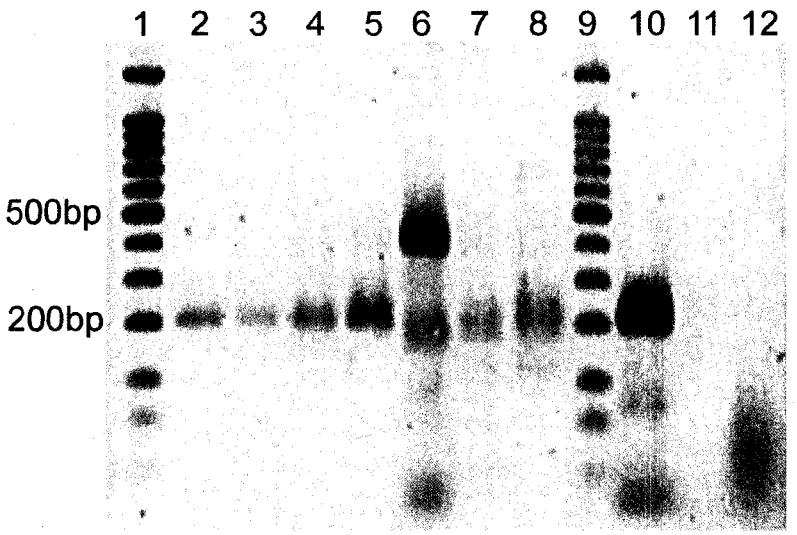


Figure 3. Agarose gel electrophoresis of trHb PCR products (202bp) amplified using TrHb-F and TrHb-R primers and DNA template from *Frankia* strains and controls. Lanes: (1 and 9) 100bp DNA ladder, (2) EAN1pec, (3) EUN1f, (4) Eul1c, (5) CN3, (6) Ccl.17, (7) ACN1ag, (8) Cpl1-P, (10) *S. coelicolor*, positive control, (11) no template, negative control and (12) *E. coli* , negative control.

Sequencing TrHbO Amplicons from *Frankia*.

From these samples, four (CN3, EAN1pec, Eul1c, and EUN1f) were sequenced and compared to the trHbO gene sequence of *Streptomyces coelicolor* (Fig. 4). During the course of this study, three genome sequencing projects were initiated for *Frankia* strains EAN1pec, Ccl3, and ACN14a, and the full trHbO genes for these isolates were identified from the draft sequences. The overall DNA similarities among the *Frankia* trHbO partial sequences ranged from 75% for strains ACN14a and CN3 to 87% for strains CN3 and Eul1c (Table 3). Similarities observed between *S. coelicolor* and the *Frankia* isolates ranged from 69% (EUN1f, Ccl3) to 76% (Eul1c). A database search using BLAST (Altschul et al., 1997) also showed high DNA similarities between these sequences and trHbO genes of other actinomycetes (Table 3).

Wittenberg et al. (2002) identified several important amino acid residues of trHbs in relation to heme coordination. Frameplot analysis (Ishikawa and Hotta, 1999) of the *Frankia* trHbO partial sequences determined their predicted amino acid sequences. The alignment of six predicted full and partial amino acid sequences is shown in Figure 5. All six *Frankia* sequences contained residues (B9F, B10Y, CD1Y, E7A, and E14F) that were unique to the actinomycete trHbO subgroup. These data including their high DNA sequence similarity to other trHbO genes, supports the presence of a trHbO in *Frankia*.

Phylogenetic analysis of truncated Hb sequences.

An analysis of the draft genome sequences for the three *Frankia* strains

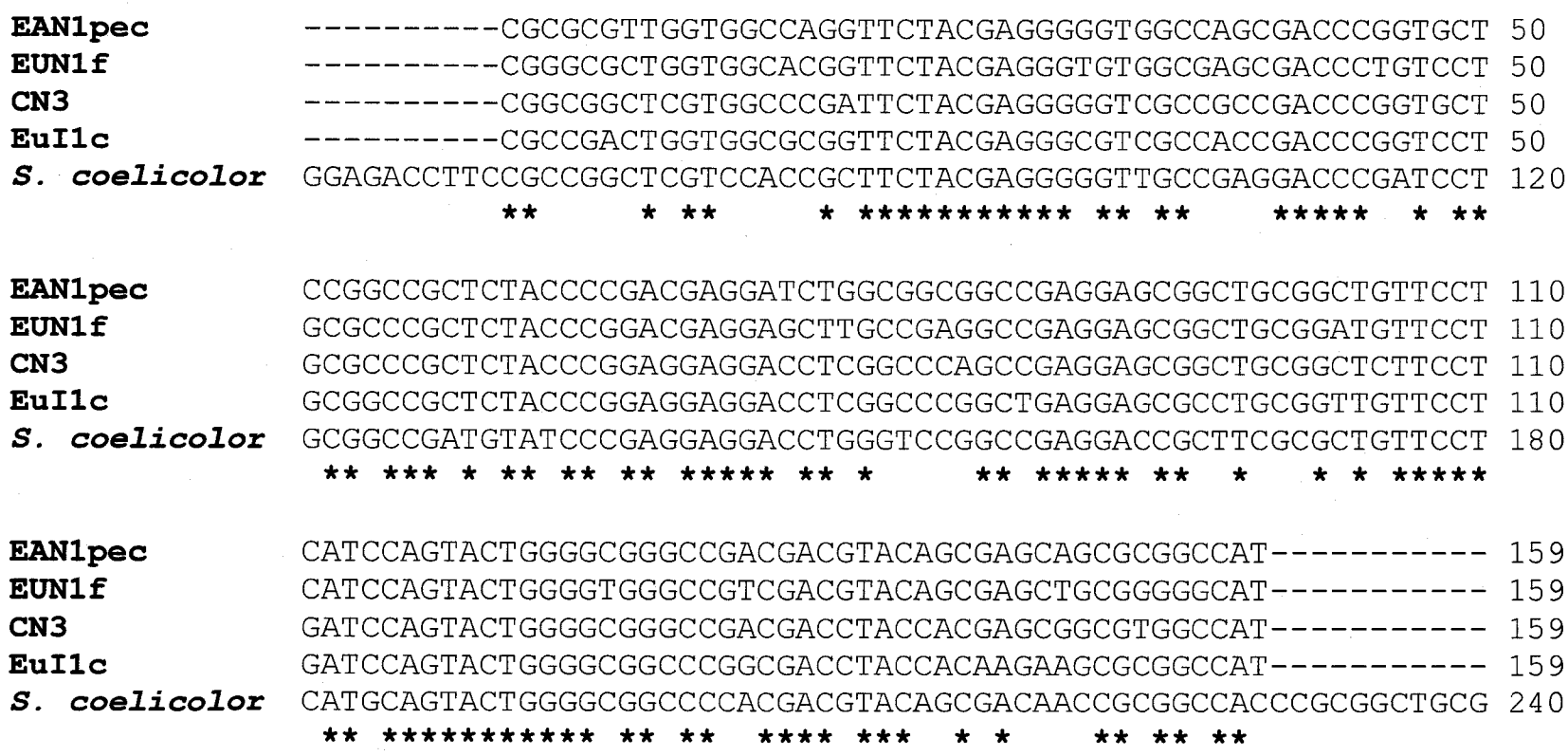
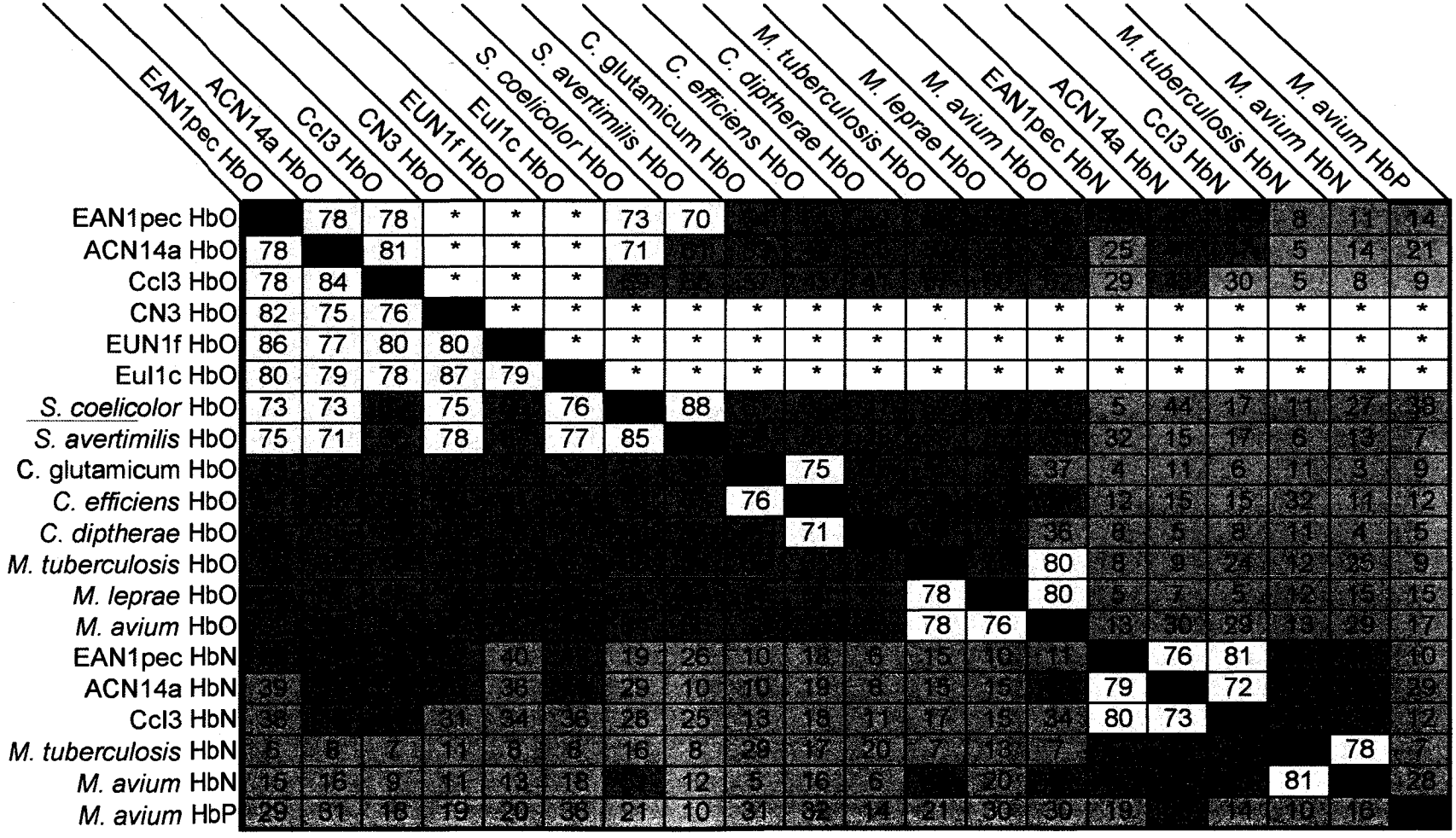


Figure 4. Alignment of *Frankia* trHb amplicon sequences with the *S. coelicolor* trHbO gene sequence. Conserved base pairs are denoted by asterisks (*).

Table 3: Percent identity of TrHb DNA sequences. Values on the bottom of the table are scores for partial trHb sequences. Values on the top of the table are scores for complete trHb gene sequences. Yellow (≥ 70%), green (between 41% and 69%), and blue (≤ 40%).



```
                    --------------------AAAAAA---BBBBBBBBBBBBBB--CCCCCC--EEEEEEE
EAN1pec             VAQNGPVNQPSPARP--RADFYEAVGGEATFRALVARFYEGVASDPVLRPLYPDEDLAAA 58
EUN1f               --------------------------------RALVARFYEGVASDPVLRPLYPDEELAEA 29
CcI3                VPQNGRVNQSPPRTLP-ISSFYDAAGGEPTFRKLVARFYQGVANDPVLRPLYPEEDLTGA 59
ACN14a              ------VSQPPTPAQPTTTTFFDAVGGEPTFRRLVARFYQGVANDPVLRPLYPEDDLAGA 54
CN3                 --------------------------------RRLVARFYEGVAADPVLRPLYPEEDLGPA 29
EuI1c               --------------------------------RRLVARFYEGVATDPVLRPLYPEEDLGPA 29
M. tuberculosis     --------------MP-KS-FYDAVGGAKTFDAIVSRFYAQVAEDEVLRRVYPEDDLAGA 44
S. coelicolor       ----MDGVNEIRRGTLQEQTFYEQVGGEETFRRLVHRFYEGVAEDPILRPMYPEEDLGPA 56
                                                    :* ***  ** * :** :**:::*  *

                    EEEEEEEEEEE-----------FFFFFFFF------GGGGGGGGGGGGGGGG-------
EAN1pec             EERLRLFLIQYWGGPTTYSEQRGHPRLRMRHVPFAIGPAERDAWLRIMESAVDSLG---L 115
EUN1f               EERLRMFLIQYWGGPSTYSELRGHPRLRMRH---------------------------- 60
CcI3                EERLRMFLIQYWGGPTDYQEQRGHPRLRRRHAPFAIGPTQRDAWLKIMRAAVDSLD---L 116
ACN14a              EDRLRLFLIQYWGGPSDYQELRGHPRLRMRHVPFAIGPAQRDAWLVVMRAAVDSLG---L 111
CN3                 EERLRLFLIQYWGGPTTYHERRGHPRLRMRH---------------------------- 60
EuI1c               EERLRLFLIQYWGGPATYHKKRGHPRLRMRH---------------------------- 60
M. tuberculosis     EERLRMFLEQYWGGPRTYSEQRGHPRLRMRHAPFRISLIERDAWLRCMHTAVASIDSETL 104
S. coelicolor       EDRFALFLMQYWGGPTTYSDNRGHPRLRMRHAPFAVDRAAHDAWLKHMRVALDELG---L 113
                    *:*: :** ******  * . ******* **

                    -----HHHHHHHHHHHHHHHH---
EAN1pec             APEHRAQLWDYLLMAANSLQNRPG 139
EUN1f               ------------------------ 
CcI3                PPDLDRQLWDYLSMAANSLQNRPD 140
ACN14a              PPDQYKTLWDYLQMAANSLQNRAD 135
CN3                 ------------------------
EuI1c               ------------------------
M. tuberculosis     DDEHRRELLDYLEMAAHSLVNSPF 128
S. coelicolor       SEEHEQTLWKYLTYAAASMINTPG 137
```

Figure 5. Alignment of the full and partial trHbO amino acid sequences from six *Frankia* strains with *S. coelicolor* and *M. tuberculosis*. Predicted amino acid sequences were obtained using Frameplot 2.3.2. The (3) conserved glycine motifs are denoted in **green**. Important amino acid residues in respect to heme coordination and ligand stabilization are denoted in **blue**. The conserved hydrophobic residues are denoted in **red**. The general globin fold topological positions (helices A-H) are shown above the alignments. The alignment was established using ClustalX program.

```
                    ----------AAAA-----BBBBBBBBBBBBBBB-CCCCCCC----EEEEEEEEEE
EAN1pec             ----------MSIYDAIGGASAVQAAVDEFYVRVTADPELAPFFAGKDIPRLKAHQQA 48
CcI3                ----------MSIYDTIGGATAVQAAVDDFYVRVTADPVLAPLFANRDLPRLKEHQRA 48
ACN14a              ----------MSIYQDIGGAKAVKAAVDEFYVRVLADPDLVRFFEGRDLASLKASQRE 48
M. tuberculosis     MGLLSRLRKREPISIYDKIGGHEAIEVVVEDFYVRVLADDQLSAFFSGTNMSRLKGKQVE 60
                              :***: ***  *::..*::***** **  *  :* . ::. **  *

                    EEEEEE---------FFFFFFFF----GGGGGGGGGGGGGGGG---HHHHHHHHHHHHHH
EAN1pec             FISAAIGGPEVYQGGAIASVHSGLRITDANFNAVVDHLVSALSGLGVPAETTGQIGAALA 108
CcI3                FIAAVIGGPEVYRGRDMAAVHATLGLTDAHFDAVVDHLLAALTGLGVPTETTGQIGAALA 108
ACN14a              FIGAALGGPEIYQGDAMSQVHASLGVGNAQFDGVVGHLLAAFASAGVPAEAAGQIAAVLG 108
M. tuberculosis     FFAAALGGPEPYTGAPMKQVHQGRGITMHHFSLVAGHLADALTAAGVPSETITEILGVIA 120
                    *:.*.:**** * *  :  **    :   :*. *..**  *::. ***:*:  :* ..:.

                    HHHH-----------
EAN1pec             PLRADIVTAK------ 118
CcI3                PLRSDVVTISK----- 119
ACN14a              PLRGEIVTAP------ 118
M. tuberculosis     PLAVDVTSGESTTAPV 136
                    **  ::.:
```

Figure 6. Alignment of the trHbN amino acid sequences from three *Frankia* strains with *M. tuberculosis* trHbN sequences. The (3) conserved glycine motifs are denoted in **green**. Important amino acid residues in respect to heme coordination and ligand stabilization are denoted in **blue**. The conserved hydrophobic residues of the ligand tunnel are denoted in **red**. The general globin fold topological positions (helices A-H) are shown above the alignments. The alignment was established using ClustalX program.

revealed the presence of a gene that was homologous to the *Mycobacterium* trHbN gene. Figure 6 shows the alignment of the predicted trHbN amino acid sequences from the annotated *Frankia* genomes with *glbN* of *M. tuberculosis*. The discovery of these homologous genes in the *Frankia* genome was surprising since the genus *Mycobacterium* was the only actinobacteria known to harbor trHb paralogues other than *glbO*. The complete trHb gene sequences of all 3 subgroups of trHbs (trHbN, trHbO, and trHbP) from *Frankia* and other actinomycetes were aligned to create a neighbor-joining distance tree (Fig. 7). Although my sample size was small, three distinct clades were evident that corresponded to the respective trHb subgroups. All of the *Frankia* trHbs were closest to the *Mycobacterium* sp. A second tree was constructed utilizing the partial trHbO sequences (Fig. 8). Within the *Frankia* clusters, I observed a distinct separation of trHbs that reflects the current taxonomic system for *Frankia* that groups strains into (3) distinct phylogenetic groups (I, II, and III) (Benson & Clawson, 2000; Normand et al., 1996). The details of *Frankia* taxonomy and its relatedness to trHb phylogeny will be discussed later in the discussion.

Analysis of TrHb Gene Neighbors and Potential Promoter Regions.

The Integrated Microbial Genome database (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi) from the Joint Genome Institute was utilized to compare the trHb gene neighborhoods among the three *Frankia* genomes and with several *Mycobacterium* genomes (Fig 9 and 10). TrHbN open reading frame (ORF) appears to be an isolated gene in *Frankia* genomes and does not appear to be

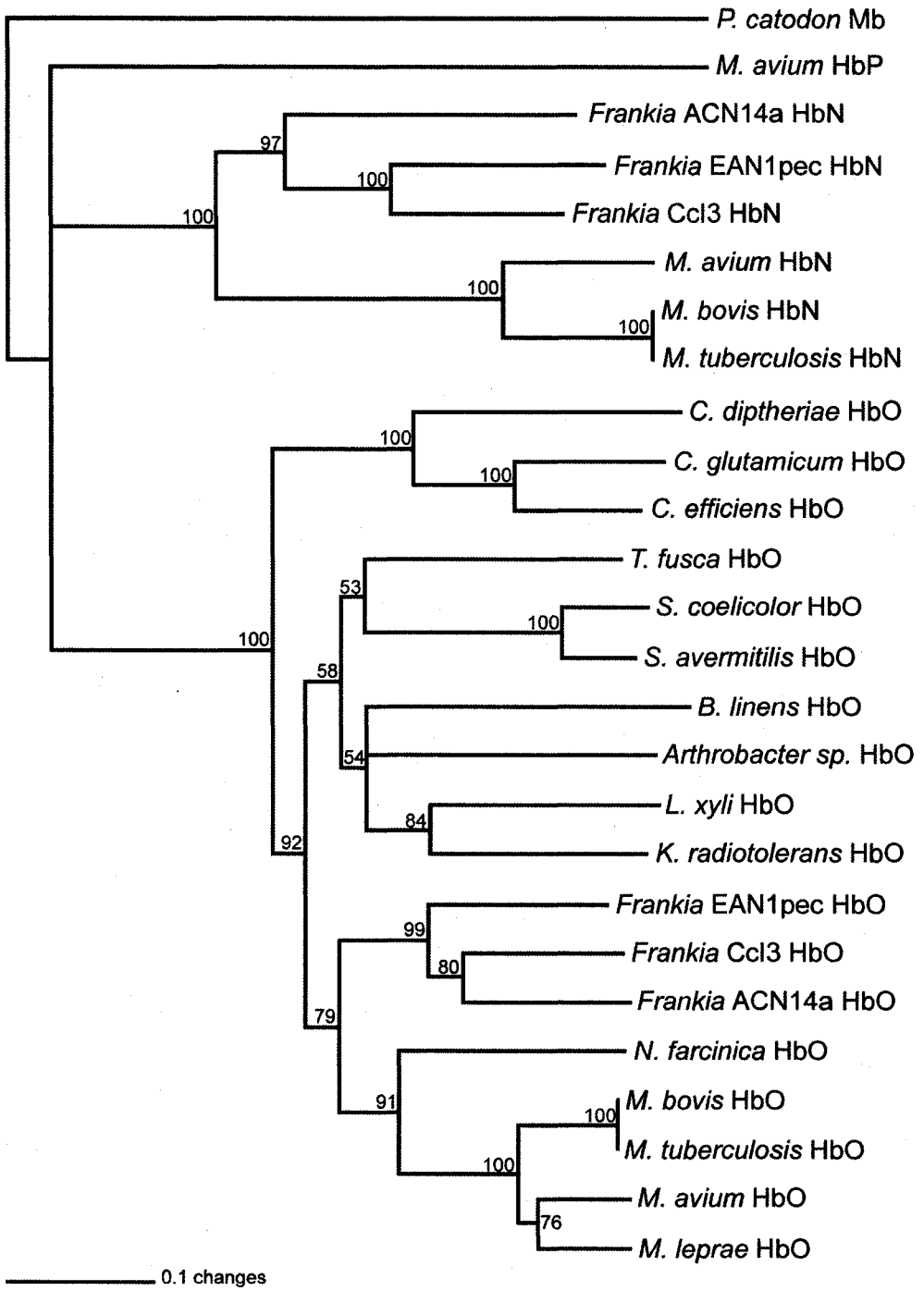


Figure 7. Phylogenetic dendogram representing neighbor-joining analysis of the complete trHb gene sequences (HbP, HbO, HbN subgroups) from *Frankia* and other actinomycetes. Sperm whale (*Physeter catodon*) Mb was included as an outgroup. The scale bar indicates 0.05 substitutions per site. Bootstrap values (above 50%) are shown as a percentage of 1000 replicates.

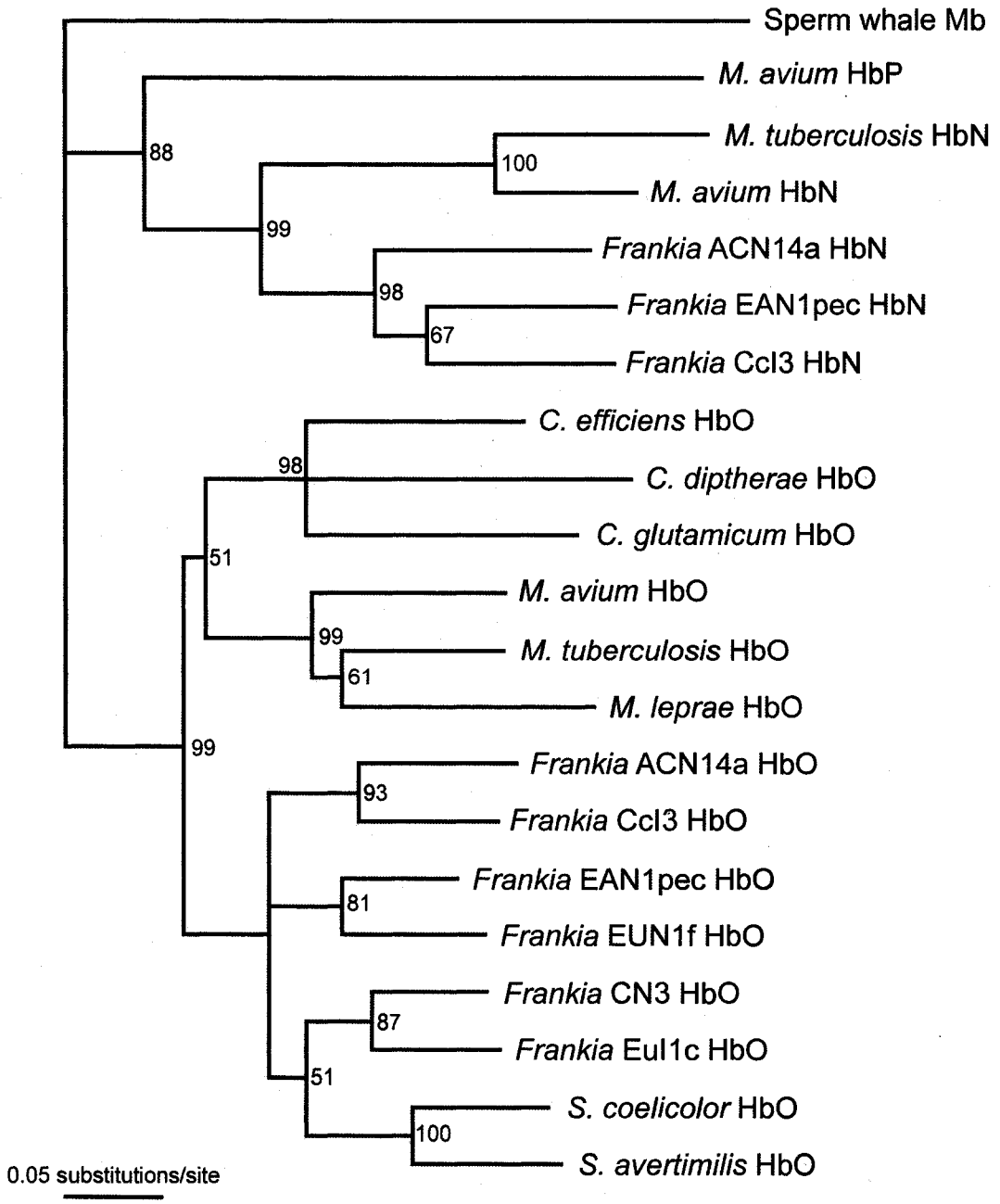


Figure 8. Phylogenetic dendogram representing neighbor-joining analysis of the partial trHb gene sequences (HbP, HbO, HbN subgroups) from *Frankia* and other actinomycetes. Sperm whale (*Physeter catodon*) Mb was included as an outgroup. The scale bar indicates 0.05 substitutions per site. Bootstrap values (above 50%) are shown as a percentage of 1000 replicates.

part of an operon. In contrast, the *Mycobacterium* sp. genomes have the trHbN gene adjacent to a predicted lipoprotein gene (Fig. 9). The shared synteny among the *Mycobacterium* trHbN gene neighborhoods suggests conserved functionality of this gene throughout that genus. There was no synteny among the *Frankia* trHbN gene neighborhoods or to *Mycobacterium* genomes. The conserved gene organization of the *Mycobacterium* chromosomes around glbN was not apparent in the *Frankia* trHbN gene neighborhoods.

The gene neighborhoods of the three *Frankia* trHbO genes are much more conserved (Fig. 10). Although gene organization was not highly conserved among the trHbO gene neighborhoods for the other actinomycete, each trHbO ORF was adjacent to a predicted alpha-glucosidase gene. Strikingly, this predicted gene was found upstream of *Frankia* trHbO genes, while it was located downstream of trHbO in other actinomycete genomes. It is not clear if the alpha-glucosidase and trHbO genes are co-regulated forming an operon.

I was interested in the identification of potential promoters for the *Frankia* trHb genes. Cournoyer and Normand (1994) analyzed potential *Frankia* promoter regions with -35 and -10 consensus sequences and identified a consensus sequence as TA(G/A)(G/A)T for the -10 promoter region and TTG(T/A)CG for the -35 region. These consensus sequences were not found upstream of the trHb genes in the three *Frankia* genomes (Fig. 11). However, conserved sequences were identified that may correspond to the -35 and -10 promoter regions. Lavire and Cournoyer (2003) compiled a list of several

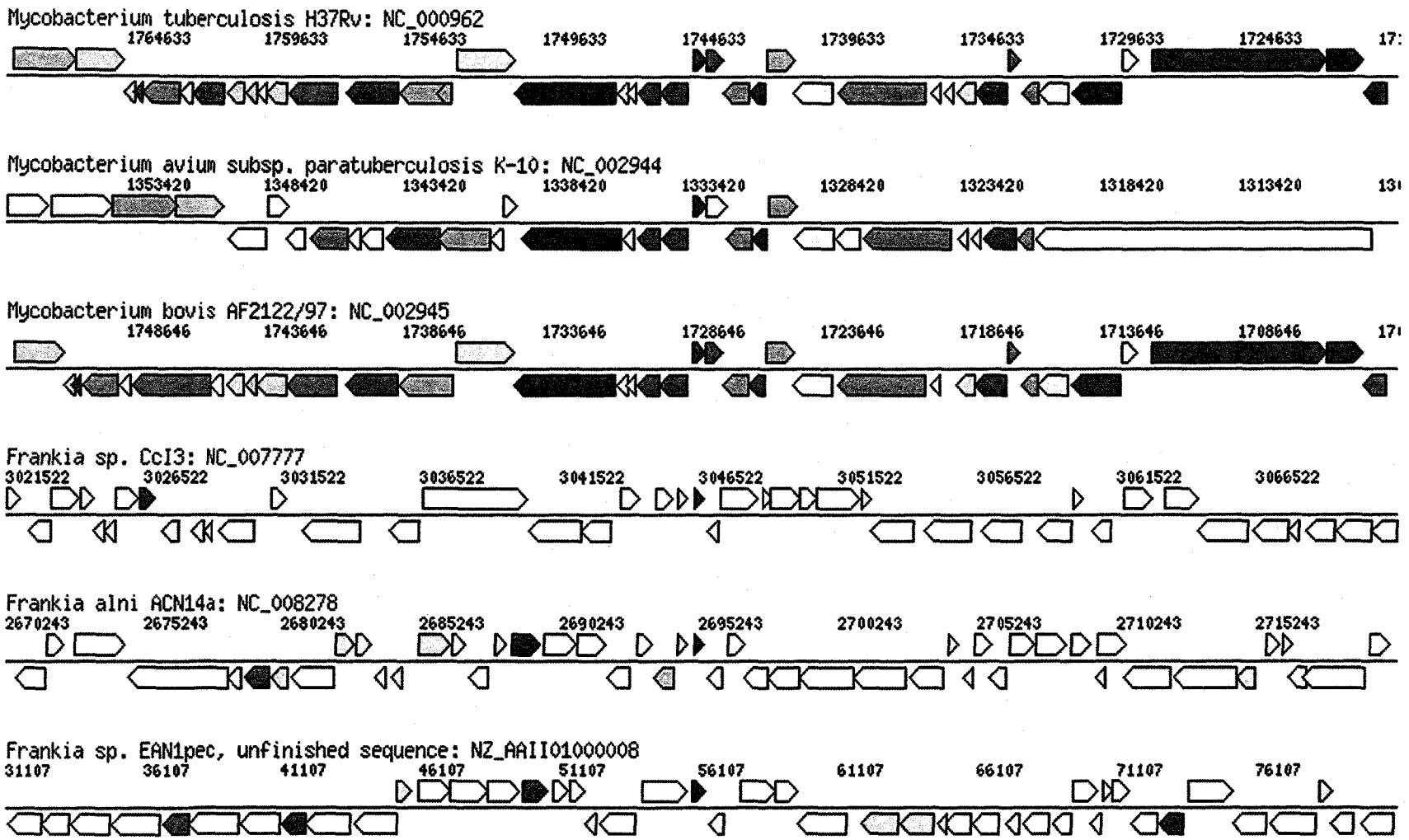


Figure 9. Comparison of trHbN gene neighborhoods of *Frankia* and *Mycobacterium* sp. TrHbN ORF indicated by red arrow. Generated using the Integrated Microbial Database from the Joint Genome Institute.

Figure 10. Comparison of trHbO gene neighborhoods of *Frankia* and *Mycobacterium sp.* TrHbO ORF indicated by red arrow, coupled with *aglA* ORF. Generated using the Integrated Microbial Database from the Joint Genome Institute.

CCI3 HbN

CTTCTGCCCACACGCGAGCGGAAGGAGCCGGACGGCCCGCGTTCGCCTGC
CATCTTTCTCTGCCGCGCGGACATTAGACTCTTTCAAGGAAATTGTGCCC
GTGGTTCAGTGAACCTGGCGGACGGGTACCAGTGCAGCCACGGTAACAAA
CATGGTCAACGACCTCCGGTCGGTCCTCTCCTGCGAAAGC**GAGA**CCTCTG
ATGAGTATCTACGACACTATCGGCGGCGCGACGGCCGTACAGGCCGCGGT

ACN14a HbN

TTCTCCTGACGCACCCGTGGAAGGAGCCAGGCGGCCTGTCCCGGCCTGCC
GTCCTTCTTCGCCGCGTGACCATCCGGCCCTGGGGCCGAACGATTCTCCT
GGGGTTCAGTGAACCCATCGGATCGTAAGCGGCGCAGCCATGCGGAAAAT
GGTGGTCACCGATCTCGCATCGGGCCACTCTTGCGAAAG**GGAGA**TCTCTG
ATGAGTATCTACCAGGATATTGGTGGCGCGAAGGCCGTGAAGGCCGCGGT

EAN1pec HbN

CGCCTGCTGATACGCGTGGCCGTGAGGAGCCGGGCGGCCCCCGACCGCGT
GCCGTCCGTCGCTCTCGCGTCGACATTCGGCTCTTTCGAACGCGTTACCT
GCGATTTAGTGGAAGTGTCGGAAATGTACCAATGCGGCCCATGGAAACGC
GCGTGGTCACCGACCTCGGCTGGGCCCATCCTGCGAAAGC**GAGA**CCTCTG
ATGAGTATCTACGACGCTATCGGTGGTGCGAGCGCCGTGCAGGCCGCGGT

CcI3 HbO

GCGACGATTCCGCTGCCCGGACGGCTGGTGCTCGCCAGCGGACCGGTGGG
GTACGACGGCGCGACCCTGACGCTGCCCCCCGACACGACGGCGTGGATCG
CACCCCGCGACGGCT**GAGA**ACACCCGTCGCCCGGATCGGACCGCACCGTC
GTGCCGCAGAATGGACGGGTGAACCAGTCTCCTCCGCGCACTCTCCCGAT

ACN14a HbO

GCGGCCGGCTGGTGCTTGCGAGCGGACCTGTGGCTTACGACGGCGCGACG
CTGACGCTGCCACCGGACACCACGGCGTGGGTCGCGCCCCGTGCGGGCT**G**
AGAGCGCAGCGGCGGGCCGGATGCGCGCCGATGTGACGCACAATGGAGGC
GTGAGCCAGCCCCCCACACCGGCCCAGCCCACCACCACGACCTTCTTCGA

EAN1pec HbO

CCCGTTCGGGTTGACCCACCCGGCTACCGCCGCACGGCGGACGTCCACCG
CCCGTCGAACCGCTCACCTCGGAAGACCCGGCCACGACCACCACCGGGCC
GGACGTCCCGTCGATC**GAGG**CCGGGGCCGCGGCCGGGGAGCGGGCGGCCA
GTGGCGCAGAATGGACCCGTGAACCAACCGTCACCCGCGCGACCACGCGC

Figure 11. Analysis of potential upstream regulatory regions of the *Frankia* trHbN and trHbO genes. ***Bold italics*** denotes the start site of the gene, **Bold** shows RBS site, underline= putative -10 sites, and double underline = putative -35 sites.

putative promoter regions that deviate from the above *Frankia* promoter consensus sequences.

Putative ribosomal binding sites (RBS) for each of the *Frankia* trHbs genes were identified (Fig. 11) and were similar to previously described *Frankia* RBS (Cournoyer and Normand, 1994; Lavire and Cournoyer, 2003). It is also interesting to note that each trHbO gene was predicted to initiate translation with GTG (valine) as a start codon in place of ATG (methionine). The use of this codon in *Frankia* is not unusual, since 32% of *Streptomyces* proteins are initiated with GTG (http://www.sanger.ac.uk/ Projects/S_coelicolor/).

The Effect of Environmental Stimuli on Truncated Hemoglobin Expression

The functions of truncated hemoglobins (trHbs) are unknown, but several potential roles have been hypothesized. For example, *Mycobacterium tuberculosis,* which produces two trHbs (trHbN and trHbO), has been hypothesized to use TrHbN (*glbN*) to detoxify nitric oxide produced by macrophages in tubercles (Ouellet et al 2002). *M. tuberculosis* trHbO *(glbO*) has been proposed to act as an oxygen delivery protein for terminal oxidases to aid in the stationary survival of this organism within hypoxic tubercles (Pathania et al. 2002, Liu et al. 2004).

The exact function(s) of hemoglobin in *Frankia* is unknown. Analysis of the *Frankia* genome elucidated the presence of two trHb genes *(trHbN* and *trHbO*).

Phylogenetic analysis grouped these *Frankia* trHbs closest to their respective *Mycobacterium* orthologs, suggesting potential analogous functions for the two *Frankia* hemoglobins (Niemann et al., 2005). Since both microbes are capable of intracellular growth during their life cycle, this hypothesis is not unreasonable. The following experiments were initiated to test several models by measuring. relative expression of both trHb genes under several environment conditions.

Optimizing RNA extraction from *Frankia*.

Prior to any measurements of relative trHb gene expression, two methods of RNA extraction were tested to evaluate which protocol would provide the highest quality and yield of *Frankia* RNA. The first method utilized RNAwiz™ (Ambion), a commercially available RNA isolation reagent, which was previously used by John et al. (2001) to isolate RNA from *Frankia* strain CpI1. The second method of RNA isolation was the Triton X-100 boiling protocol (Sung et al., 2003) that has the advantage of enriching for mRNA over stable RNA. This method was used to isolate RNA from both Gram-positive and Gram-negative bacteria, including *Mycobacterium vanbaalenii*, another high G+C actinomycete.

To test these methods, RNA was extracted from 200mg of *Frankia* EAN1pec cell mass. The Triton X-100 boiling method yielded twice the amount of RNA compared to RNAwiz™, and the purity of the RNA extracted, as determined by A_{260}/A_{280} ratios, was also improved (Table 4). Overall RNA quality of the samples was evaluated by visualization of the separated samples on an agarose gel. The Triton X-100 method proved to be superior producing a wider

Table 4. Comparison of EAN1pec RNA A_{260}/A_{280} ratios and the RNA yield by different methods.

RNA Extraction Method	A_{260}/A_{280}	**Yield (ug)**[†]
RNAwiz® (Ambion)	1.68	14.7
Triton X-100 Boiling Method (Sung et al, 2003)	1.81	30.6

[†] RNA extracted from 200 mg of cells

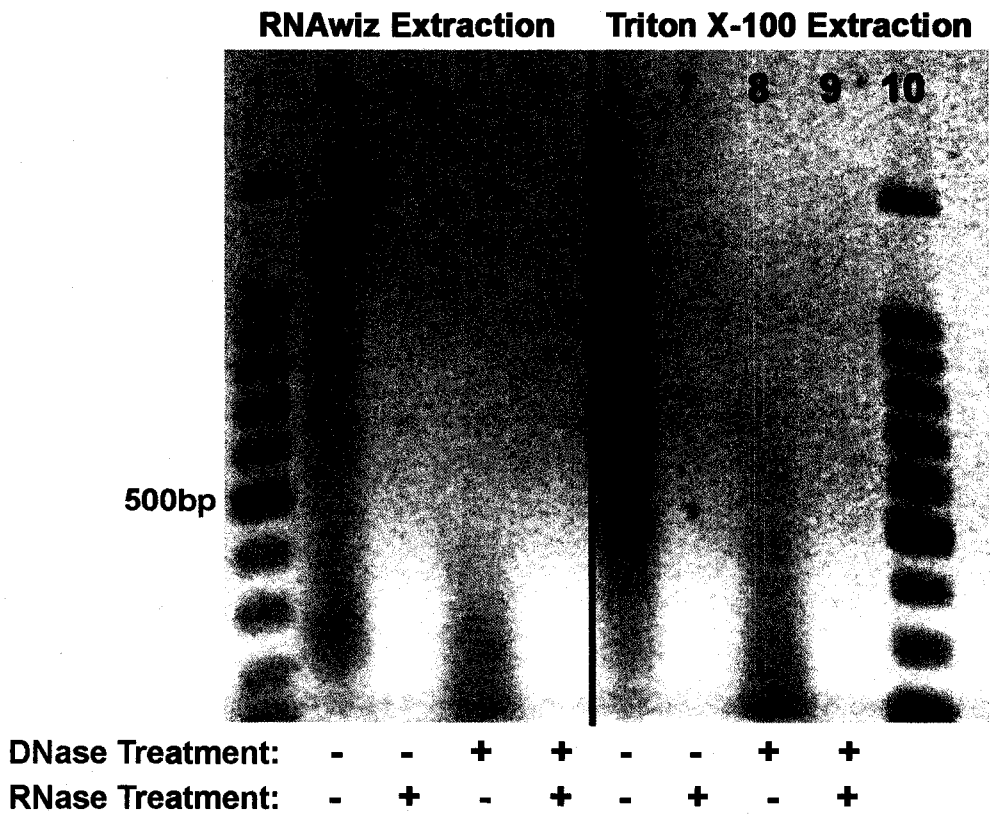


Figure 12. RNA analysis by agarose gel electrophoresis. Lanes: (1 and 10) 100 bp ladder, (2-5) RNAwiz extracted RNA, (6-9) RNA extracted by the Triton X-100 boiling method, (2 and 6) no treatment, (3-5 and 7-9) RNase and/or DNase treatment.

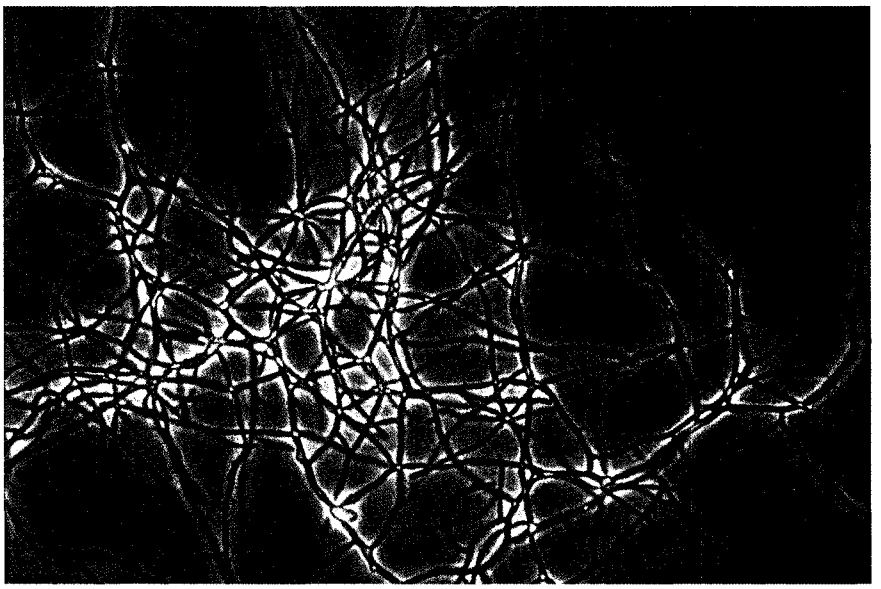
range of visible transcripts on the gel and the absence of contaminating 16S and 23S rRNA bands, which were present in the RNAwiz™ samples (Fig. 12). These preliminary experiments indicated that the Triton X-100 boiling method was the preferred protocol for RNA extraction from *Frankia* cultures. This RNA extraction method was used in all further experiments.

RT-PCR and Relative Expression of *Frankia* TrHbs.

For these experiments, mRNA levels measured by the use of Reverse Transcriptase PCR (RT-PCR) and *Frankia* trHb gene expression was compared to the level of *glnA* expression, which is constitutively expressed. Since both phylogenetically related *HbN* and *HbO* genes were present in the *Frankia* and *Mycobacterium* genomes, I predicted that their functions in *Frankia* might be analogous to their orthologues in mycobacteria.

The effect of nitrogen limitation on trHb expression. One hypothesis is that truncated hemoglobins function as oxygen scavengers, protecting the oxygen-labile nitrogenase complex during aerobic nitrogen fixation. Cyanoglobin (GlbN), a trHb present in *Nostoc* spp., is synthesized in the absence of combined nitrogen and the gene locus for *glbN* resides between two *nif* operons (Hill et al. 1996). *Frankia* strain Ccl3 was grown in media with (+N) or without (-N) a combined nitrogen source (NH_4Cl) to test the effect of nitrogen limitation on trHb expression. Under nitrogen-limiting conditions *Frankia* produces vesicles and reduces N_2 to NH_3 within the vesicle (Benson and Silvester 1993). *Frankia* Ccl3 grown for 6 days in -N medium produced vesicles, while vesicles were not

(A)



(B)

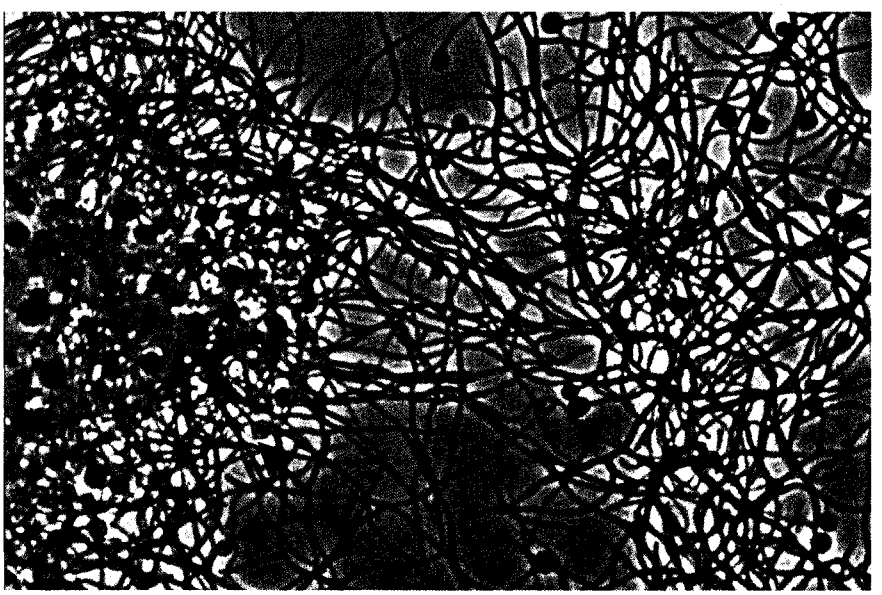


Figure 13. Photomicrographs of *Frankia* Ccl3 grown for 6 days in medium containing 5 mM NH_4Cl (A) or medium containing N_2 as a sole nitrogen source (B). The presence of vesicles are indicated by the symbol (V).

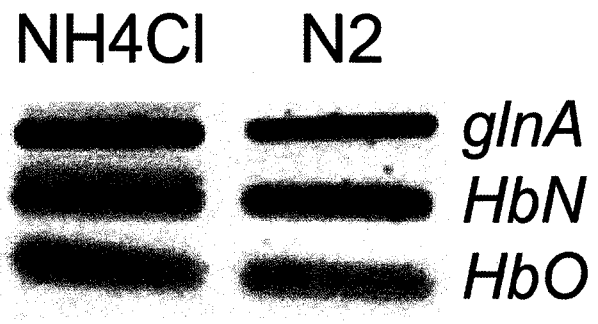


Figure 14. Nitrogen status did not alter hemoglobin expression. Transcriptional analysis of *HbN*, *HbO*, and *glnA* was determined by RT-PCR as described in the methods. *Frankia* strain Ccl3 was incubated in growth media with and without 5mM NH_4Cl as a combined nitrogen source. After 6 days, total RNA was isolated as described in the Methods and used as the template in a reverse transcriptase reaction.

produced in +N medium (Fig. 13). Under these conditions, the relative expression levels of the three genes were similar (Fig. 14). Analysis of the intensities of the bands for the three genes confirmed that the expression of both trHb genes relative to *glnA* expression were similar (data not shown) and suggest that neither of these genes may be directly involved in nitrogen fixation.

The effect of oxygen on trHb expression. Schwintzer et al. (2005) evaluated the effect of oxygen on hemoglobin production in *Frankia* strain ArI3 and observed an increase in hemoglobin levels in cultures grown under 1% oxygen atmosphere compared to cultures grown under 20% and 40% oxygen atmospheres. I hypothesized that *Frankia* trHbO may deliver oxygen to terminal oxidases to stimulate rates of respiration under hypoxic conditions. Consistent with this model, Tjepkema et al. (2002) observed rapid oxygen kinetics for *Frankia* hemoglobin in strain CcI3 suggesting a role of oxygen transport over short distances. In *M. tuberculosis,* trHbO has been shown to function as an oxygen shuttle to respiratory enzymes (Pathania et al, 2002; Liu et al, 2004, Milani et al, 2003).

Frankia CcI3 cultures were grown under oxic (cultures aerated with atmospheric oxygen) and hypoxic (static cultures) conditions for 7 days. Since high levels of hemoglobin production (Hb/protein ratio) were observed in *Frankia* strain ArI3 grown under 1% oxygen (Schwintzer & Tjepkema, 2005), RNA was extracted at 7 days. The results are shown in Figure 15. The relative expression of trHbO was greater in static cultures than aerated cultures, while trHbN expression was similar under both conditions. These results suggest that *Frankia*

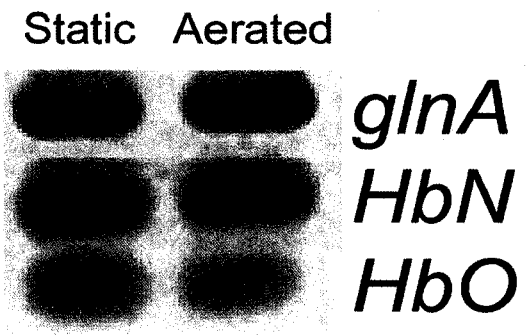


Figure 15. Hypoxic conditions increased expression of *HbO*. Transcriptional analysis of *HbN*, *HbO*, and *glnA* was determined by RT-PCR as described in the methods. *Frankia* strain Ccl3 cultures were grown in basal media supplemented with NH_4Cl under hypoxic and oxic conditions by incubating the cultures either statically or with aeration. After 7 days, total RNA was isolated as described in the Methods and used as the template in a reverse transcriptase reaction.

TrHbO may function under hypoxic conditions to shuttle oxygen to the respiratory chain, similar to the mycobacteria.

The effect of nitric oxide (NO) on trHb expression. During intracellular pathogenesis, mycobacteria are bombarded with toxic nitric oxide (NO) species generated by macrophages in tuberculosis granulomas. The trHbN protein functions as a protective molecule, catalytically reacting NO with oxygen to generate a harmless nitrate molecule (Couture et al, 1999; Milani et al, 2001; Pathania et al, 2002; Ouellet et al, 2002; Milani et al, 2004).

I predicted a similar protective function in *Frankia*. Similar to animals, plants also use NO as a defense mechanism against pathogens (Neill et al., 2003). *Frankia* may use trHbs to detoxify NO during the initiation stages of the host plant infection process to establish a symbiosis. To test my hypothesis, the spontaneous NO donor, S-nitroso-N-acetylpenicillamine (SNAP), was added to *Frankia* Ccl3 cultures and the relative expression of trHbN and trHbO was evaluated. RNA samples were taken every hour for 4 hours incubation. As a control, the NO scavenger, carboxy PT10 (cPT10), was added to cultures containing SNAP. Figure 16 shows the RT-PCR results from these experiments. Relative trHbO expression was unchanged under all test conditions. However, the relative expression of trHbN increased through the first 3 hours of exposure to SNAP and the signal decreased after 4 hours. Preliminary work measuring expression of trHbN in *Frankia* strain EAN1pec grown in 500uM SNAP cultures for 2 hours also produced a strong band with RT-PCR (data not shown). The addition of carboxy PT10 (cPT10), a NO scavenger, to cultures growing in the

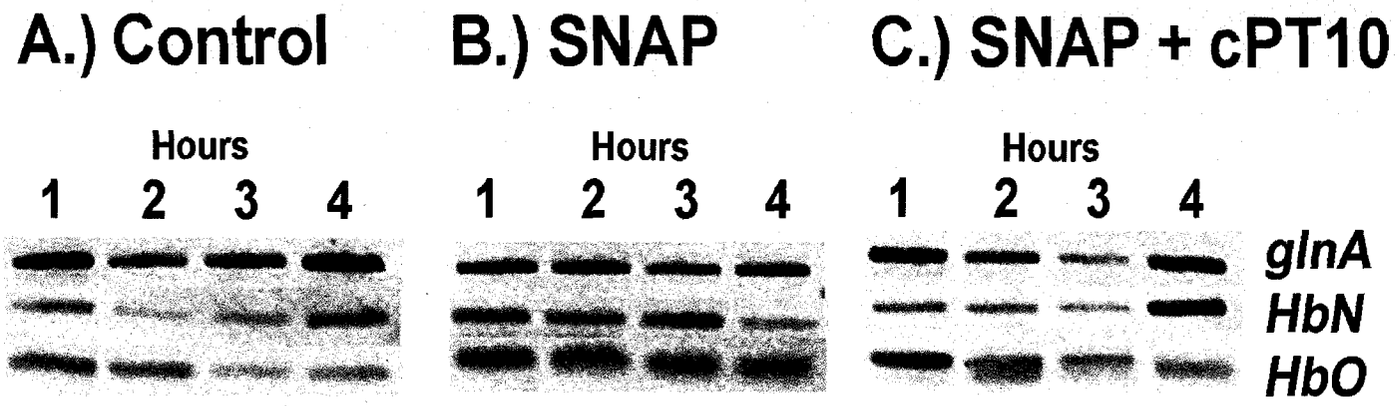


Figure 16. Nitric oxide (NO) release stimulates *trHbN* expression. Transcriptional analysis of *trHbN*, *trHbO*, and *glnA* was determined by RT-PCR as described in the methods. Cultures were incubated in growth medium in the presence of a spontaneous NO donor, SNAP. Total RNA was isolated every h and used as the template in a reverse transcriptase reaction as described in the Methods. **(A)** Control (no addition) **(B)** 400 μM SNAP, and **(C)** 400 μM SNAP + 400 μM carboxy-PT10 (a NO scavenger).

presence of SNAP decreased the level of trHbN expression. Control cultures in the absence of SNAP showed no change in the relative expression over the same time period. These results indicate that *Frankia* trHbN gene expression was stimulated by NO and suggests that *Frankia* trHbN may be involved in nitric oxide detoxification.

Real-Time RT-PCR and Expression of *Frankia* trHbs.

To support the results from my preliminary expression studies I utilized quantitative Reverse Transcriptase PCR (qRT-PCR) to measure trHb expression in *Frankia* strain Ccl3. In addition to repeating my previous experiments, H_2O_2 and paraquat were also added to *Frankia* cultures to evaluate trHb expression under oxidative stress conditions. However, it should be noted that I did not repeat the oxygen studies with qRT-PCR since our research collaborators were performing these experiments. Figure 17 shows the results of these experiments. For each experimental condition, the expression levels of trHbN and trHbO transcripts are shown as a ratio that is relative to expression in the untreated control cultures.

The dissociation curves that were generated with the end products of each qRT-PCR reaction verified the specificity of each primer set. The presence of non-specific amplicons and/or primer-dimers were not detected in the melting curves (data not shown). The efficiency of each primer set was also determined using standard curves and proved that the RT-PCR conditions worked well. The

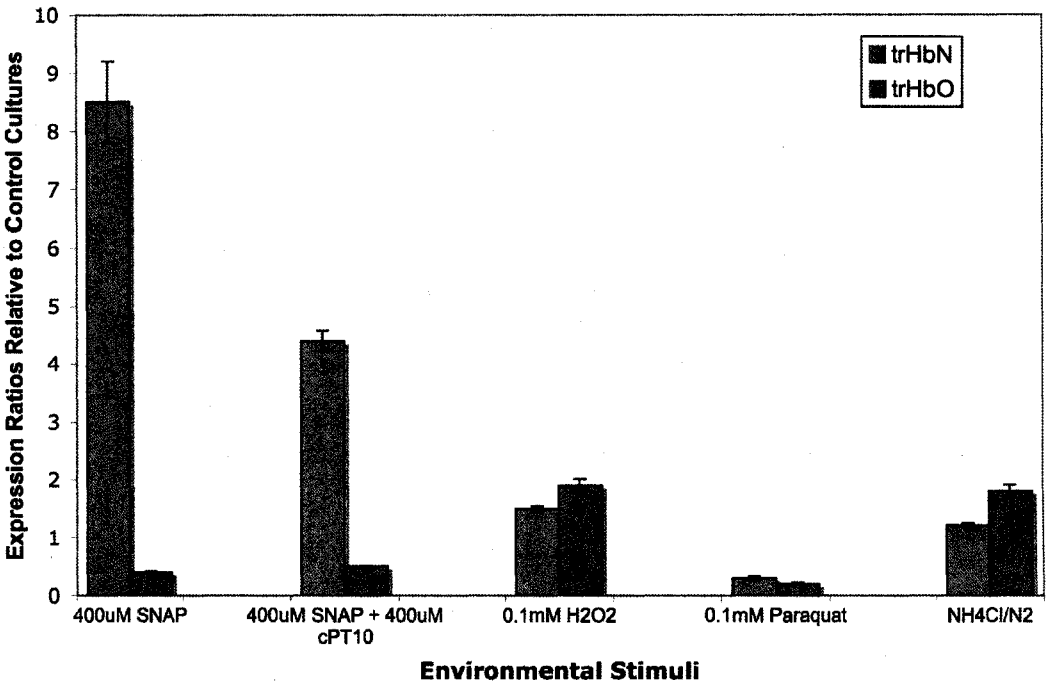
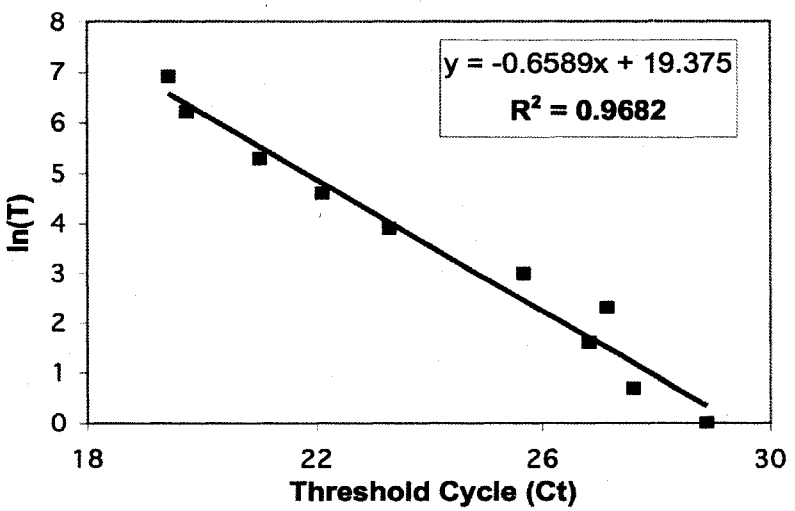


Figure 17: Transcriptional analysis of *Frankia* Ccl3 trHbN and trHbO gene expression using qRT-PCR. The relative expression ratio reflects the expression of each gene under that environmental stimulus relative to the expression under control conditions (no stimulus).

(A) TrHbN Standard Curve



(B) TrHbO Standard Curve

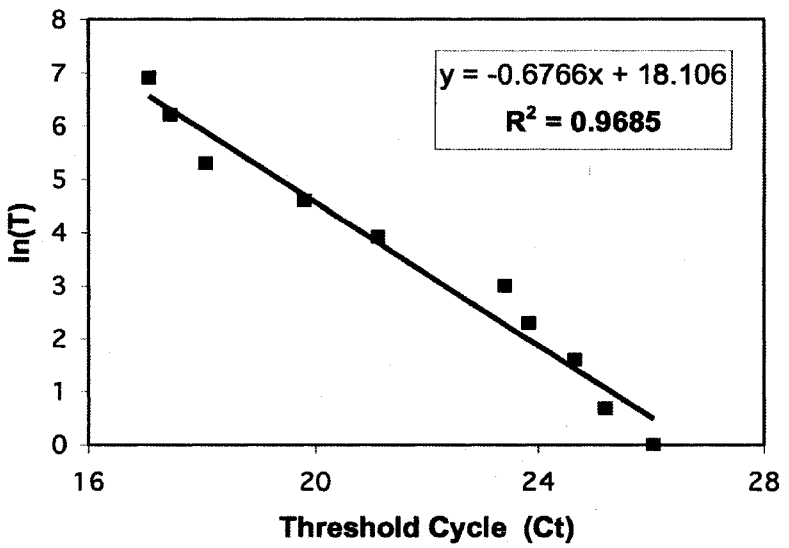


Figure 18: Standard curves generated for each trHb qRT-PCR primer set used with SYBR Green PCR Master Mix. Correlation coefficients (R^2) and equations are provided within each plot. (A) trHbN and (B) trHbO.

correlation coefficients (R^2) for each primer set are included with their respective standard curve (Fig. 18).

The effect of nitrogen limitation on trHb expression. The semi-quantitative RT-PCR data of *Frankia* strain Ccl3 cultured under nitrogen limiting conditions supported the earlier RT-PCR results. There was no substantial difference in the gene expression level for either HbO or HbN under nitrogen-sufficient and -limiting conditions (Fig. 17). HbN expression was equivalent under both conditions, while HbO expression showed small increase under growth with NH_4Cl. These data, in addition to the previous RT-PCR results, suggests that these two hemoglobin genes were not be up-regulated during nitrogen fixation conditions.

The effect of nitric oxide (NO) on trHb expression. The effect of nitrosative stress on trHb expression was also quantified by qRT-PCR. *Frankia* Ccl3 cultures were grown in media containing SNAP and/or cPT10 as described above. After 3 hours of exposure, RNA samples were extracted since this time point yielded the strongest band intensity in the RT-PCR experiments for HbN (Fig. 16). HbN gene expression increased nearly 10-fold in cells exposed to 400uM SNAP compared to control (untreated) cells (Fig. 17). The addition of 400uM cPT-10 reduced the effect of 400uM SNAP, causing only a 4-fold increase of HbN expression. This result indicates that cPT-10 was unable to scavenge all of the NO generated by SNAP under these conditions. HbO expression was reduced by 0.5-fold under these conditions. Since cPT-10 was

unable to relieve this effect, this would suggest that this reduction was not a NO effect but may be due to a change of oxygen levels in the culture.

The effect of oxidative stress on trHb expression. The oxidative burst of reactive oxygen species (ROS) by plants is another common defense mechanism against invading pathogens (Tavares et al, 2007). One problem microbial symbionts face is that they may not be initially distinguished by the host as a "friend" but instead as a "foe". Thus, a defense mechanism must be utilized to help establish this early stage of symbiosis. The use of catalases and superoxide dismutases in *Frankia* has been evaluated as a means to tolerate ROS from plant hosts (Hammad et al., 2001; Tavares et al., 2003; Santos et al., 2007). I predicted that one of the trHbs from *Frankia* might be up- regulated during oxidative stress as another mechanism to protect against ROS.

Studies of oxidative stress often use both paraquat and hydrogen peroxide (H_2O_2) so the effects of intracellular verses extracellular H_2O_2 exposure can be evaluated. When paraquat is added to the culture medium it is metabolized to produce endogenous H_2O_2. In this experiment, the addition of hydrogen peroxide to the media did not result in a noticeable change in HbN (< 1.5 fold) or HbO (< 2 fold) gene expression levels compared to the control culture (Fig. 17). Interestingly, the addition of paraquat decreased the expression of both trHbN and trHbO by nearly 10-fold. It is unclear whether or not this decrease in expression was due to the down-regulation of these genes or possibly due to cytotoxic effects of intracellular concentrations of H_2O_2.

CHAPTER IV

DISCUSSION

The research described in this thesis concentrated on the identification and expression of truncated hemoglobin genes in *Frankia* using molecular methods. Understanding the function(s) of trHbs in *Frankia* may give us better insight into *Frankia*-actinorhizal symbiosis and will allow to efficiently exploit this relationship for its ecological and economical benefits. Studying trHbs and their role in the intracellular mutually beneficial symbiosis of *Frankia* may also allow a better understanding of the intracellular pathogenic symbiosis of the closely related actinomycetes of the *Mycobacterium* genus.

Upon its introduction into host lung tissue, *Mycobacterium tuberculosis* is met by the host immune system and it is able to infect and begin to multiply intracellularly in inactivated macrophages (Ouellett et al., 2002; Pathania et al., 2002ab). These infected macrophages are then surrounded by other activated macrophages recruited to surround the infected tissue and subsequently form the characteristic granuloma. Upon their arrival the macrophages bombard the infected cells with nitric oxide (NO) to attempt to kill the pathogenic bacteria. The cytotoxic effects of NO stem from the ability of this compound to exert damage to DNA and to inhibit the activity of biologically important enzymes (Ouellett et al., 2002; Pathania et al., 2002ab). This immune response in essence forces the

microbe into a latent state which may last for several years and eventually become reactivated later in the infected individuals life.

The trHbs of *M. tuberculosis* are believed to aid in the long-term survival of the microbe in this dormant state. During long-term infection, the tubercle bacteria are competing with lung tissue for oxygen so they are essentially maintained in a low oxygen environment. The *glb*O gene product (HbO) is believed to act as a shuttle, to sequester and carry oxygen to terminal respiratory enzymes at the membrane in order to increase respiration rates (Pathania et al. 2002a; Liu et al., 2004). The *glb*N gene product (HbN) of *M. tuberculosis* has been proven to function as an oxygen-dependent nitric oxide dioxygenase that reacts host NO with an oxygen molecule to convert the compound to harmless nitrate (Ouellett et al., 2002; Pathania et al., 2002b).

Though each of these microbes are involved in very different relationships these actinomycetes live very similar lifestyles. They both rely on intracellular infection of their respective hosts, both overcome host defense strategies either upon or during infection, and both are maintained in specialized structures/tissue produced by their host. In the case of *Frankia,* the microbe persists in root nodules while *Mycobacterium* sp. are contained in granulomas within lung tissue. It is not surprising that each of these bacteria possess multiple trHbs that potentially serve different functions to aid in their survival in various niches. After all, each microbe may also exist in a free-living state. This research was the first step in showing that *Frankia* trHbs have functions analogous to their orthologs in *Mycobacterium* sp.

Molecular Identification of Truncated Hemoglobin in *Frankia*

Sequencing the TrHb Amplicons

Wittenberg et al. (2002) had previously performed phylogenetic analyses of truncated hemoglobin with a number of organisms. Three distinct groups (groups I, II, and III; or trHbN, trHbO, and trHbP, respectively) were apparent within the trHb family. Diversity of trHbs among the high G+C actinomycetes was restricted to trHbO where a distinct cluster of organisms (*S. coelicolor*, *C. diptherae*, *M. smegmatis*, *M. tuberculosis*, *M. avium*, and *M. leprae*) was formed within the trHbO group as a whole. However, the presence of trHbN or trHbP in any of the actinomycetes was strictly confined to *Mycobacterium* sp. It was assumed that if any trHb genes were present in *Frankia*, it would most likely be a group II (trHbO) hemoglobin. This would also be in agreement with the theory that all trHbs evolved from trHbO of a last common ancestor of actinomycetes and Proteobacteria and also explains the phenomenon that most bacteria possessing one or multiple trHbs have one trHb that is a group II trHb (Vuletich & Lecomte, 2006).

The sequencing of *Frankia* PCR products amplified from primers designed from conserved actinomycete trHbO regions confirmed the presence of this gene in a number of genetically diverse *Frankia* strains (Figure 4). Upon the evaluation of draft genome sequences of *Frankia* strains EAN1pec, ACN14a, and Ccl3 the presence of a second trHb gene (trHbN) in *Frankia* was also confirmed. Having the sequenced genomes also allowed me to obtain the complete gene

sequences for trHbO in these strains. The presence of trHbs in these various diverse strains suggests that these genes are likely conserved in most *Frankia* isolates. There are a number of important amino acid residues in respect to heme coordination and ligand stabilization that are conserved in the sequences of *Frankia* and other actinomycete trHbs (Figs. 5 and 6). Specifically, these residues are B9F, B10Y, CD1Y, E7A, and E14F for trHbO and B9F, B10Y, CD1F, E7L, and E14F for trHbN actinomycete sequences (Wittenberg et al., 2002). The conservation of these residues has evolved differently in various bacterial groups and suggests specific functional adaptation for each trHb (Milani et al., 2005). There are also other important amino acid residues that are conserved in *Frankia* and *Mycobacterium* trHbs that have been shown in *M. tuberculosis* to be important in respect to structure and function. These residues in *Frankia* trHbs will be discussed later in greater detail with the various expression experiments to support my hypotheses.

The Phylogeny of TrHbN and TrHbO in *Frankia*

The neighbor-joining distance trees that were created (Fig. 7 and 8) using *Frankia* and other known actinomycete trHb gene sequences revealed that of all the known actinomycetes, *Frankia* trHbs are more phylogenetically related to *Mycobacterium* trHbs. A high bootstrap value at each of these nodes strongly infers this relationship. This was the first evidence to support that trHbs in *Frankia* may have similar function to those in *Mycobacterium*.

Three distinct phylogenetic groups (I, II, III) of *Frankia* have been identified based on 16S rRNA sequence studies (Benson and Clawson 2000; Normand et al., 1996). For the Group II representatives (Ccl3 and ACN14a), the HbO genes formed a distinct group that was separate from the other *Frankia* strains (Fig. 8). The three representative of Group III (EAN1pec, EUN1f and Eul1c) and representative (CN3) from a fourth group of related "*Frankia*-like" nodulation and fixation defective (Nod^-/Fix^-) actinomycetes formed separate subgroup of the *Frankia* strains. This clade branched out into the two groups: effective strains and defective strain. Strain Eul1c forms nodules on its host plant, but generates ineffective nodules (Baker et al., 1980).

The Gene Neighborhoods of TrHbN and TrHbO in *Frankia*

The gene neighborhoods of trHbs in *Frankia* were examined using the Integrated Microbial Database from The Joint Genome Institute (Fig. 9 and 10). As observed in Figure 9, there is a lack of synteny in the gene neighborhoods surrounding the trHbN gene of *Frankia* as compared to *Mycobacterium* sp. The gene neighborhoods of *Mycobacterium* sp. surrounding *glb*N are highly conserved. Unlike the *Frankia* genomes, the location of the HbN gene in *Mycobacterium* sp. is directly adjacent to a predicted lipoprotein gene (*lprI*) (Fig. 9). My first intuition was that this suggested co-transcription of these two genes in the pathogen. Co-transcription of these two genes might have also suggested a difference in regulation of *glb*N between *Mycobacterium* and *Frankia*. Studies performed by Ouellett et al. (2002) showed that these two genes are in fact co-

transcribed in *Mycobacterium bovis*, the model organism for *M. tuberculosis* studies. However, the activity of HbN was proved to be independent of the predicted lipoprotein. Knockout mutants of HbN were unable to metabolize NO in comparison to wild-type *M. bovis*. Since these two genes are co-transcribed they wanted to show that the decrease in HbN activity was not also due to a lack of *lprI* expression from the knockout since they are co-transcribed. By introducing a plasmid (*glbN*$^{+}$, *lprI*$^{-}$) into the HbN mutant they were able to restore HbN activity without the expression of lipoprotein and NO metabolism was equivalent to that of the wild-type. Thus HbN function, though co-regulated with the *lprI* gene, is independent of the lipoprotein. This gave me further confidence in an analogous function of the trHbN gene in *Frankia*, which is not co-transcribed with *lprI*.

It should also be noted that the trHbN gene regions of the Ccl3 and ACN14a genomes have much fewer ORFs when compared to EAN1pec (Fig. 9). This observation was not surprising when one compares the smaller 5.43 Mbp genome of Ccl3 (a narrow host range strain) and the 7.50 Mbp genome of ACN14a (a medium host range strain) to the more promiscuous 9.04 Mbp genome of EAN1pec (a broad host strain) (Normand et al., 2007ab). It is frequently observed in nature that organisms occupying narrow niches undergo genome contractions to rid of unneeded genes, whereas diversification favors genome expansion (Normand et al., 2007ab). This example of reductive evolution can also be observed in the *glbO* gene neighborhood of *M. leprae* compared to the other *Mycobacterium* sp. genomes (Fig. 10). *M. leprae* is an

obligate intracellular pathogen unlike the other mycobacteria represented which are facultative intracellular pathogens or opportunistic pathogens (Wittenberg et al., 2002). It is not clear if the difference observed in gene organization around trHbN of the three *Frankia* genomes reflects differences in functionality.

Unlike their trHbN paralogs, the gene neighborhoods of trHbO in the three *Frankia* strains were more highly conserved (Fig. 10). However, a similar conservation of synteny did not carry over to *Mycobacterium* sp. though within these genomes there was good conservation of gene organization within the genus. The major difference of the mycobacteria was observed in the gene neighborhood for *M. leprae*. This phenomenon could be a direct result of reductive evolution as mentioned previously, but can also explain why the function of HbO in this pathogen is believed to be different than the HbO orthologs of other *Mycobacterium* sp. Much like HbN of *M. tuberculosis* and *M. bovis*, HbO in *M. leprae* is believed to function in the protection from NO and not as a carrier of oxygen to increase rates of respiration (Visca et al., 2002; Ascenzi et al., 2006). Though the mechanism of NO detoxification is slightly different in this species. Here, the HbO protein is believed to react with peroxynitrite, a toxic intermediate of NO reactions with superoxide radicles. The function of trHbO in *Frankia* will most likely reflect the function of HbO in *M. tuberculosis* considering that *Frankia*, like the pathogen, possesses 2 trHb genes. It would be unlikely during the course of evolution for any microbe that 2 genes of similar function would be maintained. The varied function of HbO in *M. leprae* is most likely a direct result of reductive evolution.

The other interesting characteristic of the *M. leprae* gene neighborhood surrounding HbO was that this member of the mycobacteria was the only one to not have the *glb*O gene grouped with a putative alpha-glucosidase gene (*agl*A) (Fig. 10). The trHbO gene products of the 3 *Frankia* strains were also grouped adjacent to a putative alpha-glucosidase gene, however, the trHbO ORF was oriented downstream of *agl*A. Unfortunately no work has been performed to determine whether or not *glb*O of *Mycobacterium* sp. is co-regulated with *agl*A and subsequently dependent on its gene product to function.

Characterizing the Promoter Regions of TrHbN and TrHbO in *Frankia*

The potential -35 and -10 promoter regions that are highlighted in Figure 10 are simply possible consensus sequences for the promoter of *Frankia* trHbs. These sequences were chosen solely on their conservation across *Frankia* trHb subgroups and relative distances from each other upstream of the start codon. There are 2 possible sets of conserved promoter regions present in the trHbN upstream regions. None of the sequences match the -35 and -10 consensus sequences TTG(T/A)CG and TA(G/A)(G/A)T, respectively, that have been described previously for *Frankia* genes (Cournoyer & Normand, 1994; Lavire & Cournoyer, 2003). Whether or not these sites are actual -35 and -10 promoter regions is still to be determined.

The putative ribosomal binding sites (RBS) that were identified upstream of the start codons for each of the trHbs were typical of RBS found previously in *Frankia* (Fig. 11) (Cournoyer & Normand, 1994; Lavire & Cournoyer, 2003). The

sequences of the putative RBS shown in Figure 11 are mostly GAGA for the *Frankia* trHbs and are quite similar to the GGAG consensus sequence suggested previously for *Frankia* RBS (Rochefort & Benson, 1990). Without further experimentation I am unsure if these are true RBS for these genes.

The Effect of Environmental Stimuli on TrHb Expression

The basis of the expression studies I performed for trHbs in *Frankia* stemmed from preliminary biochemical work performed by our collaborators (Beckwith et al., 2002; Tjepkema et al., 2002; Schwintzer & Tjepkema, 2005). The majority of the research they performed relied on the extraction of total heme protein. At the time of their studies it was not known that *Frankia* expressed 2 trHb genes. By evaluating total heme protein that was being isolated simultaneously from each of the two trHbs, the up-regulation or down-regulation of one trHb may have been potentially masked by the presence of the other (and *vice versa*). Any direct correlations between the observations I have noted and their results should be compared with caution.

The Triton X-100 Boiling RNA Extraction Method

The Triton X-100 Boiling Method described by Sung et al. (2003) proved to be a quick, efficient, and inexpensive method for extracting RNA from *Frankia* hyphae. The yield of total RNA extracted from frozen cells was nearly twice that

of the commercially available method and was much cleaner (Table 4). Not only did this method yield better quality mRNA, but also extracted RNA that was free from contaminating 16S and 23S rRNA (Fig. 12). Typically, these stable RNA species make up 70% of an RNA prep when using older methods (Neidhardt & Umbarger, 1996). An RNA extraction method that provides a high yield of quality RNA is of great interest to future transcriptional analysis studies for *Frankia*. Being able to extract a large amount of mRNA can be useful for microarray studies. Until, a more superior protocol is discovered, this should be the method of choice for the extraction of mRNA from *Frankia*.

Nitrogen Limitation Does Not Affect TrHb Expression in *Frankia*

Both the RT-PCR and Real-time RT-PCR data support the conclusion that neither of these Hbs function in nitrogen fixation. The bands for trHbN and trHbO RT-PCR amplicons (Fig. 14) were equal in intensity and the transcription analysis (Fig. 17) revealed a nearly equal expression of trHbN and trHbO under each condition. This supports the earlier observations of Beckwith et al. (2002) who isolated comparable levels of hemoglobin from both N+ and N- cultures of EAN1pec, Arl3, EUN1f, Cc1.17, and Ccl3 strains.

A role of *Frankia* trHbs in nitrogen fixation would not have been surprising. An initial instinct would be to assume the role of truncated hemoglobin in the *Frankia*-Actinorhizal symbiosis to be analogous to the function of leghemoglobin to protect nitrogenase in *Rhizobium*-legume symbiosis by maintaining low oxygen tension within the legume nodule (Gallon et al., 1992). *Frankia*, however,

has evolved to localize nitrogen fixation within vesicles. These specialized structures are multi-laminated with hopanoid lipids to restrict oxygen diffusion into the cell (Berry et al., 1993). In this sense, a role for trHbs in nitrogenase protection would be redundant. Interestingly, the cyanoglobin (GlbO) trHb of the nitrogen-fixing cyanobacterium *Nostoc commune* is believed to function in nitrogen fixation (Hill et al., 1996). Not only is this gene located within the intergenic region of two nitrogen-fixing genes (*nif*U and *nif*H), but hemoglobin synthesis increases when *Nostoc* is grown without a combined nitrogen source. Also, immediately upstream of *glb*N in *Nostoc* is a binding site for NtcA, a well-known transcription factor involved with other nitrogen-regulated genes. It should also be noted that no other nitrogen-fixing related genes were observed in either of the trHb gene neighborhoods for any of the *Frankia* strains.

Oxidative Stress Does Not Affect TrHb Expression in *Frankia*

Transcriptional analysis from Ccl3 cultures that were challenged with paraquat and H_2O_2 revealed no significant up regulation of expression for either trHb in comparison to control cultures (Fig. 17). Cultures that were grown with hydrogen peroxide only exhibited a slight increase in expression for both trHbN and trHbO. This could simply be explained by a relative increase in trHb expression as the cells approached stationary phase. Schwintzer & Tjepkema (2005) had observed an increase in Hb production of *Frankia* at stationary phase. This complements earlier findings of HbN expression in *Mycobacterium* that reaches its maximum when cultures are growing in stationary phase (Ouellett et

al., 2002; Pathania et al., 2002). Why the genes for trHbN and trHbO in *Frankia* were down regulated by nearly a factor of 10 when paraquat was added to cultures is unknown. One explanation for this decrease in expression could be due to metabolism of paraquat that leads to intracellular cytotoxic levels of H_2O_2.

The host plants of the *Frankia*-actinorizal symbiosis not only use reactive oxygen species (ROS) and reactive oxygen intermediates (ROI) for defense against pathogenic microbes, but these molecules are also used for signal transduction, regulating developmental processes, and programmed cell death (Tavares et al., 2007). Thus, the bacterial partner of this symbiosis requires defense against ROS and ROI upon the initial infection of the host and also while persisting in the root nodule. ROS and ROI produced by the host plant are tolerated in *Frankia* by means of catalases and superoxide dismutases (Hammad et al., 2001; Tavares et al., 2003; Santos et al., 2007). The role of trHb in *Frankia* to protect the microbe from reactive oxygen species is not likely, nor do these genes seem to be globally regulated in response to oxidative stress.

TrHbO Expression in *Frankia* Increases Under Low O_2 Environments

Schwintzer et al. (2005) evaluated the effect of oxygen on hemoglobin concentrations in *Frankia* strain Arl3. They found an increase in hemoglobin concentration in cultures grown under 1% oxygen in media with a combined nitrogen source, as compared to cultures grown under 20% and 40% oxygen. Previous studies by Beckwith et al. (2002) found a similar increase in hemoglobin production when cultures of EAN1pec were grown in 2% oxygen in comparison

to 20% oxygen; however, this was observed in cultures not supplemented with NH_4Cl. It is unclear why this difference was observed between these two strains though in either case it was clear that a low oxygen environment caused an increase in oxygen expression. In my RT-PCR experiments, NH_4Cl had been added to the media of both the aerated and static cultures. Amplification of trHbO transcripts from these cultures showed a relative increase of expression for this gene (Figure 15) in the static cultures where available oxygen was limited in comparison to the aerated cultures. Our findings were consistent with those of Schwintzer et al. (2005). These conditions had no effect on trHbN expression as observed in the band intensities that were relatively equal (Figure 15).

As described previously, this increase in trHbO expression in response to hypoxia has been reported in *Mycobacterium* sp. (Pathania et al. 2002a; Liu et al., 2004). Not only did these studies show that cultures of recombinant *E. coli* expressing the *glb*O gene from *M. tuberculosis* show greater oxygen uptake than control cultures, they also showed that the function of *glb*O in these recombinant cells was dependent on the delivery of oxygen to the membrane-associated terminal oxidase, cytochrome *o*. In other words, the hemoglobin acts as an oxygen shuttle to cytochromes to increase respiration rates, which is crucial in low oxygen environments. Consistent with this, Tjepkema et al. (2002) observed rapid oxygen kinetics for *Frankia* hemoglobin in strain CcI3 suggesting a role of oxygen transport over short distances.

It is also believed that a series of hydrophobic residues present on the molecule surface of HbO is responsible for the interaction with acidic

phospholipids at the cell membrane to allow for the delivery of oxygen to membrane-associated cytochromes (Pathania et al. 2002a; Liu et al., 2004). These same residues (mostly arginine) are conserved in the known actinomycete trHbO sequences and the *Frankia* trHbO sequences are no exception (Figure 5). The presence of these residues in *Frankia* would also support a similar function in oxygen delivery as described for *Mycobacterium* sp. This could be a useful for *Frankia* growing in any environment of low oxygen tension. Of particular interest is the strain that infects *Myrica gale* where zones of low oxygen in the nodules of this plant have been reported (Tjepkema, 1983). As stated prior, hemoglobin corresponding to the molecular weight of trHbs has been isolated from *Myrica gale* nodules (Pathirana & Tjepkema, 1995). This hemoglobin was most likely being produced by the bacterial partner.

TrHbN Expression in *Frankia* Increases Under Nitrosative Stress

For both the RT-PCR and Real-time RT-PCR experiments we observed an increase in the expression of trHbN from Ccl3 in response to a spontaneous nitric oxide donor being added to the culture medium (Figs. 16 and 17). The band intensity for the RT-PCR experiment showed the highest relative expression of trHbN at 3 hours after the addition 400uM SNAP to the medium. For the transcriptional analysis I observed a 10-fold induction of trHbN expression when 400uM SNAP was added in comparison to control cultures (Fig. 17). A nitric oxide scavenger (400uM cPT-10) was also added along with the SNAP to show that the expression of trHbN would go down in comparison to

SNAP alone cultures. The expression of trHbN was decreased with the addition cPT-10 + SNAP, however, there was still about a 4-fold induction of trHbN for these cultures in comparison to the controls (Fig. 17). This may have been due to inefficient NO scavenging by cPT-10. By using a higher concentration of this compound a reduction in trHbN transcript may have been observed in the cPT-10 + SNAP cultures. The expression of trHbO in either of these cultures did not seem to be affected. A role for *Frankia* trHbN in the detoxification of nitric oxide is plausible.

The oxygenated form of HbN for *M. tuberculosis* is believed to aid in its defense from NO produced by macrophages in infected lung tissue (Couture et al., 1999; Milani et al., 2001b; Ouellett et al., 2002; Pathania et al., 2002b). The HbN hemoglobin functions as an oxygen-dependent nitric oxide dioxygenase using bound oxygen to react with a nitric oxide molecule to form nitrate. Pathania et al. (2002b) showed that oxygen was required for NO detoxification and recombinant *E. coli* possessing the HbN gene was able to grow in the presence of NO, where as the growth of control cultures was reduced. Similarly, HbN mutants of *Mycobacterium bovis* were unable to grow in the presence of NO in comparison to the wild-type (Ouellett et al., 2002).

The heme-ligand tunnel of *Mycobacterium* HbN proteins is lined with several hydrophobic residues that are well suited for the diffusion of oxygen and nitric oxide toward the heme distal pocket (Milani et al., 2001b; Ouellett et al., 2002). This same tunnel may also function to store multiple nitric oxide molecules for more efficient catalytic activity. These hydrophobic residues are

conserved in *Frankia* trHbN sequences compared to HbN of *M. tuberculosis* (Fig. 6). From a structural standpoint, it would make sense that a similar role for *Frankia* trHbN might be conserved.

In addition to ROS, plants also use reactive nitrogen species (RNS) such as nitric oxide as key regulatory molecules (Neill et al., 2003; Tavares et al., 2007). There are many biological processes such as signal transduction and the hypersensitivity response that are regulated by NO in plants. Thus, it is certain that this molecule is present in the infected actinorhizal nodule where a great deal of differentiation occurs for the plant. It can then be understood why *Frankia* would have evolved a means to cope with this cytotoxic compound. The plant hosts of *Frankia*-actinorhizhal symbiosis also evolved a means by which to cope with cytotoxic levels of internal NO. The class I non-symbiotic hemoglobin (AfHb1) from *Alnus firma* is believed to function in the detoxification of nitric oxide in both symbiotic and non-symbiotic plant tissue (Sasakura et al., 2006).

Currently, there are more unpublished biochemical studies being performed by our collaborators that also confirm my findings (Tjepkema, 2007). Using chromatography they have been able to isolate and measure the amounts of the two truncated hemoglobins from *Frankia* cultures grown under various conditions. When cultures were grown in the presence of a nitric oxide donor they measured an increase in trHbN production. When cultures were grown in a reduced oxygen environment they measured an increase in trHbO production. An increase in hemoglobin protein in either of these examples is further support that more protein is being translated in response to their respective

environmental stimulus. These recent studies by Tjepkema (2007) correlate well with my earlier observations.

The lack of reliable genetic transfer systems and mutagenesis protocols in *Frankia* research is a major drawback to genetic studies. Having a definitive method for developing knockout phenotypes for genes of interest is a must to prove causality in gene expression studies. Though *Frankia* genetics was once considered recently to be in its infancy, the sequencing of three *Frankia* genomes will reveal an enormous amount of information and the advent of more genetic tools will have this field well on its way to "adolescence".

LIST OF REFERENCES

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

Ascenzi, P., A. Bocedi, M. Bolognesi, G. Fabozzi, M. Milani, and P. Visca. 2006. Nitric oxide scavenging by *Mycobacterium leprae* GlbO involves the formation of the ferric heme-bound peroxynitrite intermediate. Biochem. Biophys. Res. Comm. **339:**450-456.

Baker, D., Newcomb, W., and Torrey, J.G. 1980. Characterization of an ineffective actinorhizal microsymbiont, *Frankia* sp. Eul1 (Actinomycetales). Can. J. Microbiol. **26:**1072–1089.

Beckwith, J., J.D. Tjepkema, R.E. Cashon, C.R. Schwintzer, and L.S. Tisa. 2002. Hemoglobin in five genetically diverse *Frankia* strains. Can. J. Microbiol. **48:**1048-1055.

Benson, D.R. and Clawson, M.L. 2000. Evolution of the actinorhizal plant symbiosis. In: Prokaryotic nitrogen fixation: a model system for analysis of a biological process. Triplett, E.W., ed. Horizon Scientific Press, Norfolk, United Kingdom, pp. 207–224.

Benson, D.R. and Silvester, W.B. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiol. Rev. **57:**293–319.

Berry, A.M., O.T. Harriot, R.A. Moreau, S.F. Osman, D.R. Benson, and A.D. Jones. 1993. Hopanoid lipids compose the *Frankia* vesicle envelope, presumptive barrier of oxygen diffusion to nitrogenase. PNAS. **90:**6091-6094.

Bonamore, A., A. Ilari, L. Giangiacomo, A. Bellini, V. Morea, and A. Boffi. 2005. A novel thermostable hemoglobin from the actinobacterium *Thermobifida fusca*. FEBS Journal. 272:4189-4201.

Callaham, D., Del Tridici, P., and Torrey, J.G. 1978. Isolation and cultivation in vitro of the actinomycete causing root nodulation in *Comptonia*. Science. **199:**899–902.

Choudhary, M.L., S. Jawaid, M.K. Ahuja, N.K. Shiva, P. Gupta, A.K. Bhuyan, and G.S. Khatri. 2005. Open reading frame *yjbI* of *Bacillus subtilis* codes for truncated hemoglobin. Prot. Express. Purific. **41:**363-372.

Clawson, M.L., A. Bourret, and D.R. Benson. 2004. Assessing the phylogeny of *Frankia*-actinorhizal plant nitrogen-fixing root nodule symbioses with *Frankia* 16S rRNA and glutamine synthetase gene sequences. Mol Phylogenet Evol. **31:**131-138.

Cournoyer, B. and P. Normand. 1994. Gene expression in *Frankia*: characterization of promoters. Microbios. **78:**229-236.

Couture, M., S. Yeh, B.A. Wittenberg, J.B. Wittenberg, Y. Ouellet, D.L. Rousseau, and M. Guertin. 1999a. A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis*. PNAS. **96:**11223-11228.

Couture, M., T.K. Das, H.C. Lee, J. Peisach, D.L. Rousseau, B.A. Wittenberg, J.B. Wittenberg, and M. Guertin. 1999b. *Chlamydomonas* chloroplast ferrous hemoglobin. J. Biol. Chem. 274:6898-6910.

Couture, M., T.K. Das, P.Y. Savard, Y. Ouellet, J.B. Wittenberg, B.A. Wittenberg, D.L. Rousseau, and M. Guertin. 2000. Structural investigations of the hemoglobin of the cyanobacterium *Synechocystis* PCC6803 reveal a unique distal heme pocket. Eur. J. Biochem. **267:**4770-4780.

Das, T.K., R,E. Weber, S, Dewilde, J.B. Wittenberg, B.A. Wittenberg, K. Yamauchi, M.L. Van Hauwaert, L. Moens, and D.L. Rousseau. 2001. Ligand binding in the ferric and ferrous states of *Paramecium* hemoglobin. Biochem. **39:**14330-14340.

Gallon, J.R. 1992. Tansley review no. 44, reconciling the incompatible: N_2 fixation and O_2. New Phytol. **122:**571-609.

Hammad, Y., J. Marechal, B. Cournoyer, P. Normand, and A.M. Domenach. 2001. Modification of the protein expression pattern induced in the nitrogen-fixing actinomycete *Frankia* sp. strain ACN14a-trs by root exudates of its symbiotic host *Alnus glutinosa* and cloning of the *sodF* gene. Can. J. Microbiol. **47:**541-547.

Hill, D.R., T.J. Belbin, M.V. Thorsteinsson, D. Bassam, S. Brass, A. Ernst, P. Böger, H. Paerl, M.E. Mulligan, and M. Potts. 1996. GlbN (cyanoglobin) is a peripheral membrane protein that is restricted to certain *Nostoc* spp. J. Bacteriol. **178:**6587-6598.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M., and Schrempf, H. 1985. Genetic Manipulation of *Streptomyces*: A Laboratory Manual. John Innes Foundation, Norwich, England.

Huss-Danell, K. 1997. Tansley review no. 93, actinorhizal symbioses and their N_2 fixation. New Phytol. **136:**375-405.

Jacobsen-Lyon, K., E.O. Jensen, J-E. Jorgensen, K.A. Marcker, W.J. Peacock, and E.S. Dennis. 1995. Symbiotic and nonsymbiotic hemoglobin genes of *Casuarina glauca*. Plant Cell. 7:213-223.

John, T.R., J.M. Rice, and J.D. Johnson. 2001. Analysis of pFQ12, a 22.4-kb *Frankia* plasmid. Can. J. Microbiol. **47**:608-617.

Ishikawa, J. and Hotta, K. 1999. Frameplot: a new implementation of the frame analysis for predicting protein coding regions in bacterial DNA with a high G+C content. FEMS Microbiol. Lett. **174:**251–253.

Korenaga, S., J. Igarashi, A. Matsuoka, and K. Shikama. 2000. A primitive myoglobin from *Tetrahymena pyriformis*: its heme environment, autoxidizability, and genomic DNA structure. Biochim. Biophys. Acta. **1543:**131-145.

Kortt, A.A., A.S. Inglis, A.I. Fleming, and C.A. Appleby. 1988. Amino acid sequence of hemoglobin I from root nodules of the non-leguminous *Casuarina glauca-Frankia* symbiosis. FEBS Lett. **231:**341-346.

Lalonde, M., Calvert, H.E., and Pine, S. 1981. Isolation and use of *Frankia* strains in actinorhizae formation. In: Current Perspectives in Nitrogen Fixation. Gibson, A.H. and Newton, W.E., eds. Australia Academy of Science, Canberra, Australia, pp. 296–299.

Lavire, C. and B. Cournoyer. 2003. Progress on the genetics of the N_2-fixing actinorhizal symbiont *Frankia*. Plant Soil. **254:**125-137

Liu, C., Y. He, and Z. Chang. 2004. Truncated hemoglobin *o* of *Mycobacterium tuberculosis*: oligomeric state change and the interaction with membrane components. Biochem. Biophys. Res. Comm. **316:**1163-1172.

Meesters, T.M., van Genesen, S.T., and Akkermans, A.D.L. 1985. Growth, acetylene reduction activity and localization of nitrogenase in relation to vesicle formation in *Frankia* strains Cc11.7 and Cp1.2. Arch. Microbiol. **143:**137–142.

Milani, M., Pesce, A., Bolognesi, M., and Ascenzi, P. 2001a. Truncated hemoglobins: trimming the classical 'three-over-three' globin fold to a minimum size. Biochem. Mol. Biol. Educ. **29:**123–125.

Milani, M., A. Pesce, Y. Ouellet, P. Ascenzi, M. Guertin, and M. Bolognesi. 2001b. *Mycobacterium tuberculosis* hemoglobin N displays a protein tunnel suited for O_2 diffusion to the heme. EMBO Journal. **20:**3902-3909.

Milani, M., P. Savard, H. Ouellet, P. Ascenzi, M. Guertin, and M. Bolognesi. 2003. A TyrCD1/TrpG8 hydrogen bond network and a TyrB10-TyrCD1 covalent link shape the heme distal site of *Mycobacterium tuberculosis* hemoglobin O. PNAS. **100:**5766-5771.

Milani, M., A. Pesce, Y. Ouellet, S. Dewilde, J. Friedman, P. Ascenzi, M. Guertin, and M. Bolognesi. 2004. Heme-ligand tunneling in group I truncated hemoglobins. J. Biol. Chem. **279:**21520-21525.

Milani, M., A. Pesce, M. Nardini, H. Ouellet, Y. Ouellet, S. Dewilde, A. Bocedi, P. Ascenzi, M. Guertin, L. Moens, J.M. Friedman, J.B. Wittenberg, and M. Bolognesi. 2005. Structural bases for heme ligand binding and diatomic ligand recognition in truncated hemoglobins. J. Inorg. Biochem. **99:**97-109

Mirza, M.S., Akkermans, W.M., and Akkermans, A.D.L. 1994. PCR-amplified 16S rRNA sequence analysis to confirm nodulation of *Datisca cannabina* L. by the endophyte of *Coriaria nepalensis* Wall. Plant Soil. **160:**147–152.

Neidhardt, F.C., and H.E. Umbarger. 1996. *Escherichia coli* and *Salmonella* (Neidhardt, F.C., Ed.), Vol 1, pp. 13-16. ASM Press, Washington, DC.

Neill, S.J., R. Desikan, and J.T. Hancock. 2003. Tansley review, nitric oxide signaling in plants. New Phytol. **159:**11-35.

Niemann, J.M., J.D. Tjepkema, and L.S. Tisa. 2005. Identification of the truncated hemoglobin gene in *Frankia*. Symbiosis. **39:**91-95

Normand, P., Orso, S., Cournoyer, B., Jeanin, P., Chapelon, C., Dawson, J., Evtushenko, L., and Mirsa, A.K. 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family Frankiaceae. Int. J. Syst. Bacteriol. **46:**1-9.

Normand, P., P. Lapierre, L.S. Tisa, J.P. Gogarten, N. Alloisio, E. Bagnarol, C.A. Bassi, A.M. Berry, D.M. Bickhart, N. Choisne, A. Couloux, B. Cournoyer, S. Cruveiller, V. Daubin, N. Demange, M.P. Francino, E. Goltsman, Y. Huang, O.R. Kopp, L. Labarre, L. Lapidus, C. Lavire, J. Marechal, M. Martinez, J.E. Mastronunzio, B.C. Mullin, J. Niemann, P. Pujic, T. Rawnsley, Z. Rouy, C. Schenowitz, A. Sellstedt, F. Tavares, J.P. Tomkins, D. Vallenet, C. Valverde, L.G. Wall, Y. Wang, C. Medigue, and D.R. Benson. 2007a. Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. Genome Research. **17:**7-15

Normand, P., C. Queiroux, L.S. Tisa, D.R. Benson, Z. Rouy, S. Cruveiller, and C. Medigue. 2007b. Exploring the genomes of *Frankia*. Physiol. Planta. **130:**331-343.

Ouellet, H., Y. Ouellet, C. Richard, M. Labarre, B. Wittenberg, J. Wittenberg, and M. Guertin. 2002. Truncated hemoglobin HbN protects *Mycobacterium bovis* from nitric oxide. PNAS. **99:**5902-5907.

Pathania, R., N.K. Navani, G. Rajamohan, and K.L. Dikshit. 2002a. *Mycobacterium tuberculosis* hemoglobin HbO associates with membranes and stimulates cellular respiration of recombinant *Escherichia coli*. J. Biol. Chem. **277:**15293-15302.

Pathania, R., N.K. Navani, A.M. Gardner, P.R. Gardner, and K.L. Dikshit. 2002b. Nitric oxide scavenging and detoxification by the *Mycobacterium tuberculosis* haemoglobin, HbN in *Escherichia coli*. Mol. Micro. **45:**1303-1314

Pathirana, S.M., and J.D. Tjepkema. 1995. Purification of hemoglobin from the actinorhizal root nodules of *Myrica gale* L. Plant Physiol. **107:**827-831.

Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Acsenzi, P., Moens, L., and Bolognesi, M. 2000. A novel two-over-two a-helical sandwich fold is characteristic of the truncated hemoglobin family. EMBO Journal. **19:**2424–2434.

Richards, J.W., G.D. Krumholz, M.S. Chval, and L.S. Tisa. 2002. Heavy metal resistance patterns of *Frankia* strain. Appl. Environ. Microbiol. 68:923-927.

Rochefort, D.A. and D.R. Benson. 1990. Molecular cloning, sequencing, and expression of the glutamine synthase II (glnII) gene from the actinomycete root nodule symbiont *Frankia* sp. strain Cpl1. J. Bacteriol. **172:**5335-42

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Agarose gel electrophoresis In: Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 6.3-6.19.

Santos, C.L., J. Vieira, A. Sellstedt, P. Normand, P. Moradas-Ferreira, and F. Tavares. 2007. Modulation of *Frankia alni* ACN14a oxidative stress response: activity, expression and phylogeny of catalases. Physiol. Planta. **130:**454-463.

Sarma, H.K., B.K. Sharma, S.C. Tiwari, and A.K. Mishra. 2005. Truncated hemoglobins: a single structural motif with versatile functions in bacteria, plants, and unicellular eukaryotes. Symbiosis. 39:151-158.

Sasakura, F., T. Uchiumi, Y. Shimoda, A. Suzuki, K. Takenouchi, S. Higashi, and M. Abe. 2006. A class I hemoglobin gene from *Alnus firma* functions in symbiotic and nonsymbiotic tissues to detoxify nitric oxide. Mol. Plant-Micro. Interac. **19:**441-450.

Schwencke, J., and M. Carú. 2001. Advances in actinorhizal symbiosis: host plant-*Frankia* interactions, biology, and applications in arid land reclamation. A review. Arid Land Manage. **15:**285-327.

Schwintzer, C.R. and J.D. Tjepkema. 2005. Effect of oxygen concentration on growth and hemoglobin production in *Frankia*. Symbiosis. **39:**77-82.

Scott, N.L., J.T.J. Lecomte. 2000. Cloning, expression, purification, and preliminary characterization of a putative hemoglobin from the cyanobacterium *Synechosystis* sp. PCC 6803. Prot. Sci. **9:**587-597.

Suharjo, U.K.J., and J.D. Tjepkema. 1995. Occurrence of hemoglobin in the nitrogen-fixing root nodules of *Alnus glutinosa*. Physiol. Planta. **95:**247-252.

Sung, K., S.A. Khan, M.S. Nawaz, and A.A. Khan. 2003. A simple and efficient Triton X-100 boiling and chloroform extraction method of RNA isolation from gram-positive and gram-negative bacteria. FEMS Microbiol. Lett. **229:**97-101.

Swofford, D.L. 2003. PAUP: Phylogenetic analysis using parsimony (and other methods). Version 4.0b10. Sinauer Associates (Sunderland, MA).

Tavares, F., L. Bernardo, and A. Sellstedt. 2003. Identification and expression studies of a catalase and a bifunctional catalase-peroxidase in *Frankia* strain R43. Plant Soil. **254:**75-81.

Tavares, F., C.L. Santos, and A. Sellstedt. 2007. Reactive oxygen species in legume and actinorhizal symbioses: the microsymbiont's responses to an unfriendly reception. Physiol. Planta. **130:**344-356

Tjepkema, J.D. 1982. Hemoglobins in the nitrogen-fixing root nodules of actinorhizal plants. Can. J. Bot. **61:**2924-2929.

Tkepkema, J.D. 1983. Oxygen concentration within the nitrogen-fixing root nodules of *Myrica gale* L. Am. J. Bot. **70:**59-63

Tjepkema, J.D., and D.J. Asa. 1987. Total and CO-reactive heme content of actinorhizal nodules and the roots of some non-nodulated plants. Plant Soil. **100:**225-236.

Tjepkema, J.D., R.E. Cashon, J. Beckwith, and C.R. Schwintzer. 2002. Hemoglobin in *Frankia*, a nitrogen-fixing actinomycete. Appl. Environ. Microbiol. **68:**2629-2631.

Tjepkema, J.D. 2002. Direct communication.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876-4882.

Tisa, L.S., M. McBride, and J.C. Ensign. 1983. Studies of growth and morphology of *Frankia* strains EAN1pec, Eul1c, Cpl1, and ACN1AG. Can. J. Bot.. 61:2768–2773.

Tisa, L.S., M.S. Chval, G.D. Krumholz, and J. Richards. 1999. Antibiotic resistance patterns of *Frankia* strains. Can. J. of Bot.. **77:**1257-1260.

Vessey, J.K., K. Pawlowski, and B. Bergman. 2005. Root-based N_2-fixing symbioses: legumes, actinorhizal plants, *Parasponia* sp. and cycads. Plant Soil. **274:**51-78.

Vieweg, M.F., N. Hohnjec, and H. Kuster. 2005. Two genes encoding different truncated hemoglobins are regulated during root nodule and arbuscular mycorrhiza symbioses of *Medicago truncatula*. Planta. **220:**757-766.

Visca, P., G. Fabozzi, M. Milani, M. Bolognesi, and P. Ascenzi. 2002. Nitric oxide and *Mycobacterium leprae* pathogenicity. Life. **54:**95-99.

Visca, P., G. Fabozzi, A. Petrucca, C. Ciaccio, M. Coletta, G. De Sanctis, M. Bolognesi, M. Milani, and P. Ascenzi. 2002. The truncated hemoglobin from *Mycobacterium leprae*. Biochem. Biophys. Res. Comm. **294:**1064-1070.

Vuletich, D.A., and J.T.J. Lecomte. 2006. A phylogenetic and structural analysis of truncated hemoglobins. J. Mol. Evol. **62:**196-210.

Wainwright, L.M., K.T. Elvers, S.F. Park, and R.K. Poole. 2005. A truncated haemoglobin implicated in oxygen metabolism by the microaerophilic food-borne pathogen *Campylobacter jejuni*. Microbiol. **151:**4079-4091.

Wall, L.G. 2000. The actinorhizal symbiosis. J. Plant Growth Regul. **19:**167-182.

Watts, R.A., P.W. Hunt, A.N. Hvitved, M.S. Hargrove, W.J. Peacock, and E.S. Dennis. 2001. A hemoglobin from plants homologous to truncated hemoglobins of microorganisms. PNAS. **98:**10119-10124.

Wilson, K. 1989. Preparation of genomic DNA from bacteria. In: Current Protocols in Molecular Biology. Volume 1. Ausubel, F.M. *et al.*, eds. Greene Publishing Associate & Wiley Interscience, New York, NY, pp. 2.4.1–2.4.5.

Wittenberg, J.B., M. Bolognesi, B.A. Wittenberg, and M. Guertin. 2002. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. J. Biol. Chem. **277:**871-874.

Zhang, Z.M., F. Lopez, and J.G. Torrey. 1994. A comparison of cultural characteristics and infectivity of *Frankia* isolates from root nodules of *Casuarina* species. Plant Soil. **78:**79-90.